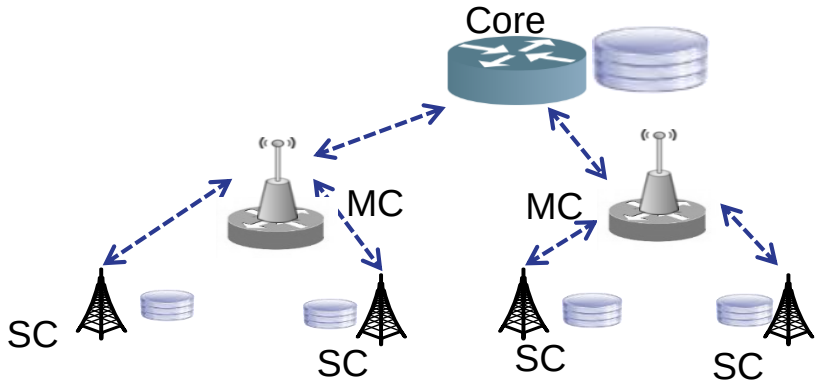
PHY-Aware Cooperative Caching for End-to-End QoS in Cellular Networks

Presenter: Thrasyvoulos Spyropoulos (Eurecom)
Collaborators: Alina Tuholukova (Eurecom), Giovanni Neglia (INRIA)

Wireless Edge Caching

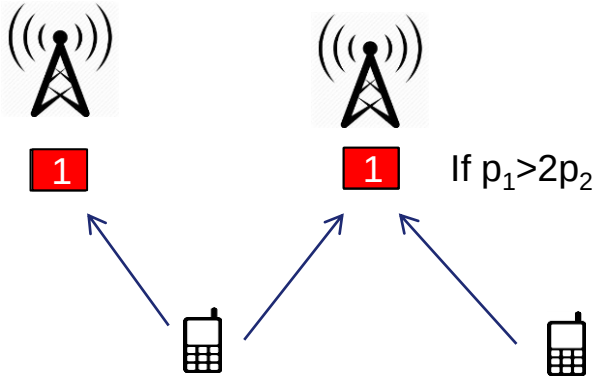


Small Cells (SC) → to improve radio access

Edge Storage → to offload backhaul (or radio) traffic

Key Question: What to cache where?

Traditional Caching: Maximize Hit Rate



Single Cache?

- ***Most popular content***

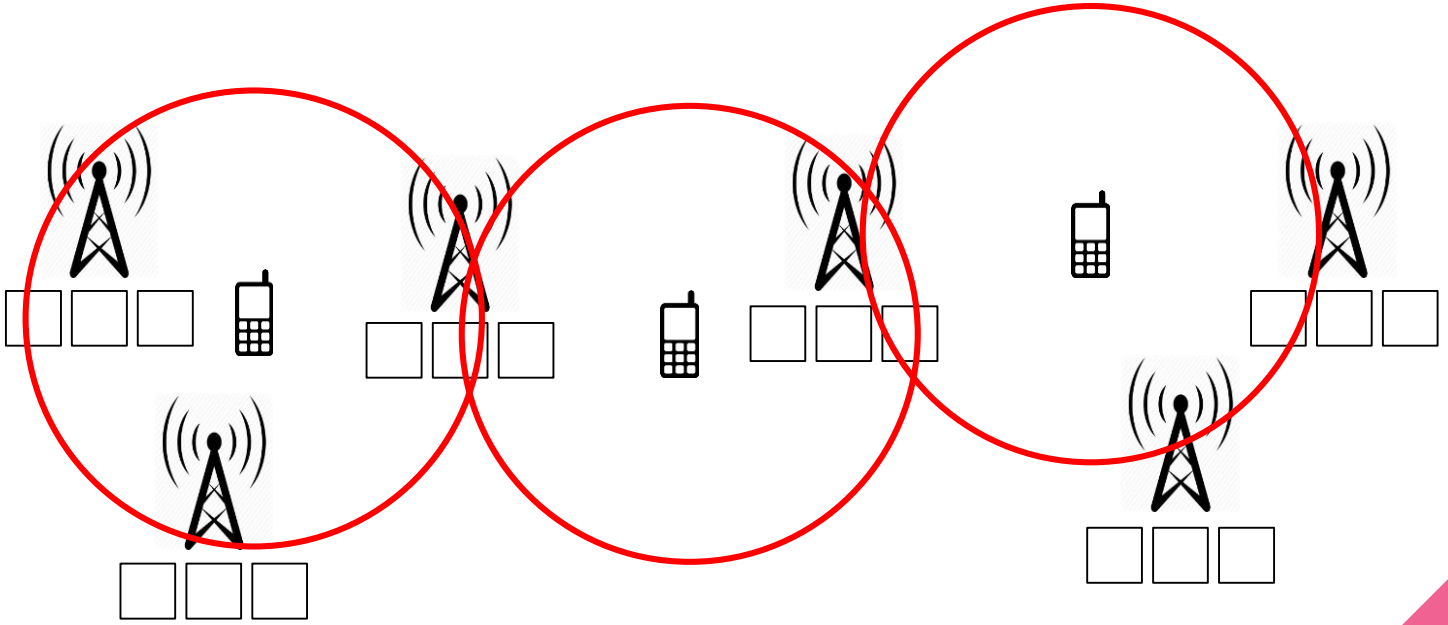
Two Caches?

- ***Diversify***

Multiple users?

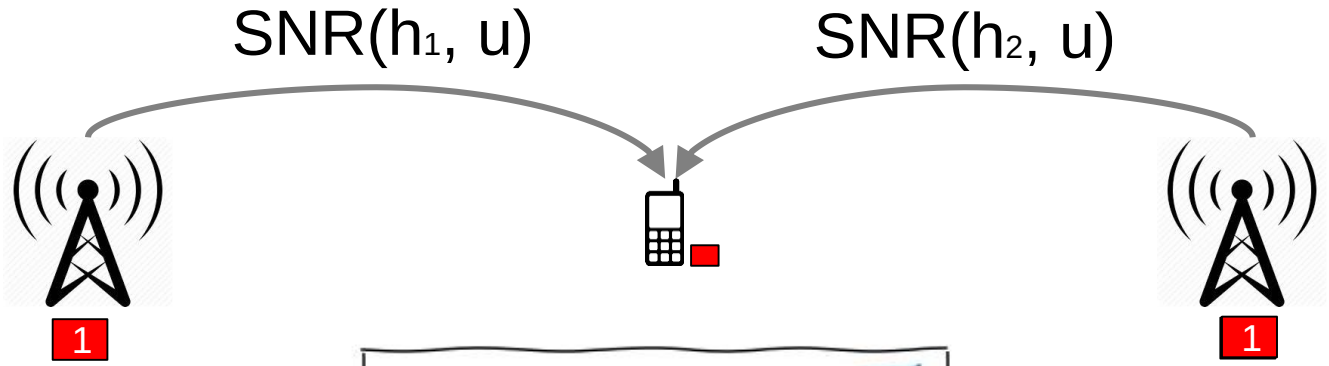
- ***Depends***

Femto-caching problem: NP-hard (Trans. on IT 2013*)



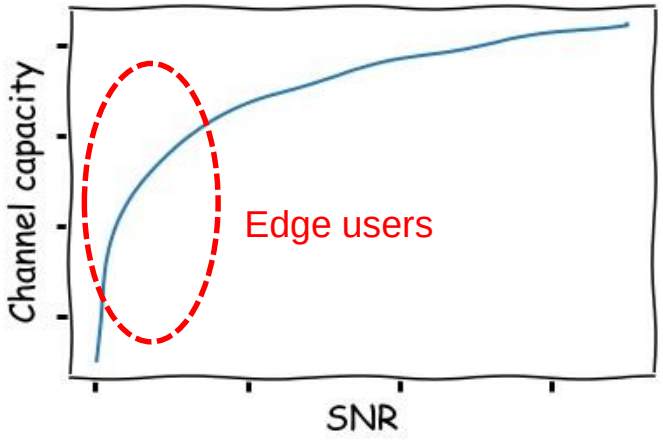
*Shanmugam et. al: FemtoCaching: Wireless Content Delivery Through Distributed Caching Helpers. IEEE Trans. Information Theory 2013

PHY-friendly Caching



Single transmitter

Joint transmission (CoMP)

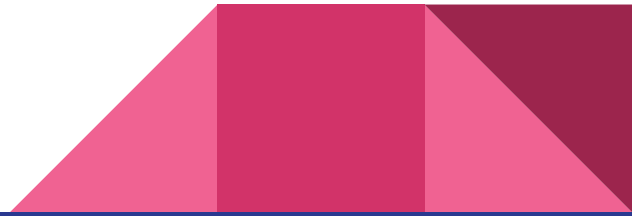


$_1, u)]$

$_1, u) + SNR(h_2, u)]$

Modeling

- Decision variables: **where** to cache **which** files
- Metric: “end-to-end” time to download a file (radio access + backhaul)
- Parameters:
  - Popularity distribution
  - Cache size
  - SNR between each user and base station



Trade-off: hit rate VS opportunities of Joint Tx

To alleviate backhaul:

❏ Increase hit rate

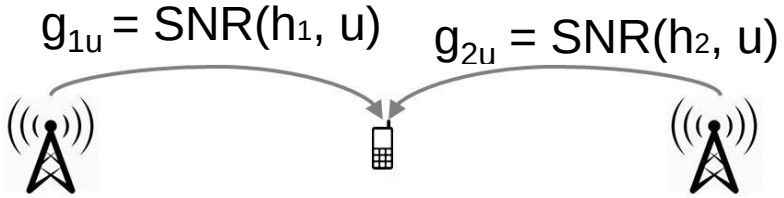
By ***caching diverse files***

To increase radio access rates:

❏ Create Joint Tx opportunities

By ***caching same (popular) files***

Caching Model



Cache miss: fetch to the strongest BS, download from this BS

Backhaul delay Radio access delay

$$D_{miss} = D_{BH} + \frac{file\ size}{(W \log [1 + \max(g_{1u}, g_{2u})]}$$

Cache hit: download from all neighbor BSs that contain the requested file

Radio access delay

$$D_{hit} = \frac{file\ size}{(W \log [1 + \sum_j x_{kj} \cdot g_{ju}]}$$

x_{kj} = 1, iff file k in BS j

Optimization Problem

Probability to request file k

Function of x_{kj}

$$\max(x_{kj}) \sum_{u} \sum_{k} p_k \cdot [P_{miss} \cdot D_{miss} + (1 - P_{miss}) \cdot D_{hit}], subj.\, capacity\ const.$$

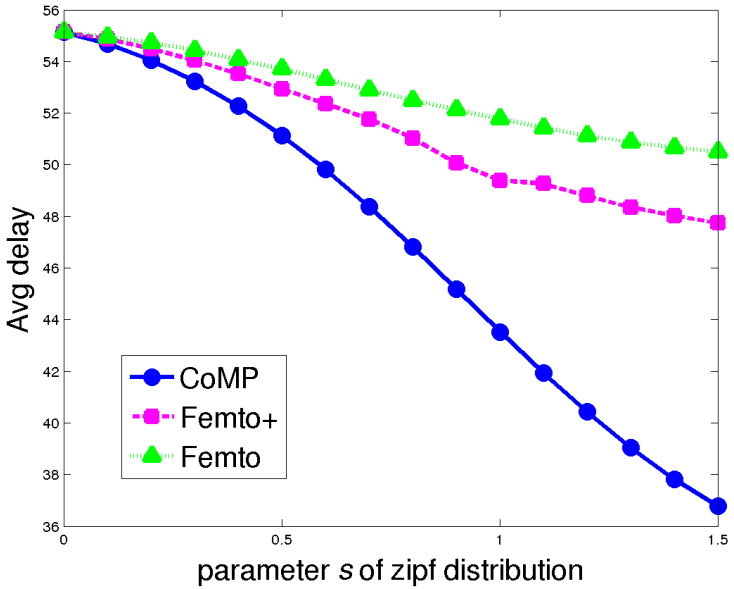
Function of x_{kj}

u: user
k: file
$x_{kj:}$ storage variables

Submodular optimization

- NP-hard problem
- Fortunately, **submodular** objective function
  - Constraints define partition matroid
- **Greedy algorithm** gives guaranteed ½ approximation
  - Iteratively choose a file that gives the largest gain

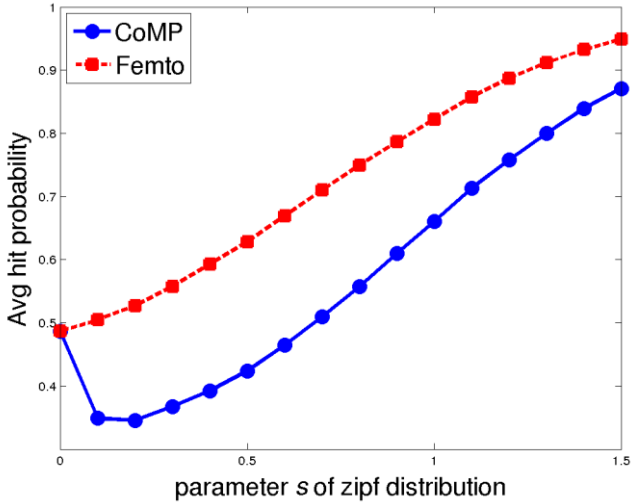
Simulation results

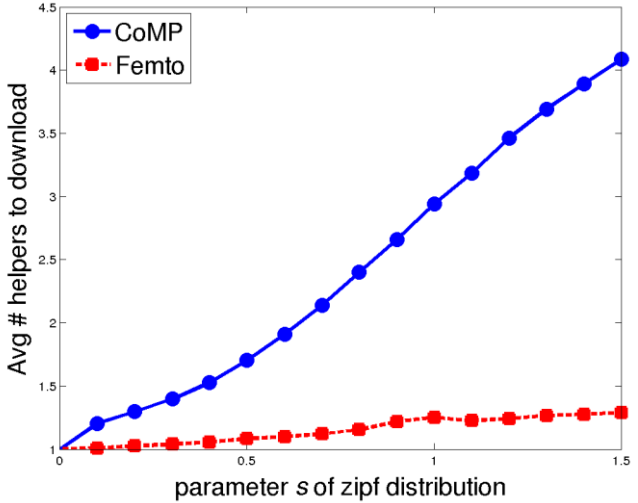


more skewed popularity distribution

- Significant improvement for more skew popularity distribution
- ! Not only due to possibility of cooperation

Simulation results: Trade-off

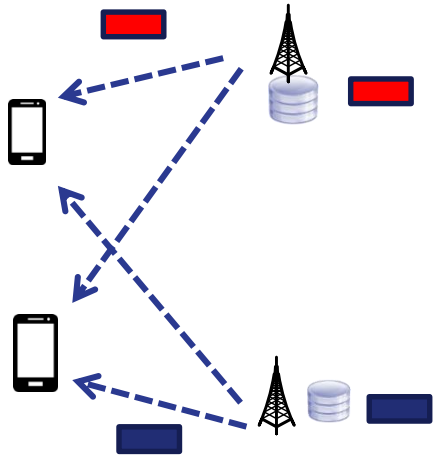




- **Femto caching** has higher hit rates
- **CoMP caching** favors opportunities for JT

Next Steps

- **More advanced PHY techniques (e.g. MU-MIMO)**
  - Higher gains (outside of log)
  - Harder to model
  - “Coded Caching” extensions?
- **Distributed/Dynamic Caching**
  - MCMC, potential games for distr. implementation
  - TTL-cache for adaptive learning



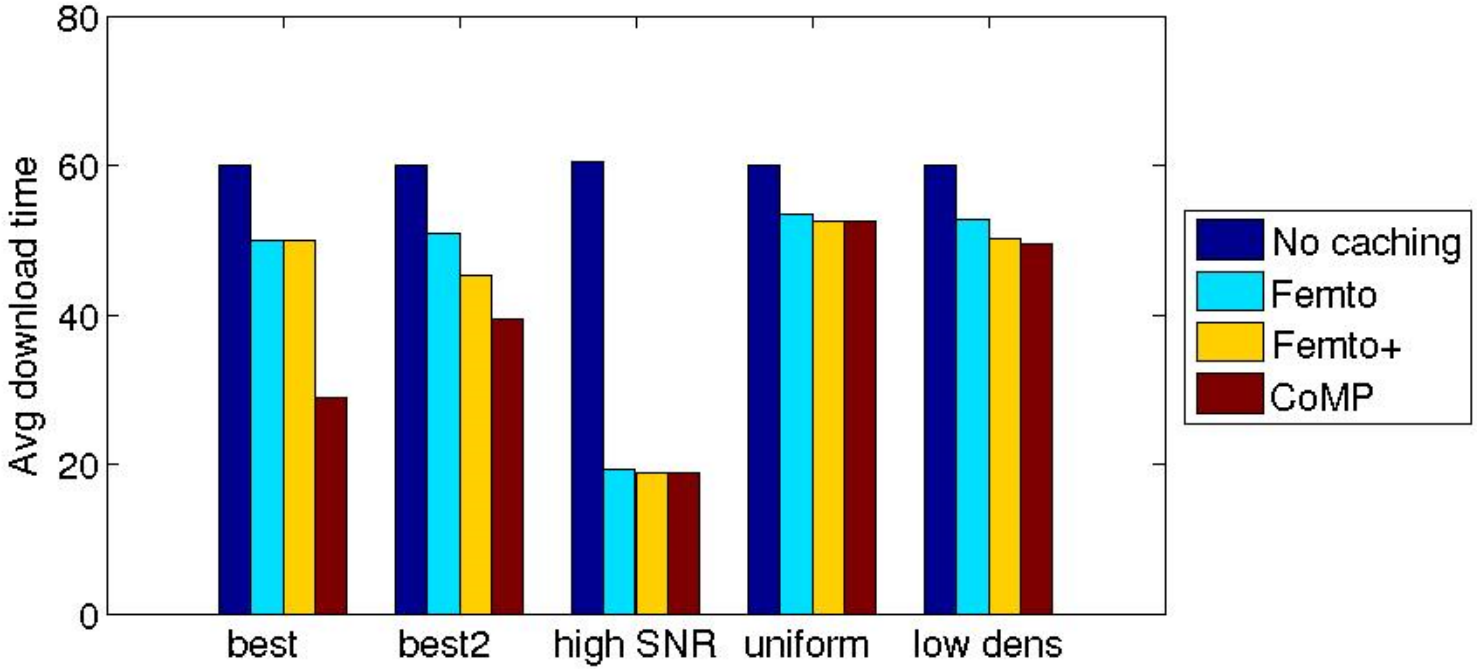
LABEX positive impact

- 2 LABEX funder projects on caching
- Led to a successful ANR JCJC project application
  - a “young investigator” type of personal grant
  - 5C-for-5G: **C**oordinated **C**aching, re**C**ommendation, and **C**ommunication of **C**ontent for 5G Cellular Networks
  - PHY-aware, but also Content and Architecture-aware (e.g. MEC)
- ERC Consolidator project (to be submitted)

Thank you!

Questions?

Simulation results



Caching Model

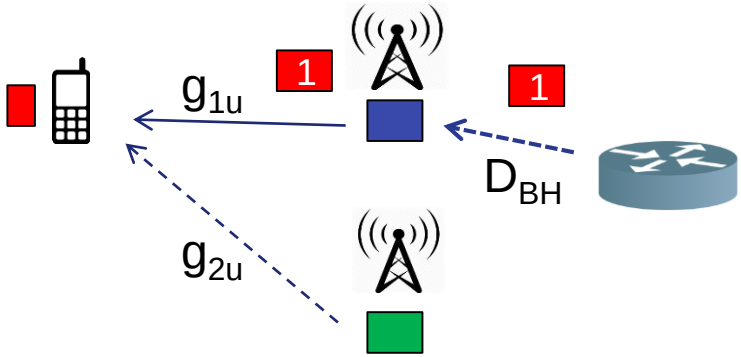
$g_{1u} = SNR(h_1, u)$

$g_{2u} = SNR(h_2, u)$

Cache miss: fetch to the strongest BS, download from this BS

Backhaul delay Radio access delay

$$D_{miss} = D_{BH} + \frac{file\ size}{(W \log [1 + \max(g_{1u}, g_{2u})]}$$



Cache hit: download from all neighbor BSs that contain the requested file

Radio access delay

$$D_{hit} = \frac{file\ size}{(W \log [1 + \sum_j x_{kj} \cdot gju]}$$

$x_{kj} = 1$, iff file k in BS j