

Heavy Hitter Detection Entirely in The Dataplane

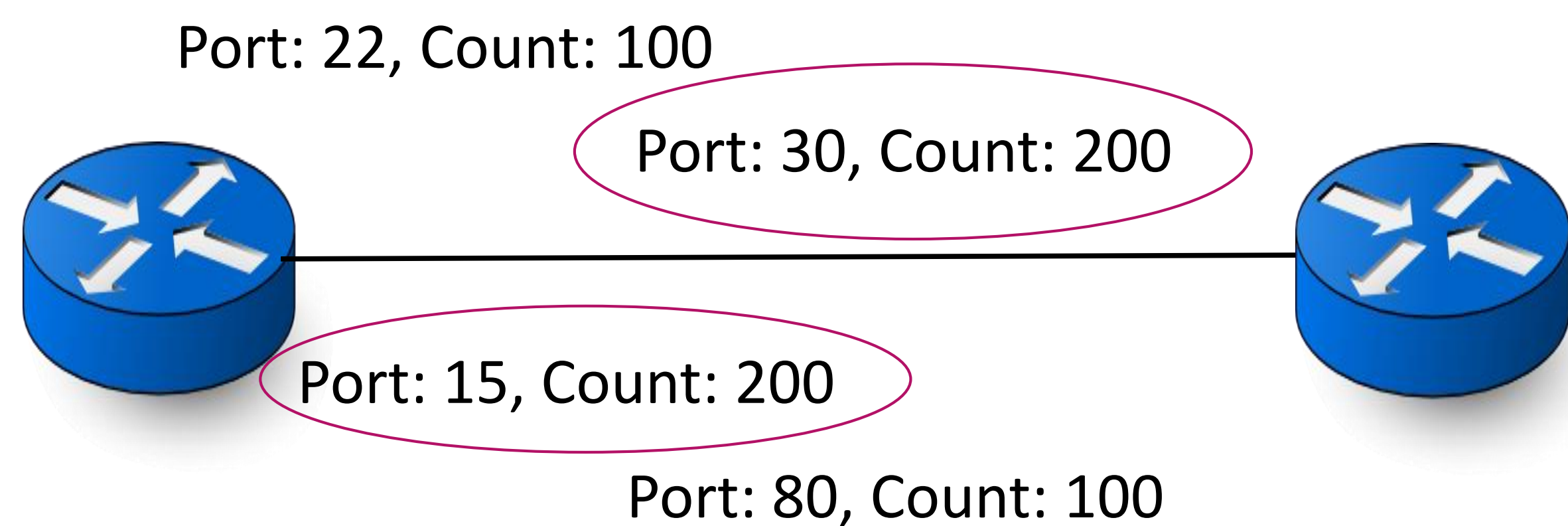
Vibhaalakshmi Sivaraman¹, Srinivas Narayana², Ori Rottenstreich¹,

S. Muthukrishnan³, and Jennifer Rexford¹

¹Princeton University, ²MIT CSAIL, ³Rutgers University

1. Problem Statement

- Efficiently identify the flows contributing the most traffic using programmable hardware
- Useful in diagnosing congestion, traffic accounting and DoS attacks



2. Challenges

Space:

- Too many flows to fit in switch memory

Speed:

- Packets traverse link at rates as high as 100 Gbps
- Packets can only be fed forward in the pipeline
- Need deterministic processing time for each packet

Accuracy:

- Measurements needed at fine-grained timescales
- Must maintain flow identifiers for heavy flows

3. Opportunities

Switching Hardware:

- Match-action tables to update per-flow statistics per packet
- New hardware with the ability to program stateful rules in the dataplane

Space Saving Algorithm:

Top-K detection by eviction of the smallest flow currently tracked

Goal: Methods with high accuracy within hardware constraints

New Key K8

Before	After
K1, 4	K1, 4
K2, 2	K2, 2
K3, 1	K8, 2
K4, 10	K4, 10
K5, 3	K5, 3
K6, 90	K6, 90
K7, 8	K7, 8

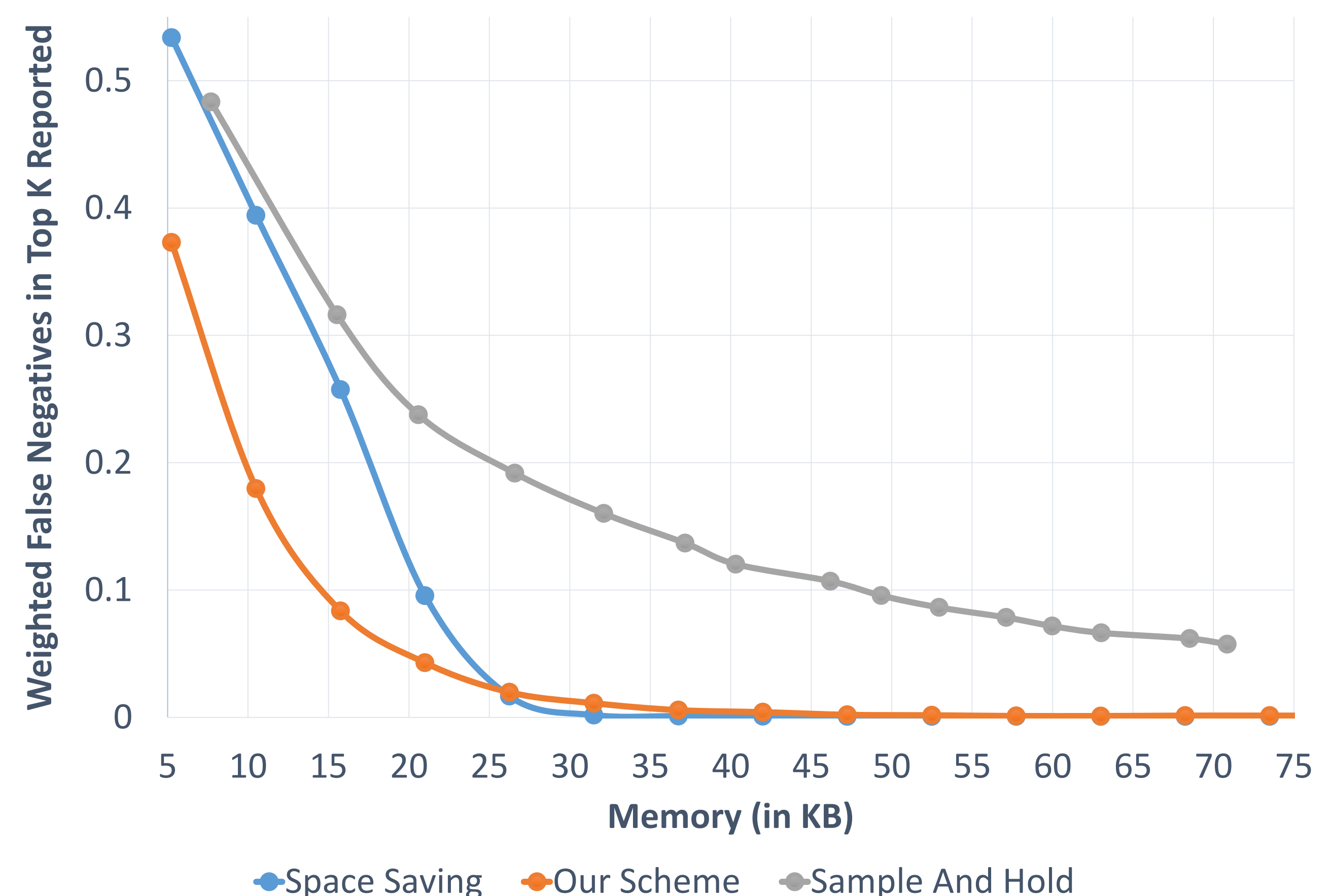
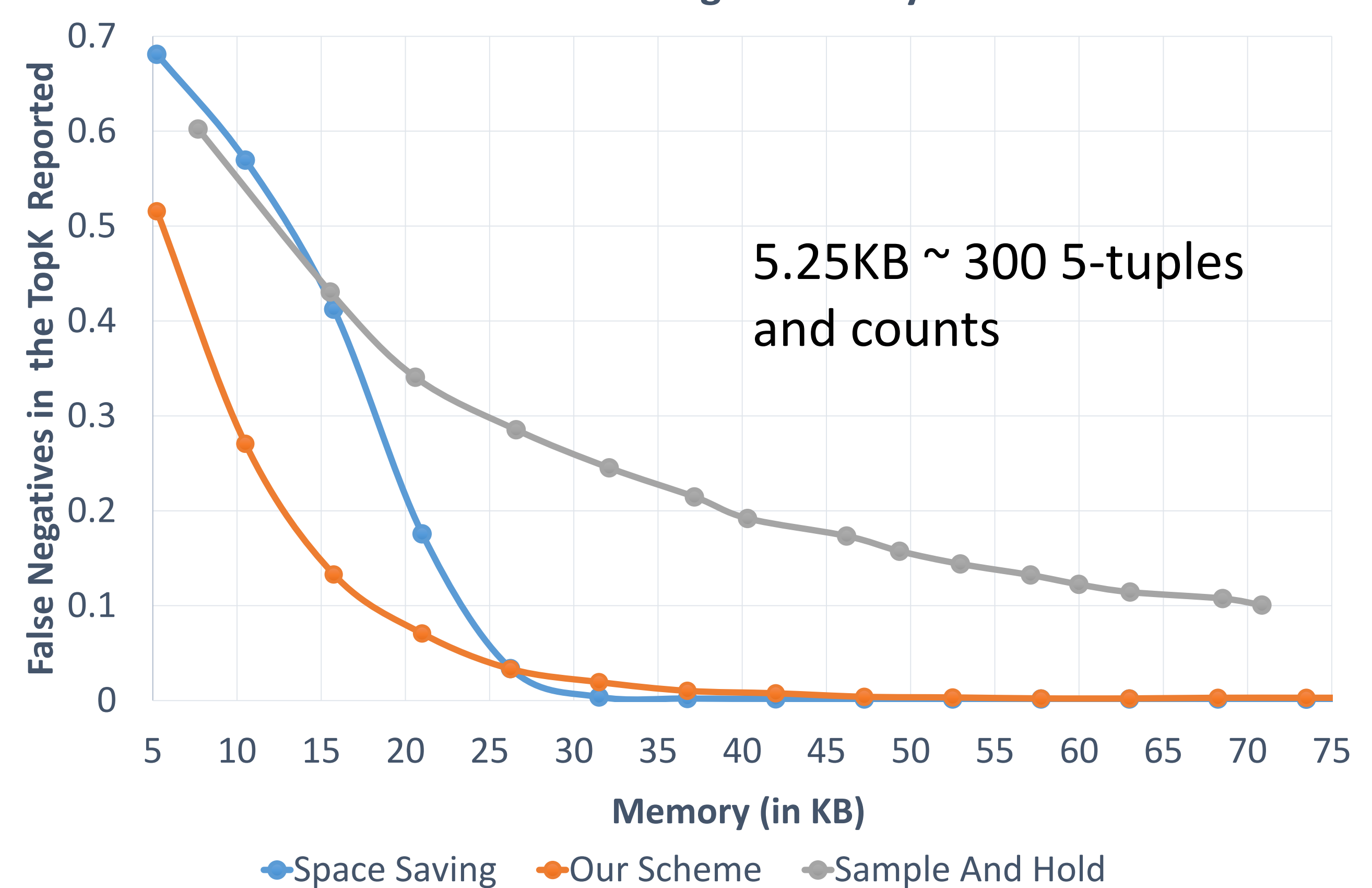
4. Insights

- Evict smaller table entries when constrained for space
- Find minimum of a small fixed number of table entries instead of global table minimum
- Spread the searched entries across multiple stages, reading exactly one entry per stage
- Approximate minimum of searched entries “as you go” to maintain feed-forward processing

5. Evaluations

Results of simulations run on CAIDA traces with 1M packets, 47337 flows

Accuracy vs Memory Trade Off across Competing Schemes in Detecting 300 Heavy hitters



Prototyped and tested the algorithm in P4

For more information:

vibhaa@princeton.edu, alephtwo@csail.mit.edu