Edit Distance Based Instruction Merging Technique to Improve Flexibility of Custom Instructions Toward *Flexible Accelerator* Design

Hui Huang*, **Taemin Kim**** and Yatin Hoskote**

UCLA* and Intel**

Outline

- Introduction
- Related Works
- Hypotheses and Verification
- Our Approach
- Experimental Results
- Conclusion

Introduction

- Flexible Accelerator
  - To respond ever decreasing time-to-market
  - To reduce NRE cost
- Traditional ASIP
  - Flexibility through base instruction-set
  - Performance through custom instructions
- Our Proposal
  - Make custom instructions more flexible

Contributions

- Analyzed how similar applications in the same domain are one another
- Used *edit distance* as a flexibility metric
- Proposed an algorithm to make custom instructions flexible

Related Works

- Adding Redundancy *Randomly*
  - Interconnect, Functional Unit and Storage
  - Loop Accelerator[1]
    - Global Bus, Register File and ALU without guidance
  - Flexible CGRA[2]
    - Decompose a computational pattern (CP) into a smaller one and then add redundant interconnects to support more CPs

Hypotheses

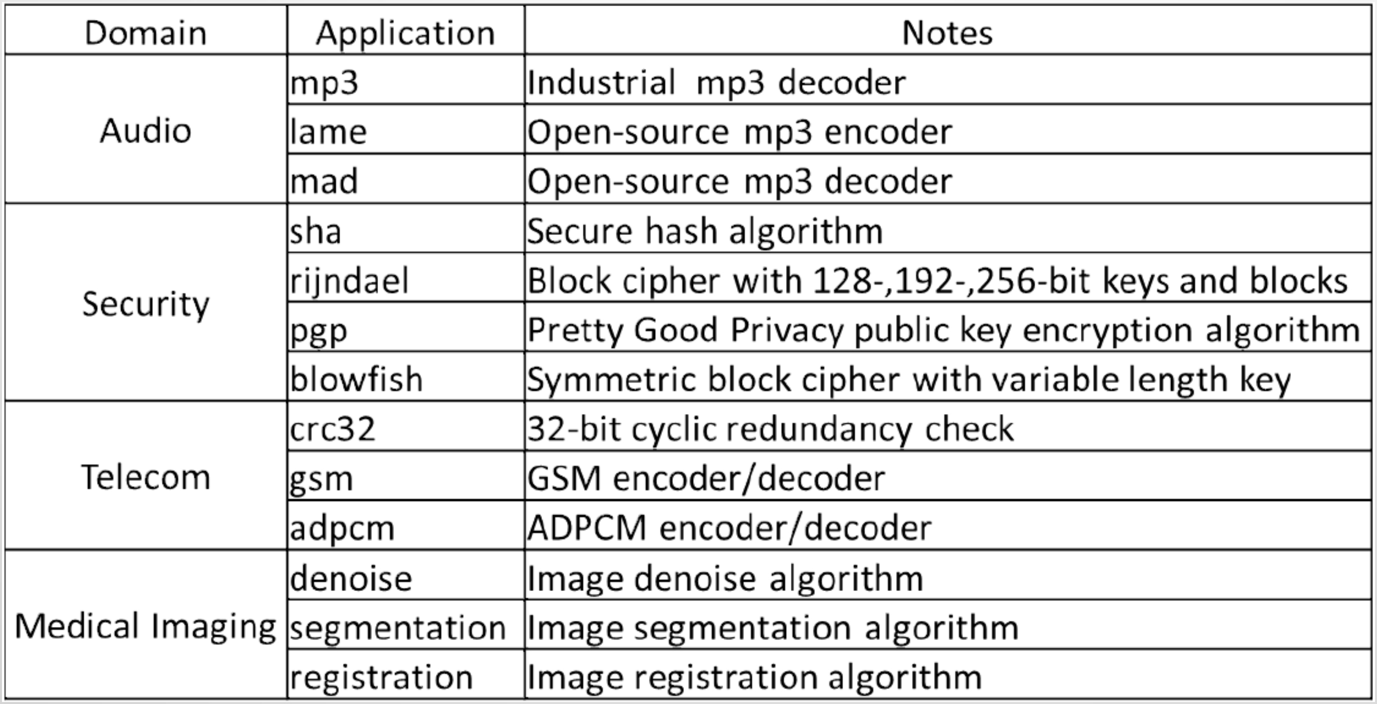
- Significant Similarity in the Same Domain*
  - Exactly same computational patterns (CPs) **(H1)**
  - Small difference **(H2)**
  - Coverage increases for one app → Same for another in the same domain **(H3)**

***Domain : Application functionality level terminology (e.g. audio, image, cryptography)**

Verification

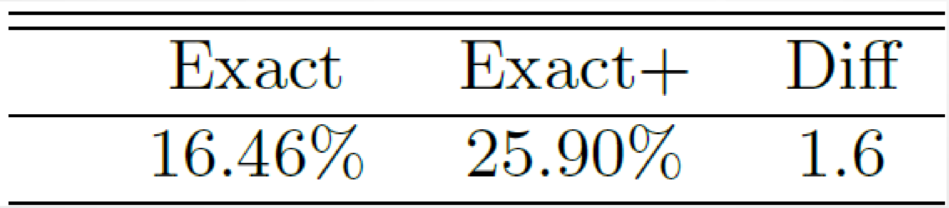
- Methodology
  - Define a current application (CA) and a future application (FA)
  - Generate custom instructions (CIs) for CA and FA separately. (CICA and CIFA)
  - **Exact** : Compare CICA to CIFA to see the difference (**H1** and **H2**)
  - **Exact+** : When |CICA| increases, see how many elements of CIFA are same as those of CICA (**H3**)

Verification (Cont'd)



Domain	Application	Notes
Audio	mp3	Industrial mp3 decoder
	lame	Open-source mp3 encoder
	mad	Open-source mp3 decoder
Security	sha	Secure hash algorithm
	rijndael	Block cipher with 128-,192-,256-bit keys and blocks
	pgp	Pretty Good Privacy public key encryption algorithm
	blowfish	Symmetric block cipher with variable length key
Telecom	crc32	32-bit cyclic redundancy check
	gsm	GSM encoder/decoder
	adpcm	ADPCM encoder/decoder
Medical Imaging	denoise	Image denoise algorithm
	segmentation	Image segmentation algorithm
	registration	Image registration algorithm

Verification Result

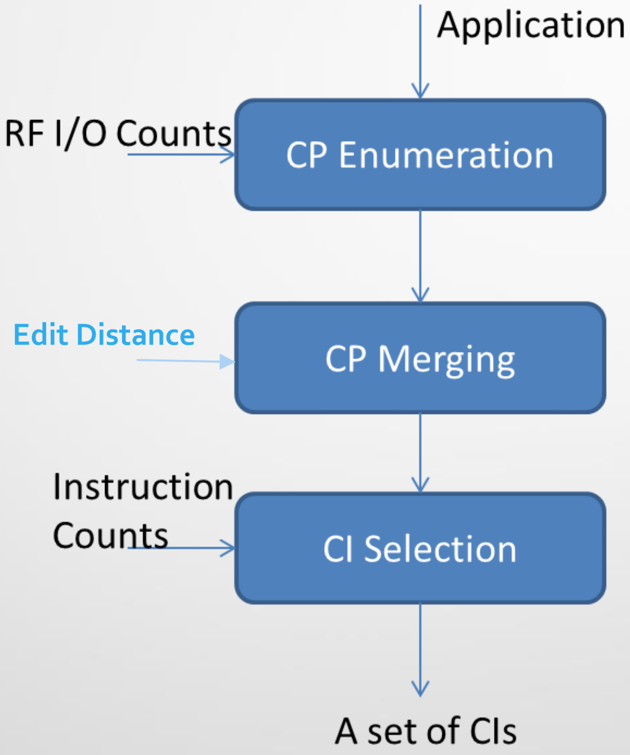


Exact	Exact+	Diff
16.46%	25.90%	1.6

Our Approach

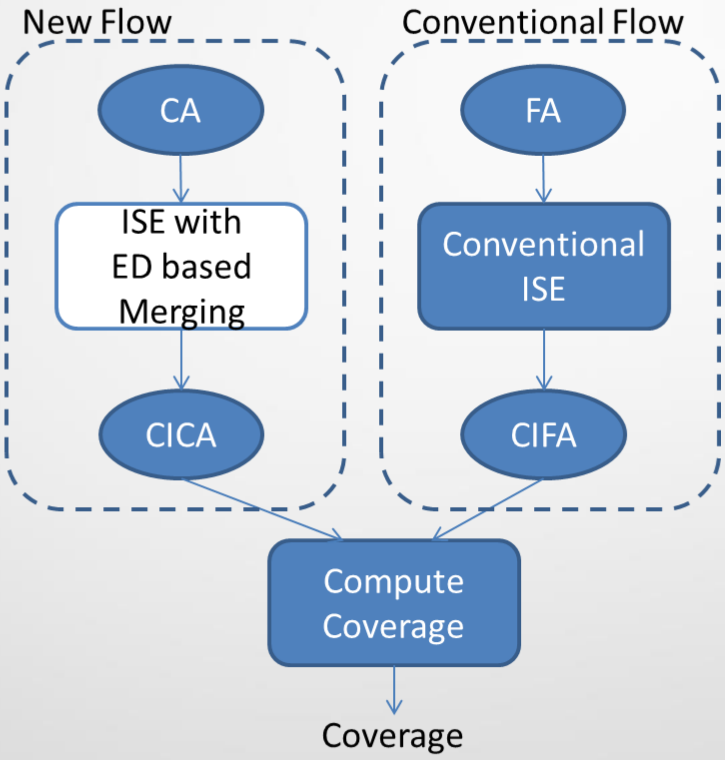
- Objective
  - Maximize the number of computational patterns (CPs) that custom instructions (CIs) can support
- Constraint
  - The number of custom instructions
- How to support multiple CPs with a single CI
  - Allow small difference among CPs
  - Use *Edit Distance* as a difference metric
  - Relax the constraint of *CP Merging* step
    - Merge *similar* CPs into a single CI

Flow

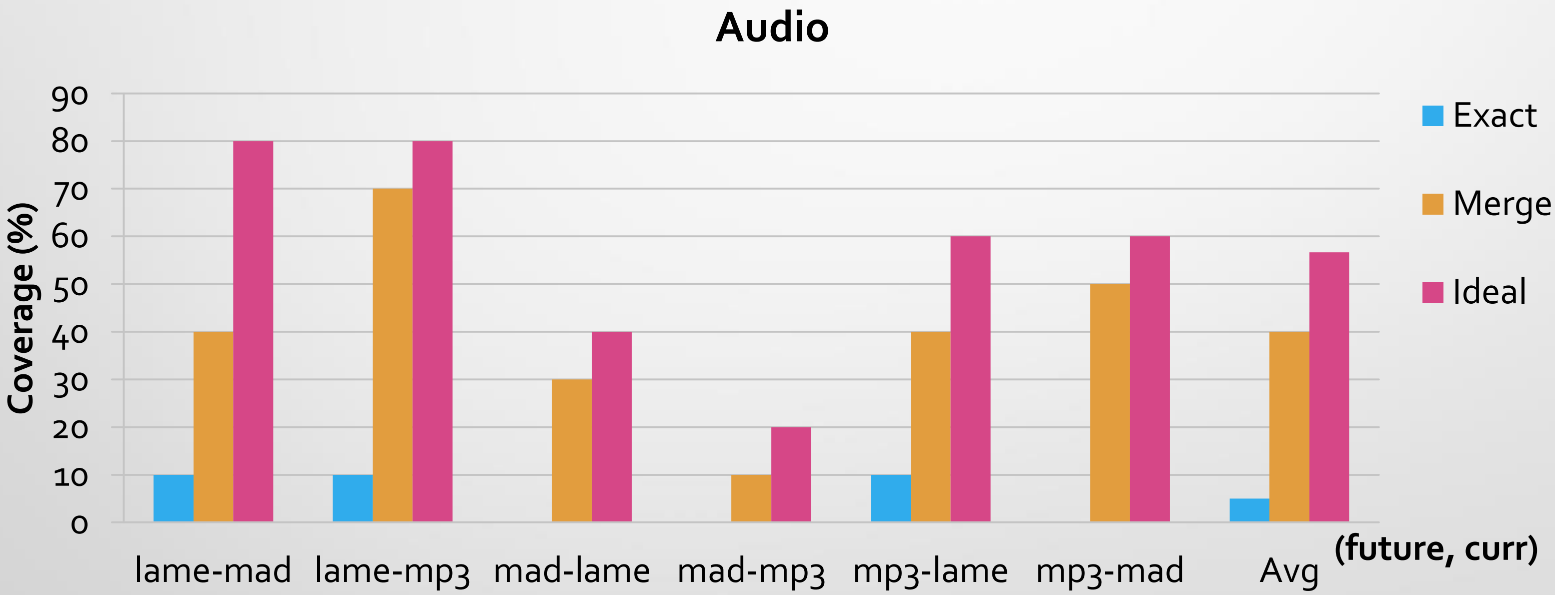


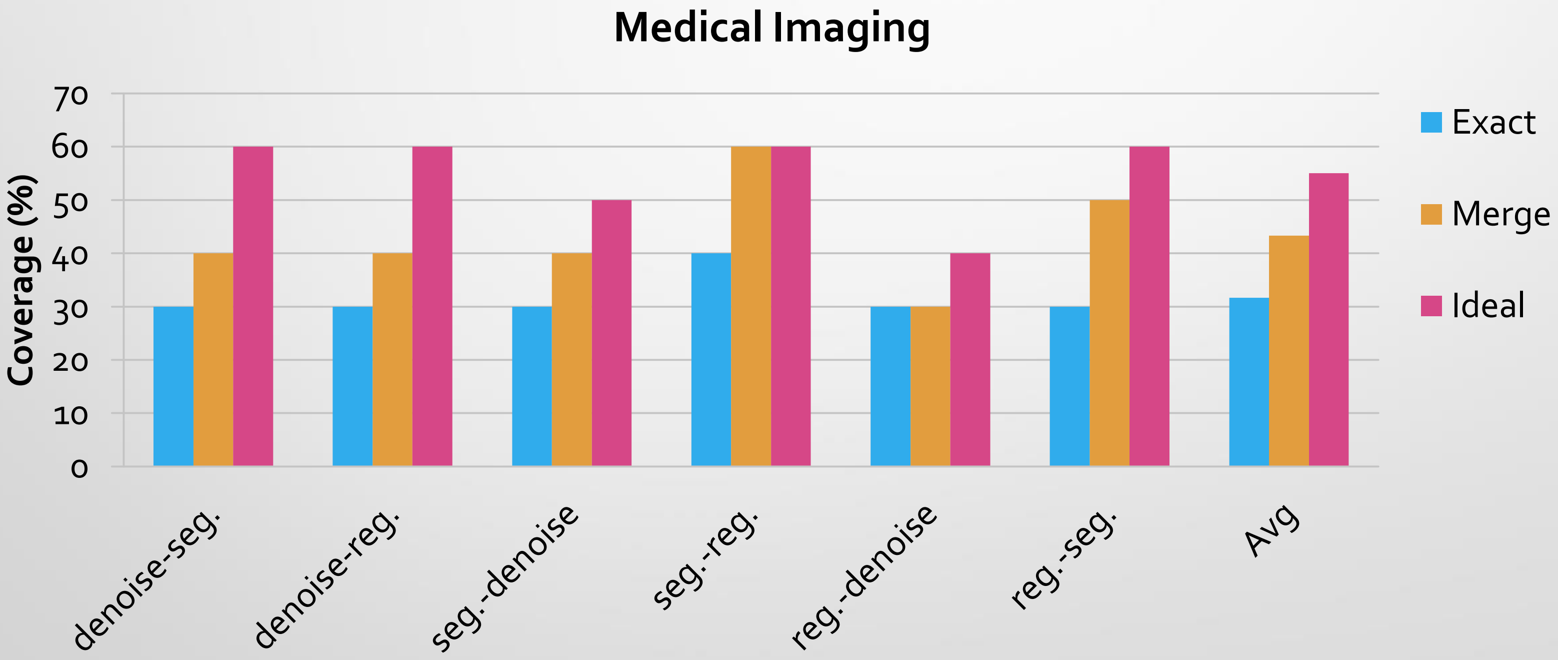
Application
RF I/O Counts
CP Enumeration
Edit Distance
CP Merging
Instruction Counts
CI Selection
A set of CIs

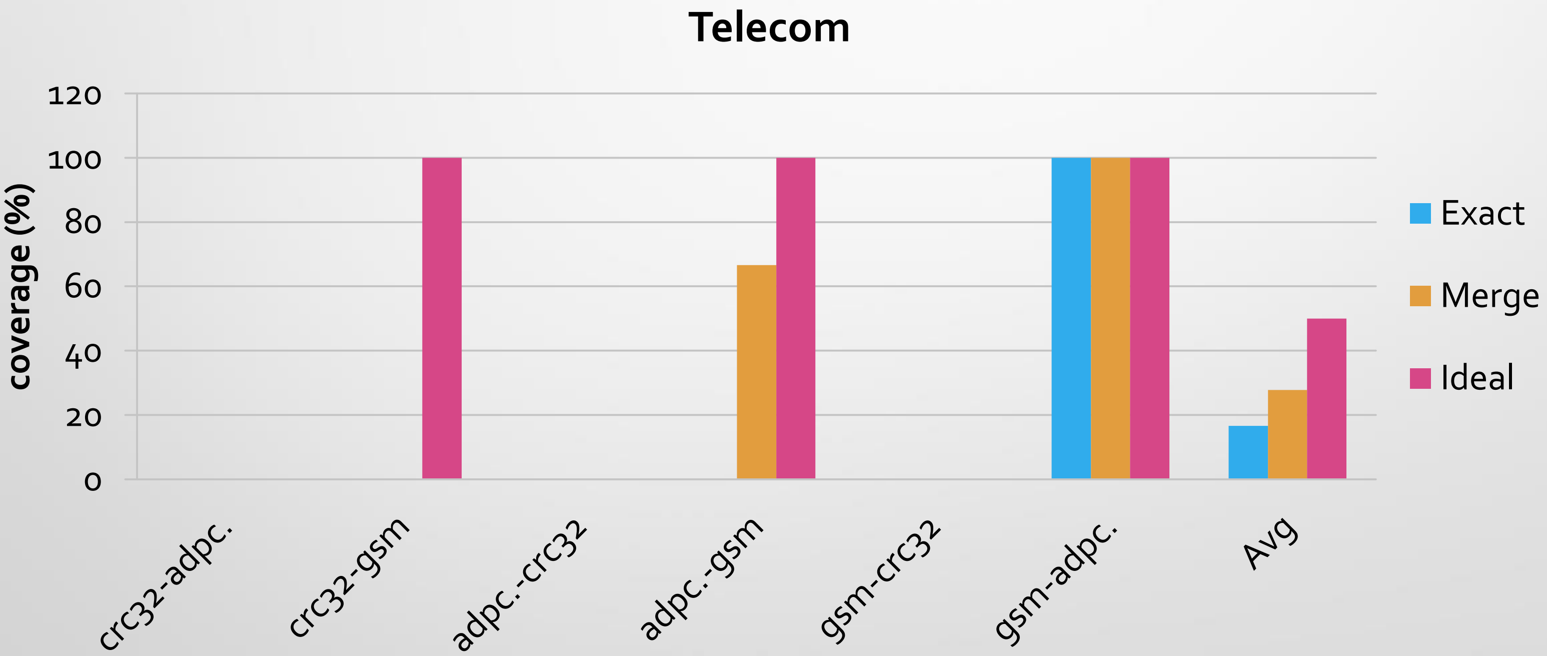
Experimental Setup

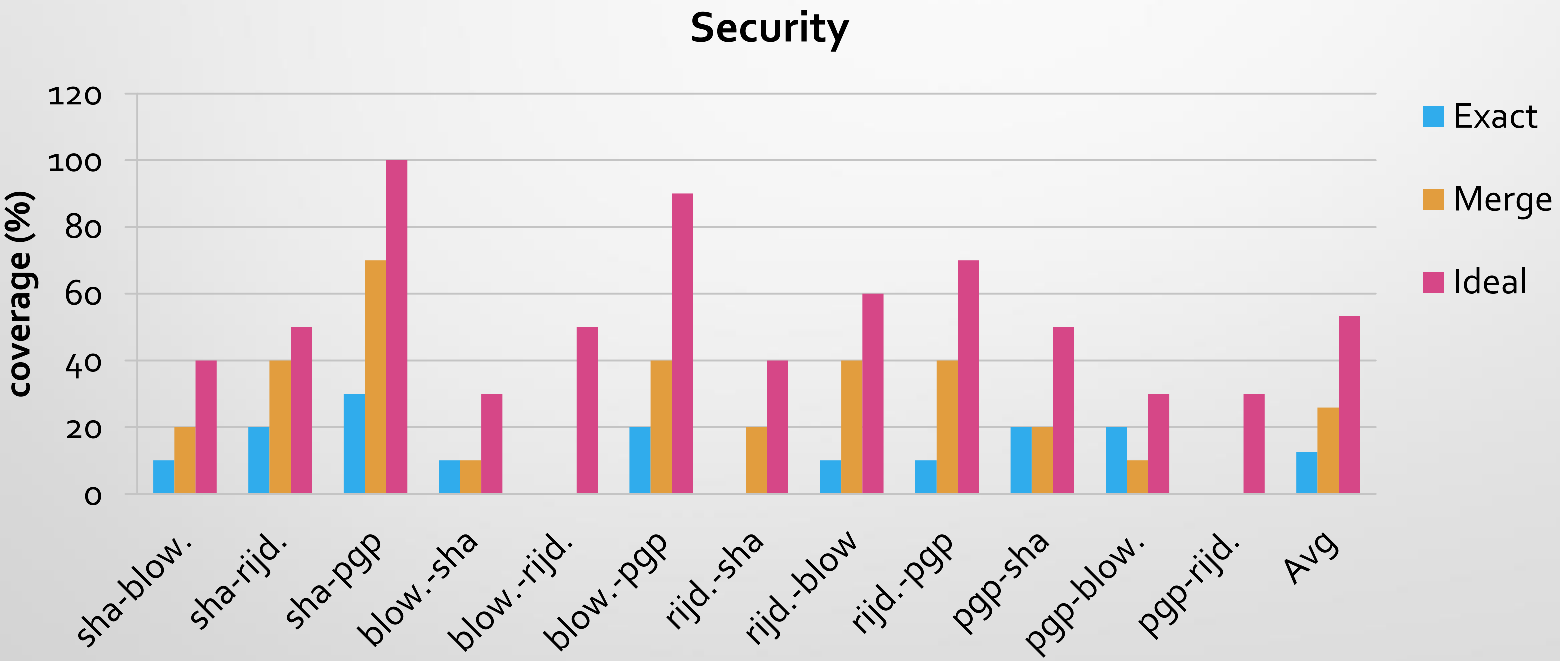


Experimental Results

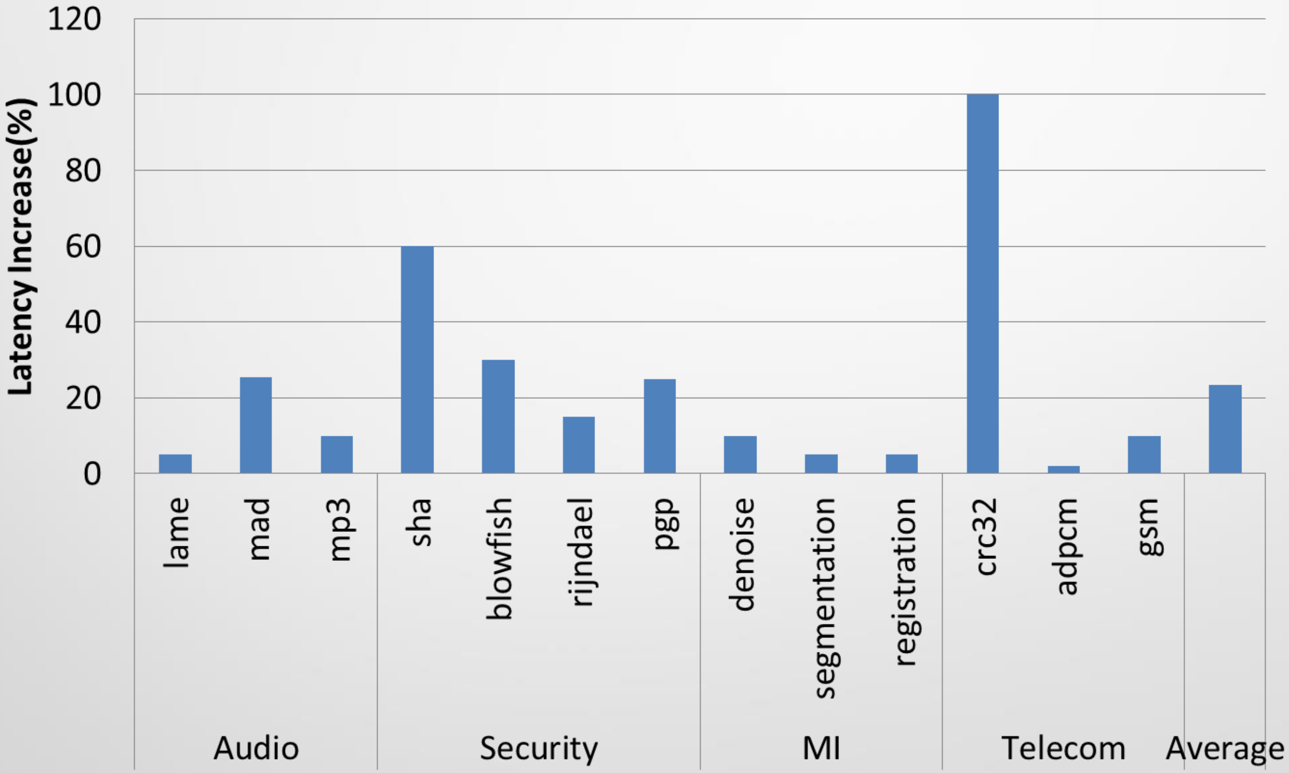



Medical Imaging
Coverage (%)
70
60
50
40
30
20
10
0
denoise-seg.
denoise-reg.
seg.-denoise
seg.-reg.
reg.-denoise
reg.-seg.
Avg
Exact
Merge
Ideal


Telecom
coverage (%)
120
100
80
60
40
20
0
crc32-adpc.
crc32-gsm
adpc.-crc32
adpc.-gsm
gsm-crc32
gsm-adpc.
Avg
Exact
Merge
Ideal


Security
coverage (%)
120
100
80
60
40
20
0
sha-blow.
sha-rijd.
sha-pgp
blow.-sha
blow.-rijd.
blow.-pgp
rijd.-sha
rijd.-blow
rijd.-pgp
pgp-sha
pgp-blow.
pgp-rijd.
Avg
Exact
Merge
Ideal

Experimental Results (Cont'd)



Conclusion

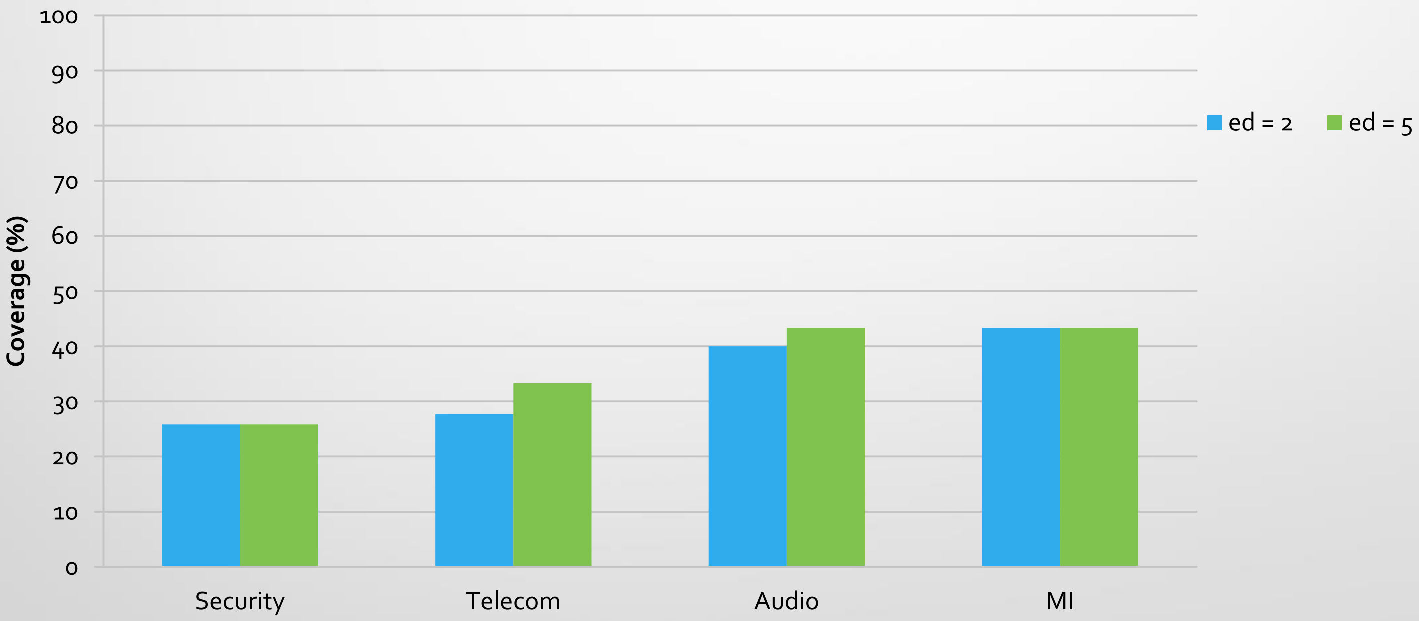
- Flexible Custom Instruction
  - Hypotheses on similarity of application in the same domain
  - Edit distance based instruction merging
    - Support future apps by supporting as many CPs of current apps as possible
- Up to 7X flexibility improvement with 23% latency increase on average

References

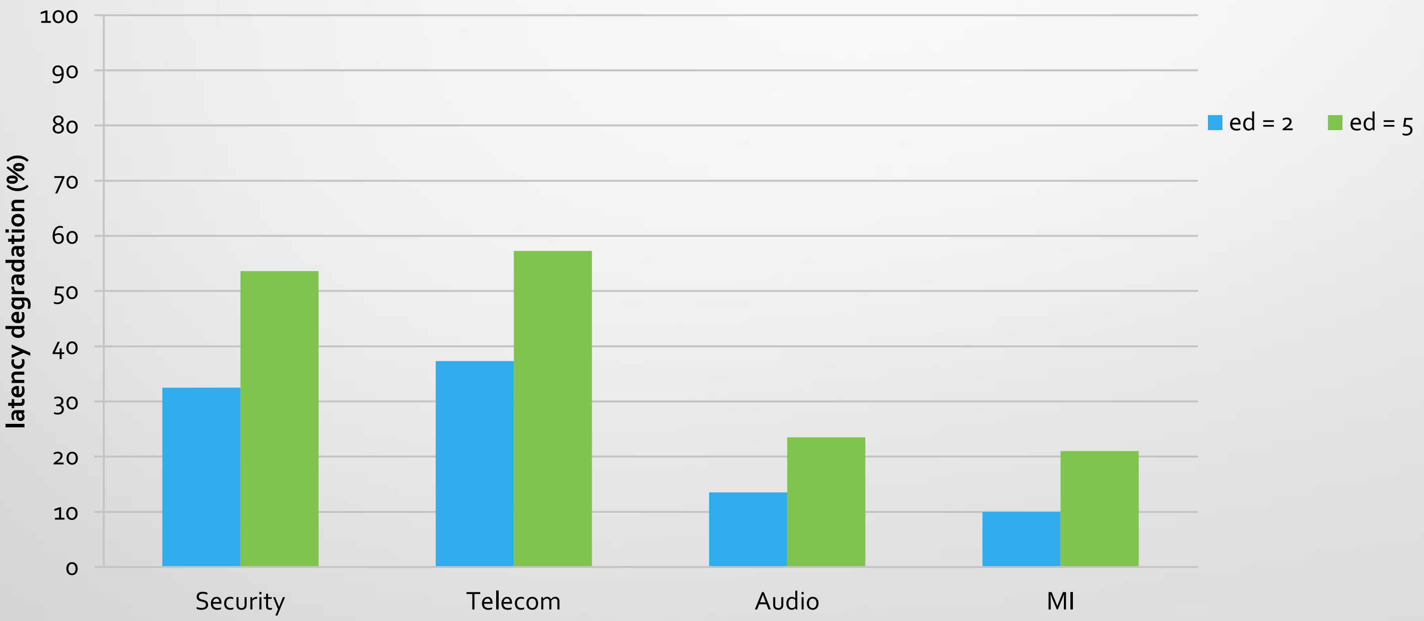
- [1] K. Fan, M. Kudlur, G. Dasika, and S. Mahlke, "Bridging the computation gap between programmable processors and hardwired accelerators," HPCA 2009
- [2] M. Stojilovic, D. Novo, L. Saranovac, P. Brisk, and P. Ienne, "Selective fexibility: Creating domain-specic reconfigurable arrays," IEEE TCAD May 2013

Backup

Coverage w.r.t Edit Distance



Latency Degradation w.r.t Edit Distance


latency degradation (%)
Security
Telecom
Audio
MI
ed = 2
ed = 5