

1 Objective

We present an algorithm (CELL) for estimating flows sizes, which combines estimators with efficient flow representation for superior memory reduction.

We also extend CELL to the sliding window model, which prioritizes recent data over old one, by presenting two variants named RAND-CELL and SHIFT-CELL.

2 CELL

Counter Estimation Levels List (CELL) is an algorithm that brings probabilistic identifiers' compaction to counter estimation.

It has two representations, Approximate Membership (AM) and Hash Table (HT).

CELL maps the flows to levels according to their frequencies.

Each flow starts at the first level and probabilistically climbs to higher levels as more packets arrive.

It stores flow fingerprints either in levels of approximate memberships (AM) or in a hash table along with their level number.

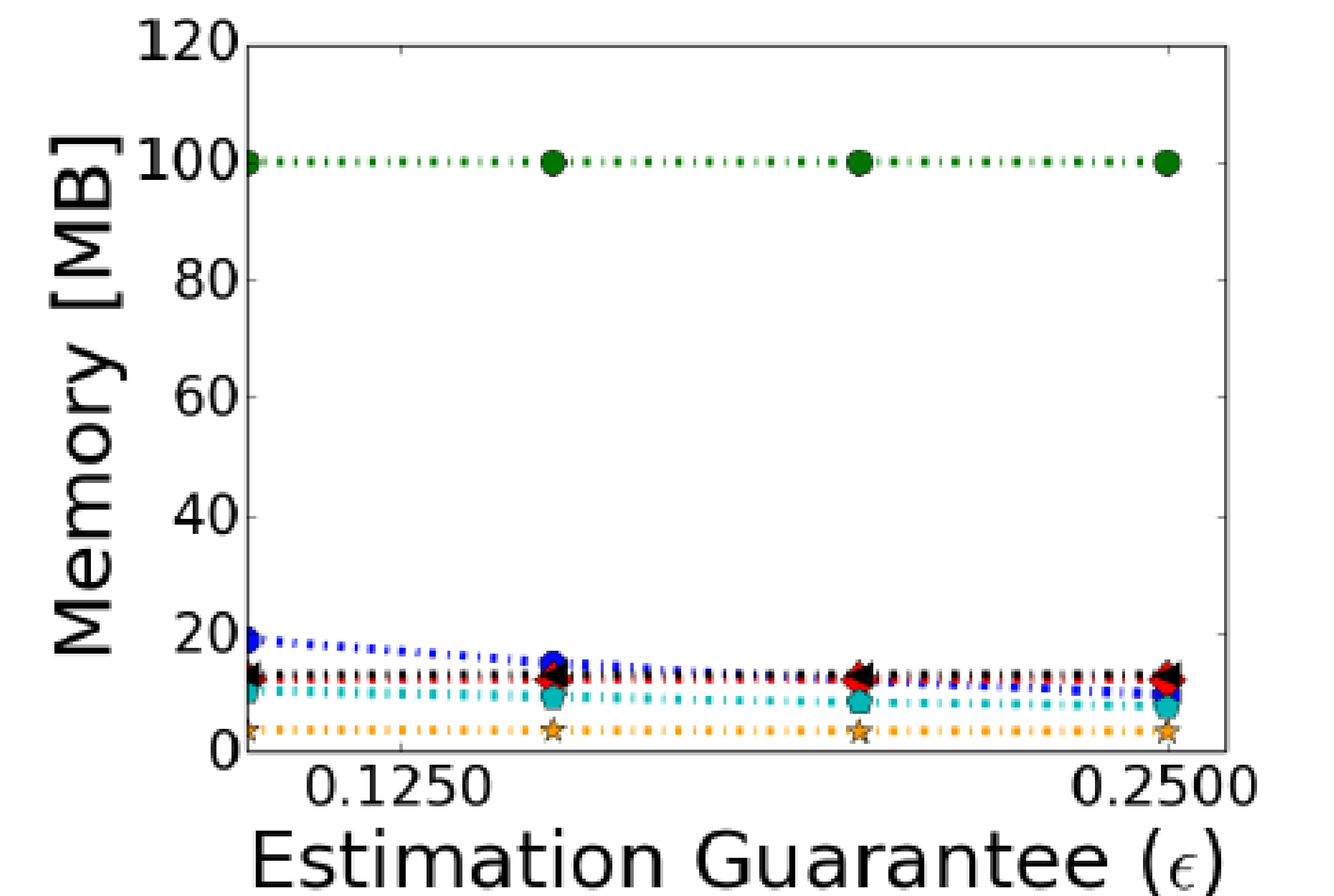
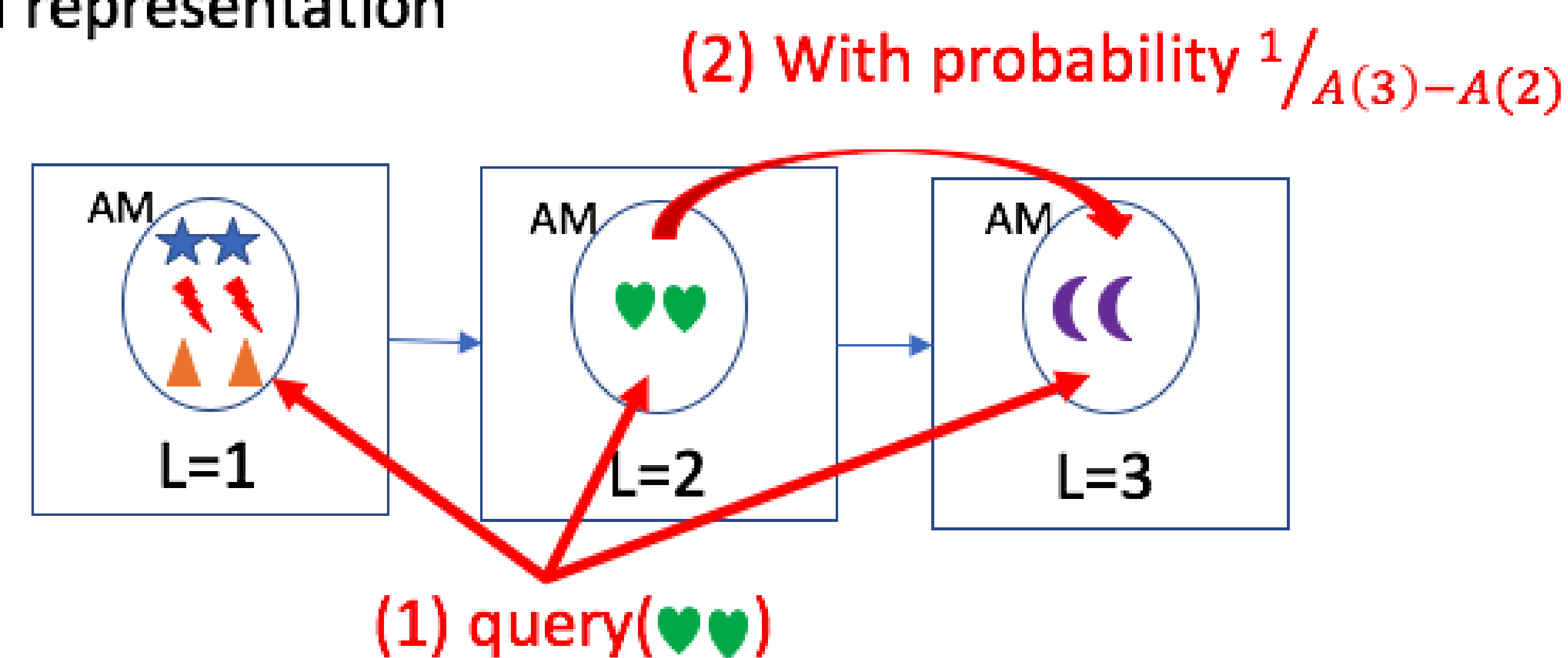
Next, it uses the level number to compute frequencies using estimators.

HT representation

Fingerprint	Level
☆☆	1
⚡⚡	2
♥♥	3 3 With probability $1/A(3)-A(2)$
☾☾	3
▲▲	1

Stream items: $X_{n-W}, X_{n-W+1}, \dots, X_{n-1}$
 Fingerprints: ▲▲, ☾☾, ..., ♥♥
 New item: $X_n, \dots, h(X_n)=♥♥$
 Estimation function: $A_\epsilon(L)$

AM representation



3 RAND-CELL

Randomized Counter Estimation Levels List (RAND_CELL) extending CELL to the sliding window model by randomly seeking a prior element arrival and decrementing one from the corresponding flow's estimation.

4 SHIFT-CELL

Shift Counter Estimation Levels List (RAND_CELL) also extending CELL to the sliding window model by applying batch reductions on all flows every C arrivals.

5 Results

