

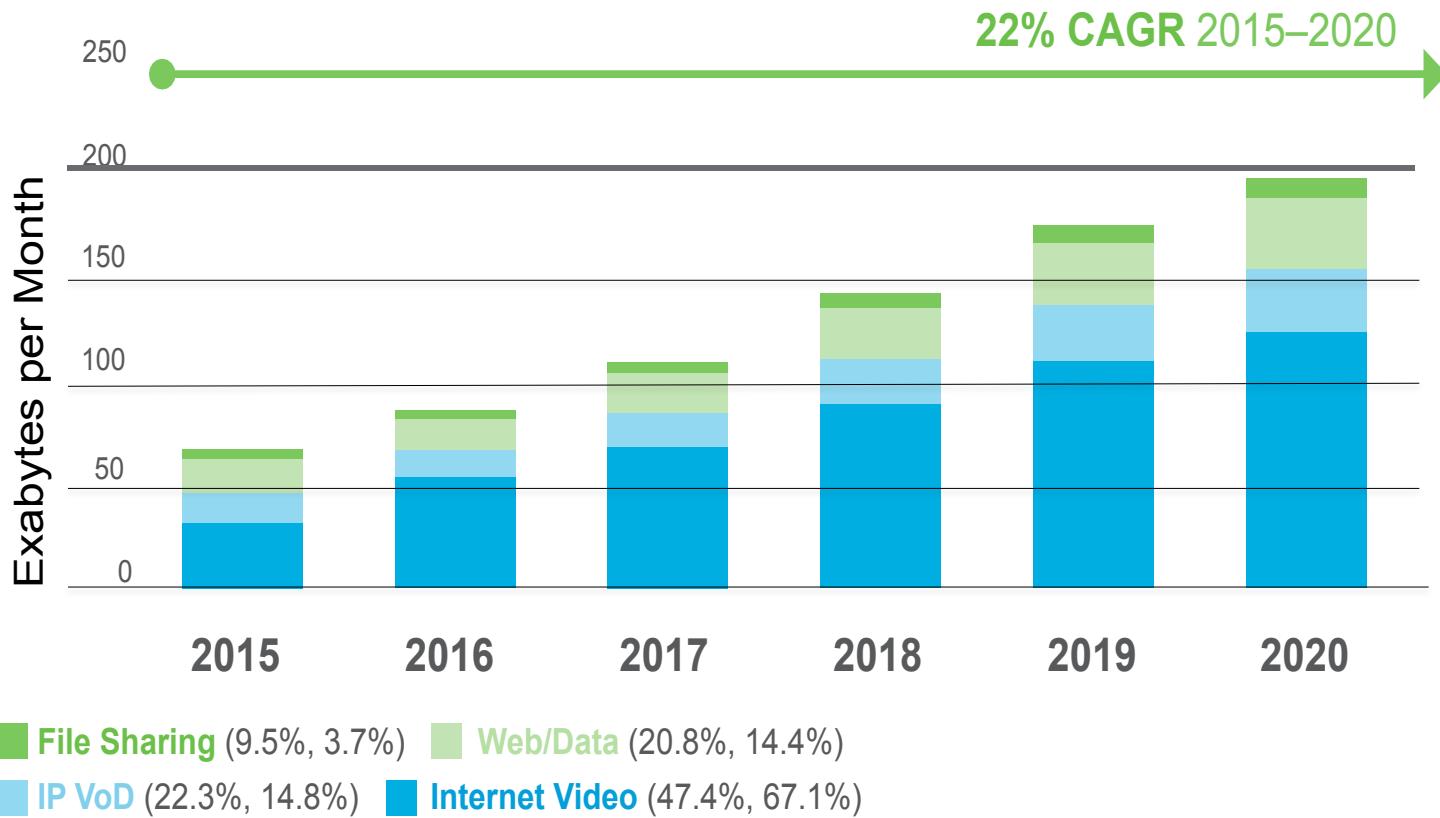
Profiling and Grouping Users to Edge Resources According to User Interest Similarity

Pengyuan Zhou, Jussi Kangasharju

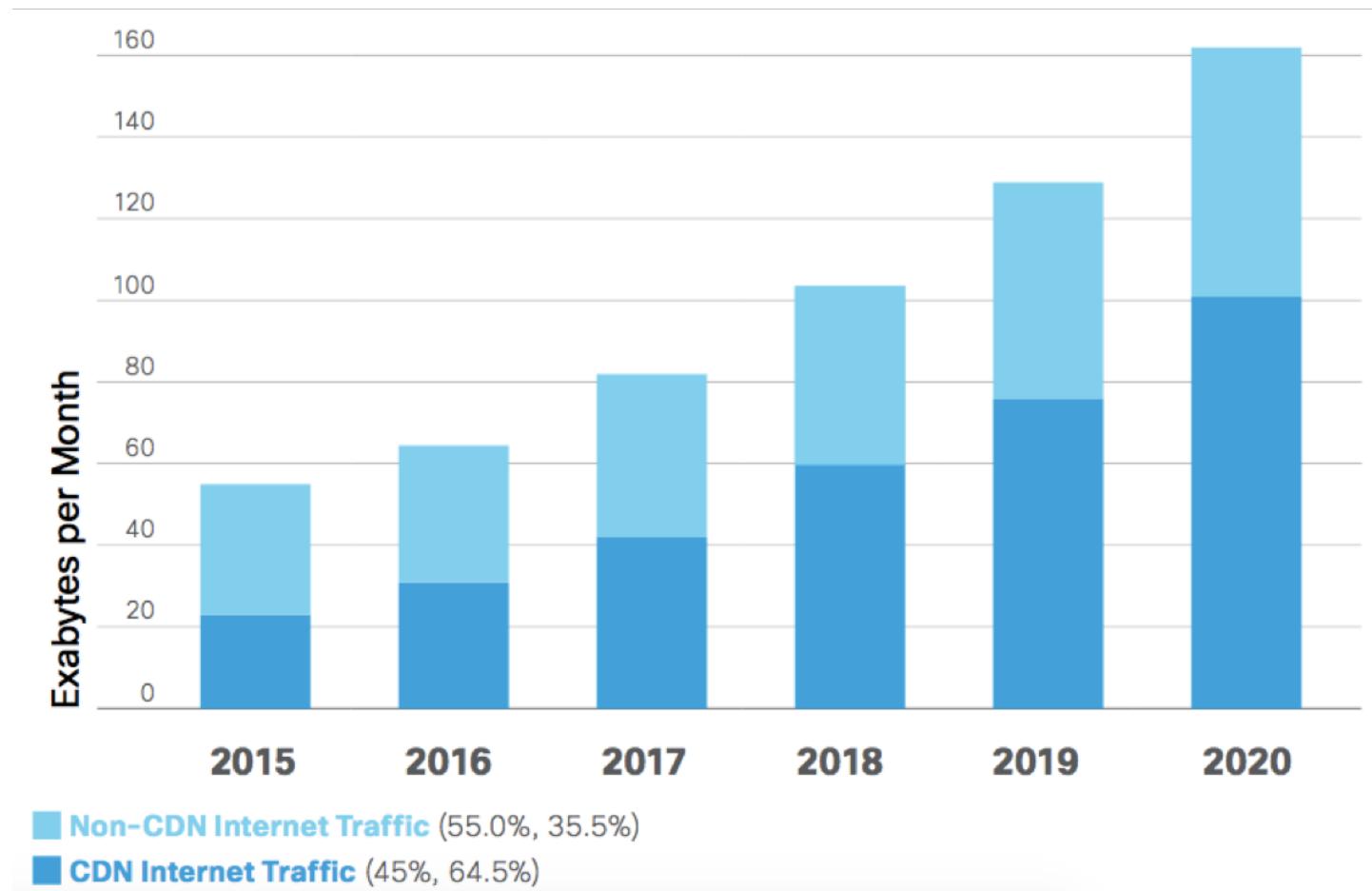
Department of Computer Science, University of Helsinki, Finland
{firstname.lastname@cs.helsinki.fi}



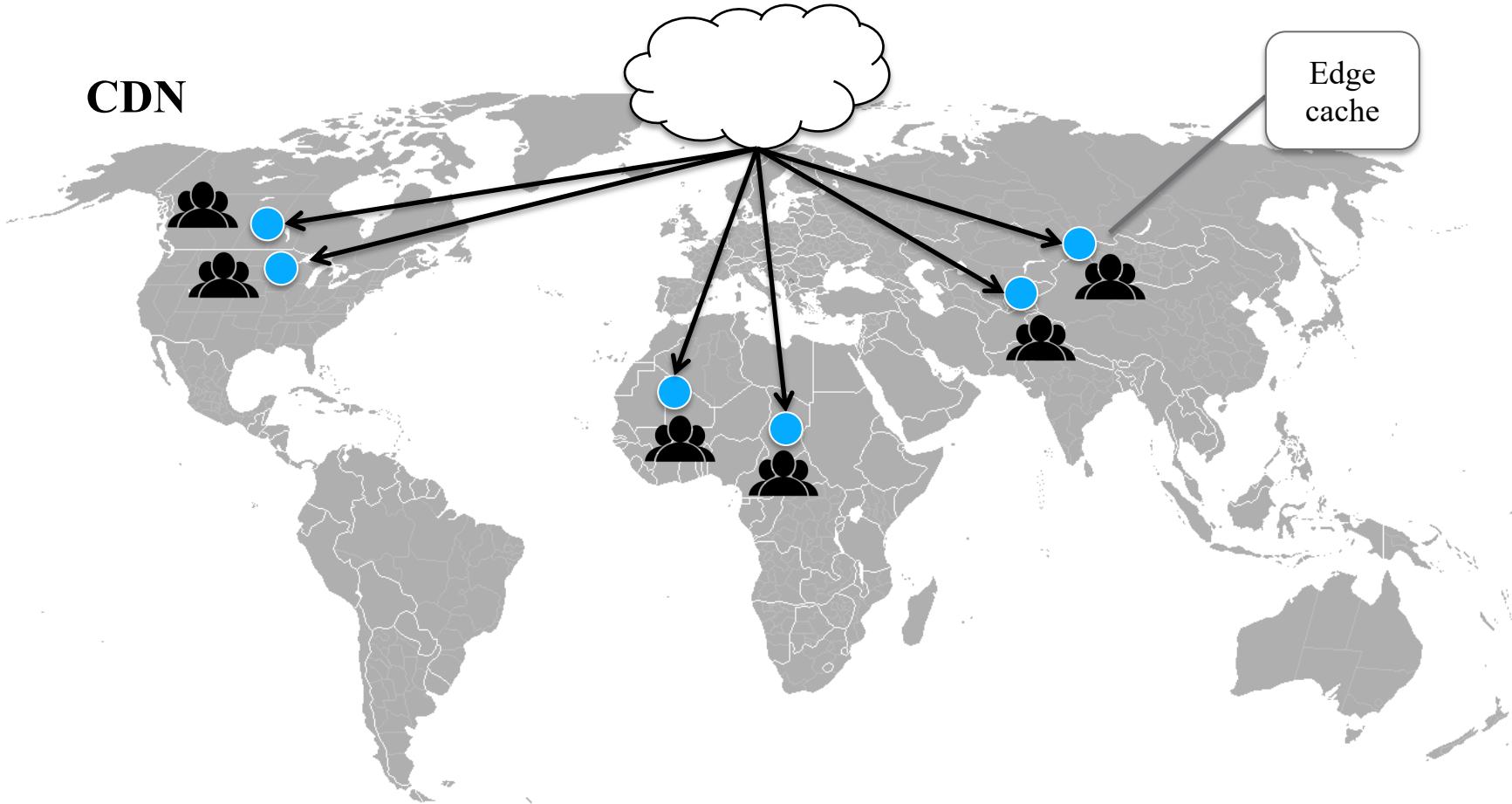
Increasing Demand



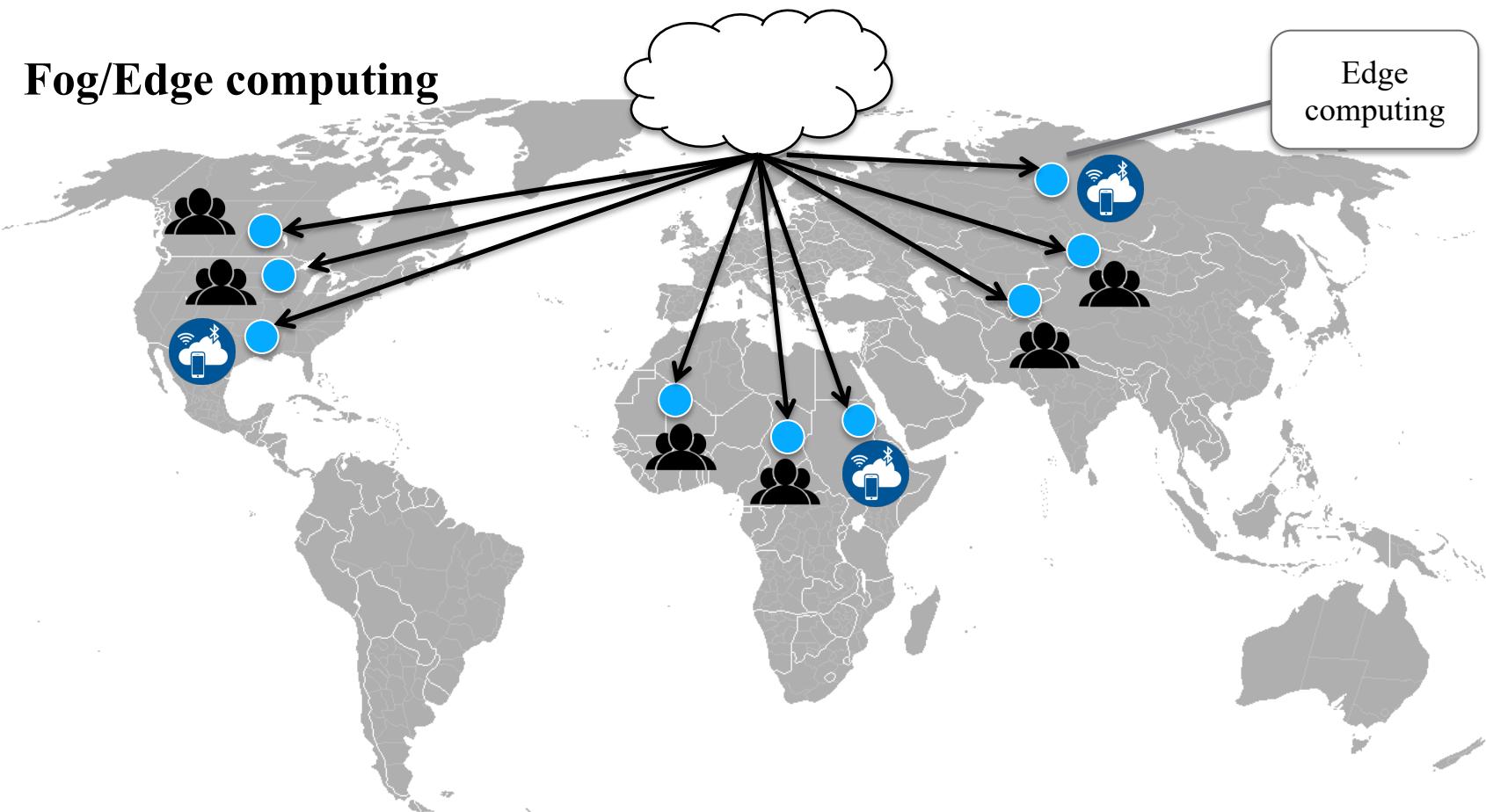
Increasing Demand



Edge Solution

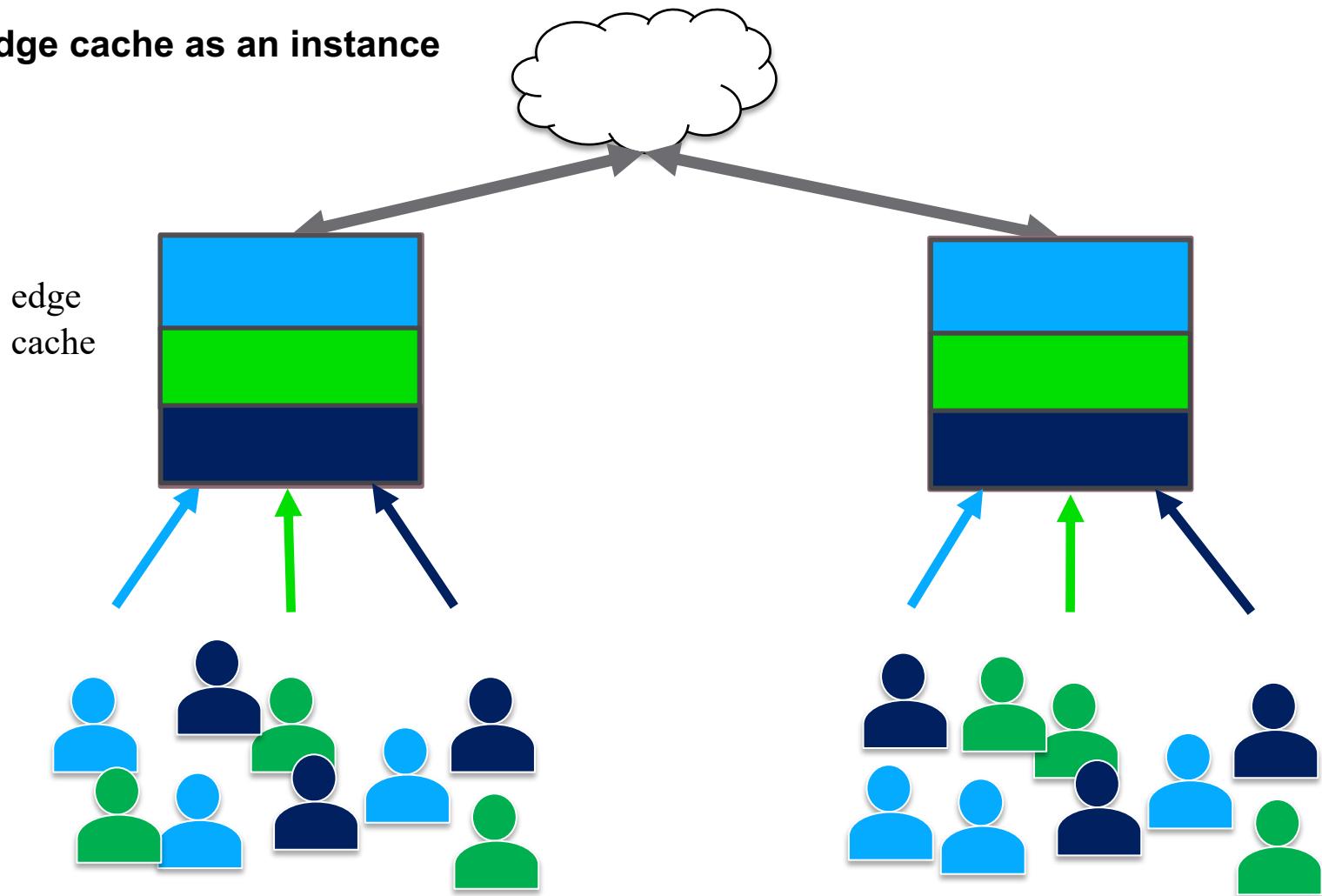


Edge Solution

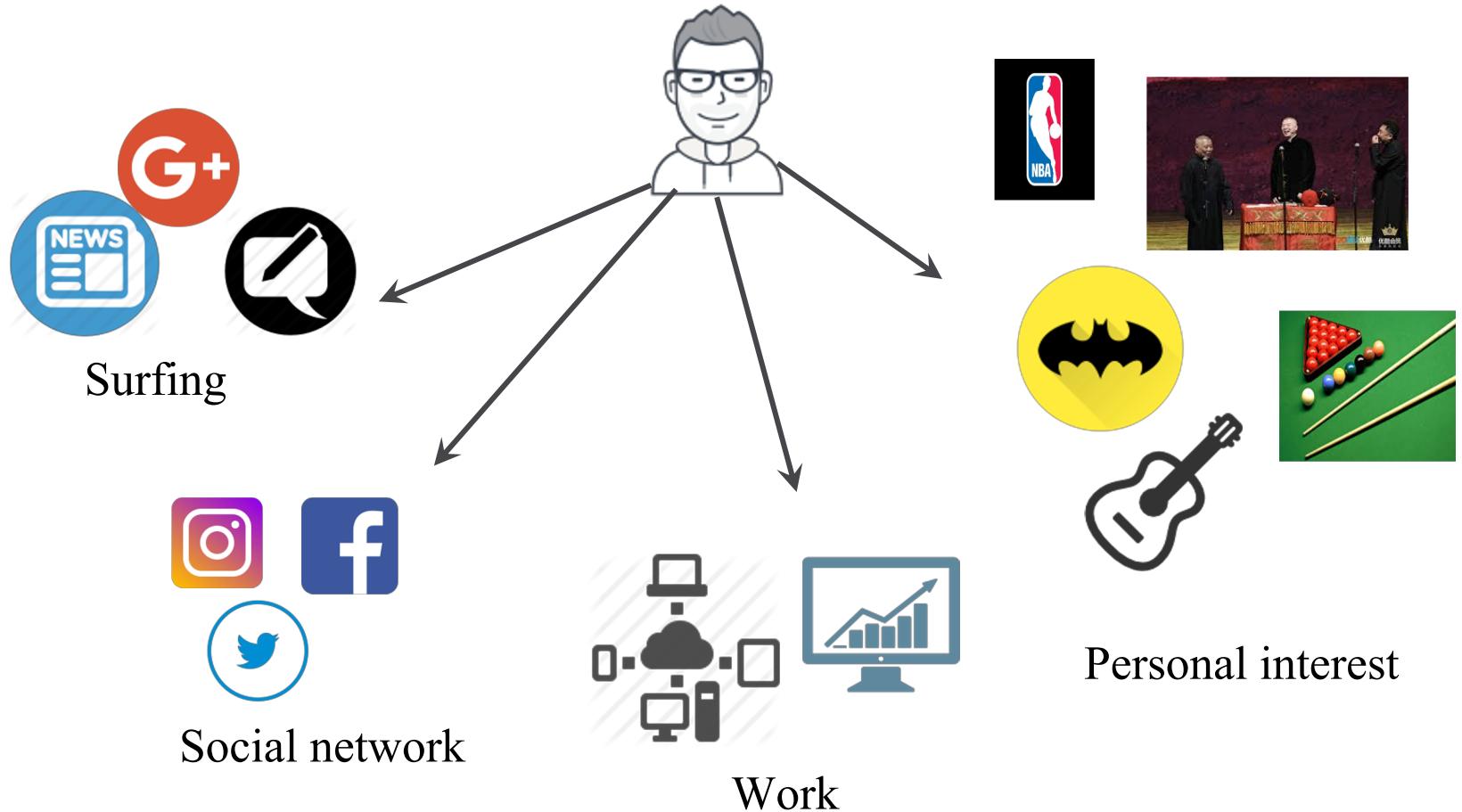


Edge Solution

Use edge cache as an instance



User Interest



User Interest

Location

Personality

Age

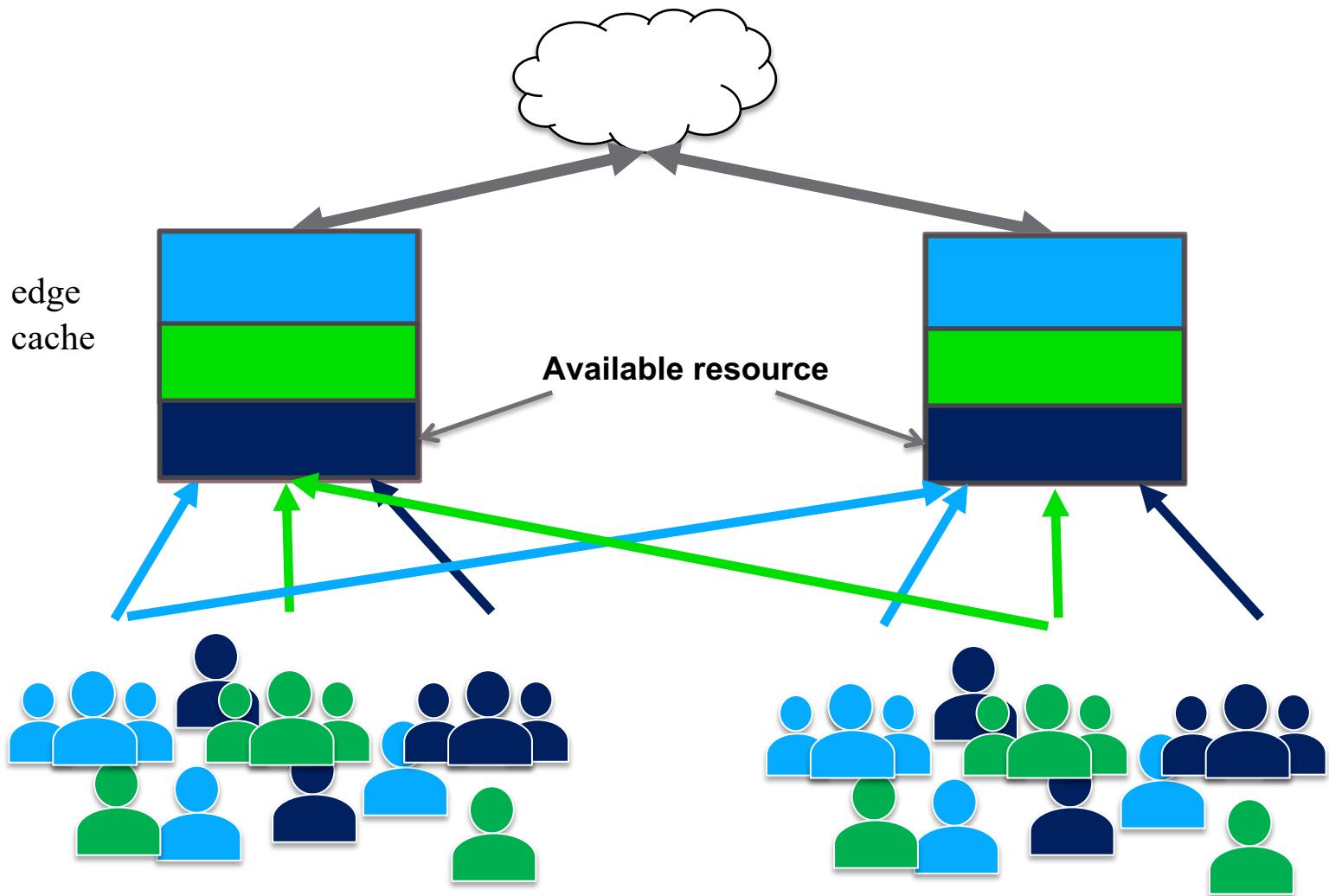
Language

Gender

Job

Education

Grouping Users to Edge Resources

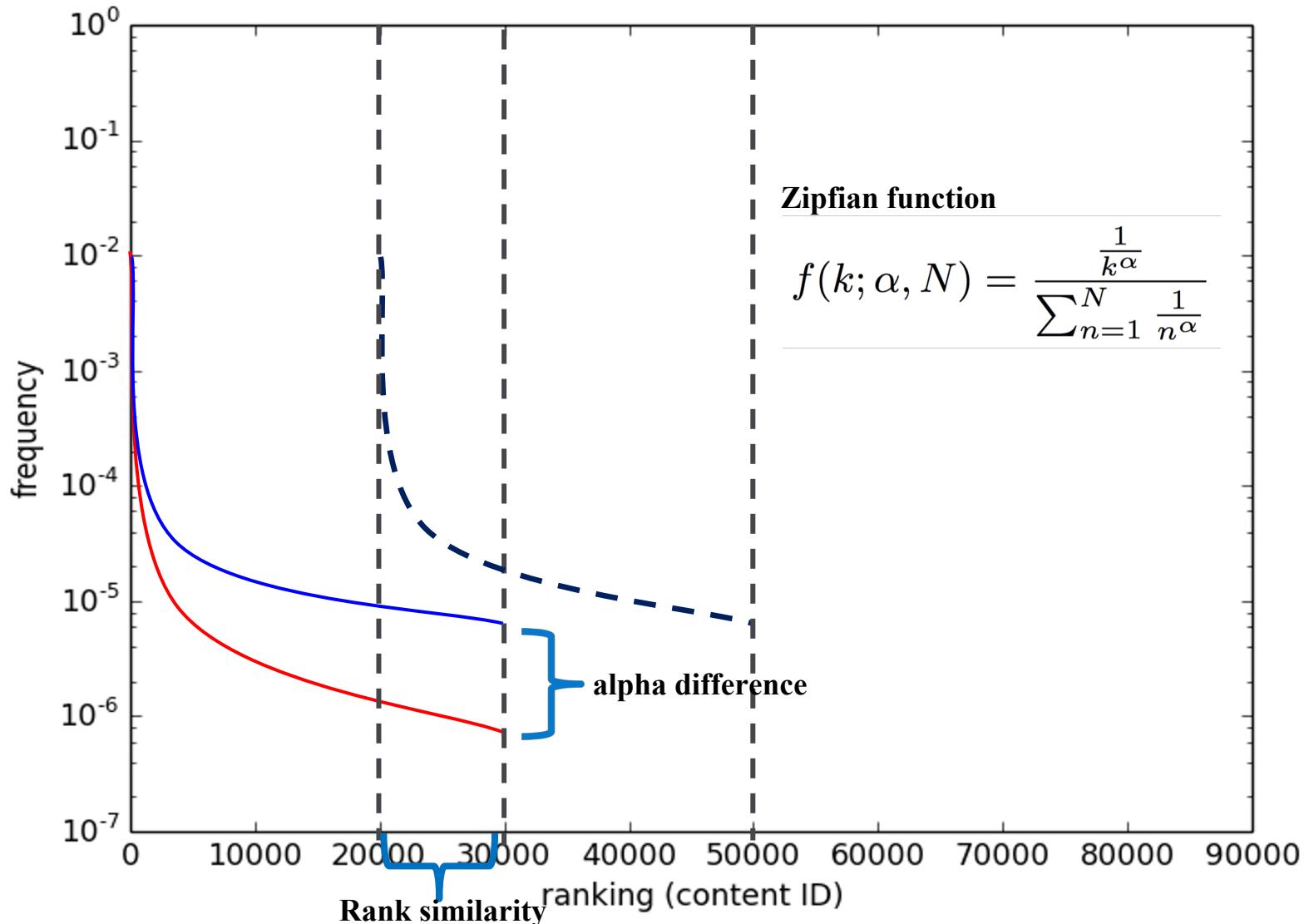


Grouping Users to Edge Resources

- Problem Statement:
 - Issue solved by this work
 - whether grouping users would improve cache performance
 - Future issues
 - clustering
 - optimization of cache improvement and forwarding latency

Evaluation

User Interest Distribution



Evaluation

Topology

Topology	Source nodes	Routers	Receivers
Tiscali	44	160	1636
Garr	13	27	291
Geant	13	32	328
Datacenter	1	16	160

Tiscali is parsed from Rocketfuel topology

Garr and Geant are topologies from The Internet Topology Zoo database

Datacenter is a simple 3 layer topology designed by ourselves.

Evaluation

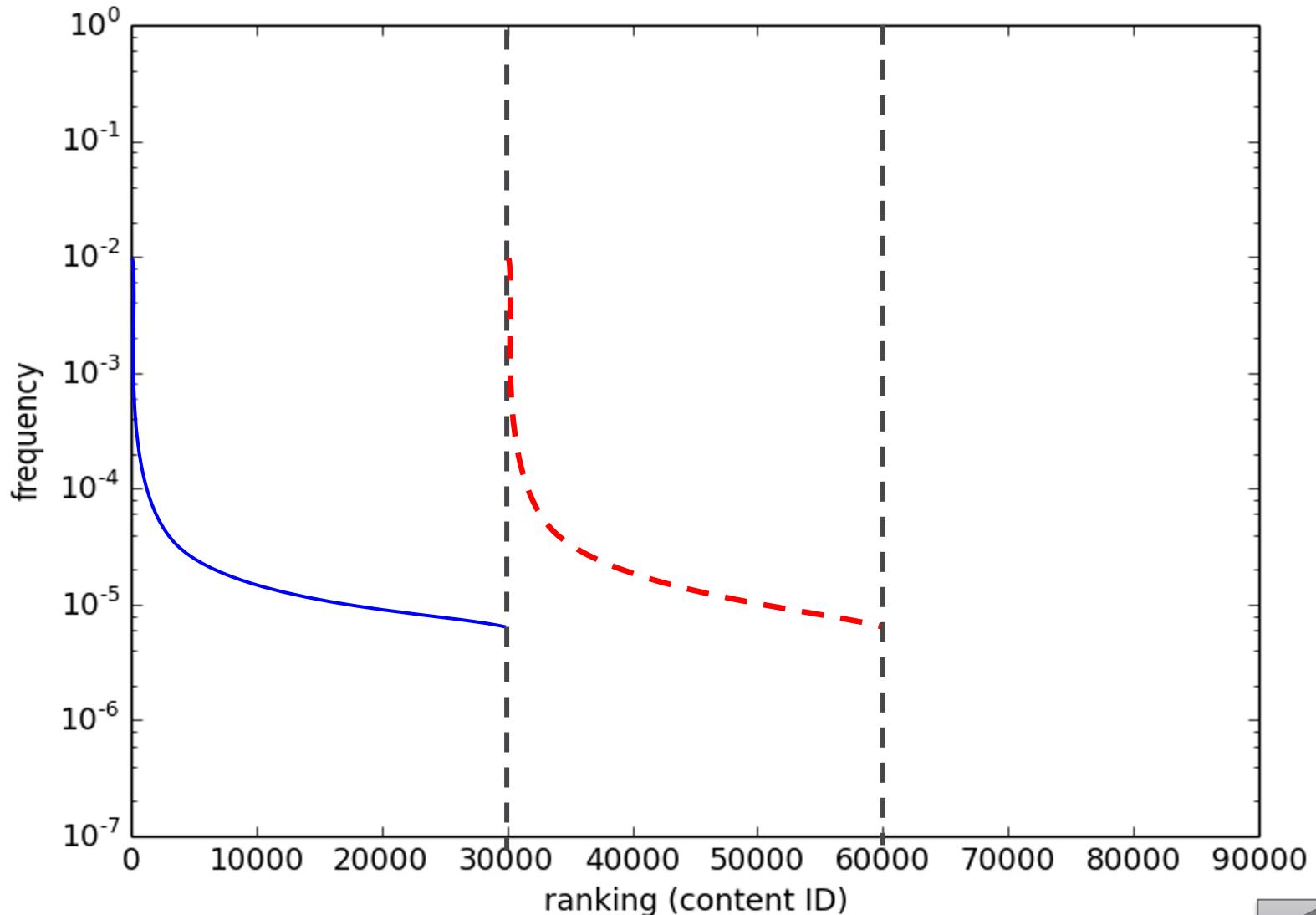
Workload Sample

Workload situations	number of ranking	number of alpha	number of distribution	annotation
<u>Diffrank</u> <u>(R)</u>	7	1	7	users have distributions with different rankings but same alpha
<u>GroupDiffrank</u> <u>(R-GR)</u>	1 (locally)	1 (locally)	1 (locally)	grouping the users with same ranking
<u>Alldiff</u> <u>(AR)</u>	7	6	42	users have distributions with different rankings and different alphas
<u>GroupAlldiffrank</u> <u>(AR-GR)</u>	1 (locally)	6	6	only grouping the users above with same rankings
<u>GroupAlldiff</u> <u>(AR-GAR)</u>	1 (locally)	1 (locally)	1 (locally)	grouping the users above with same rankings and alphas



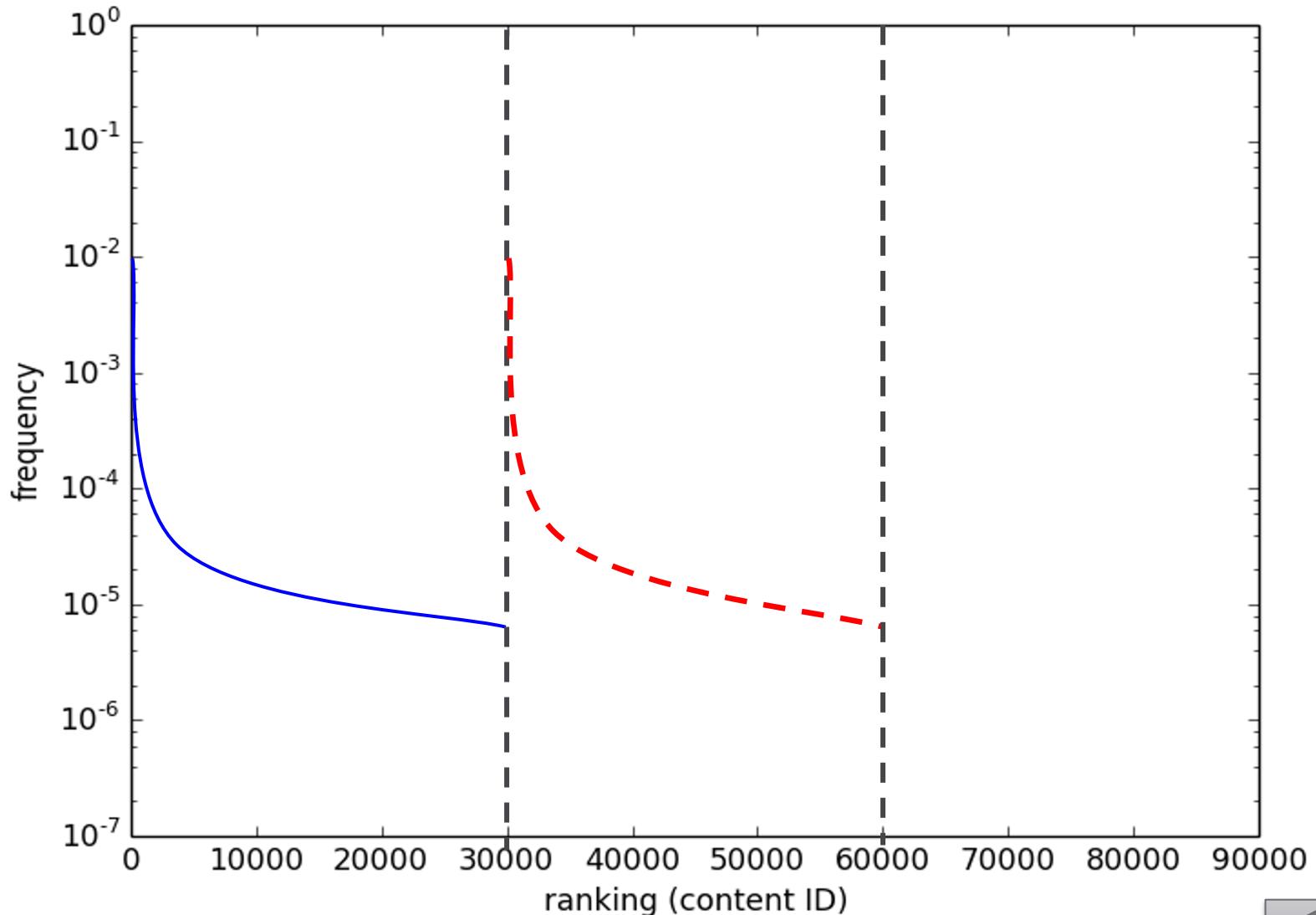
Evaluation

User Interest Distribution



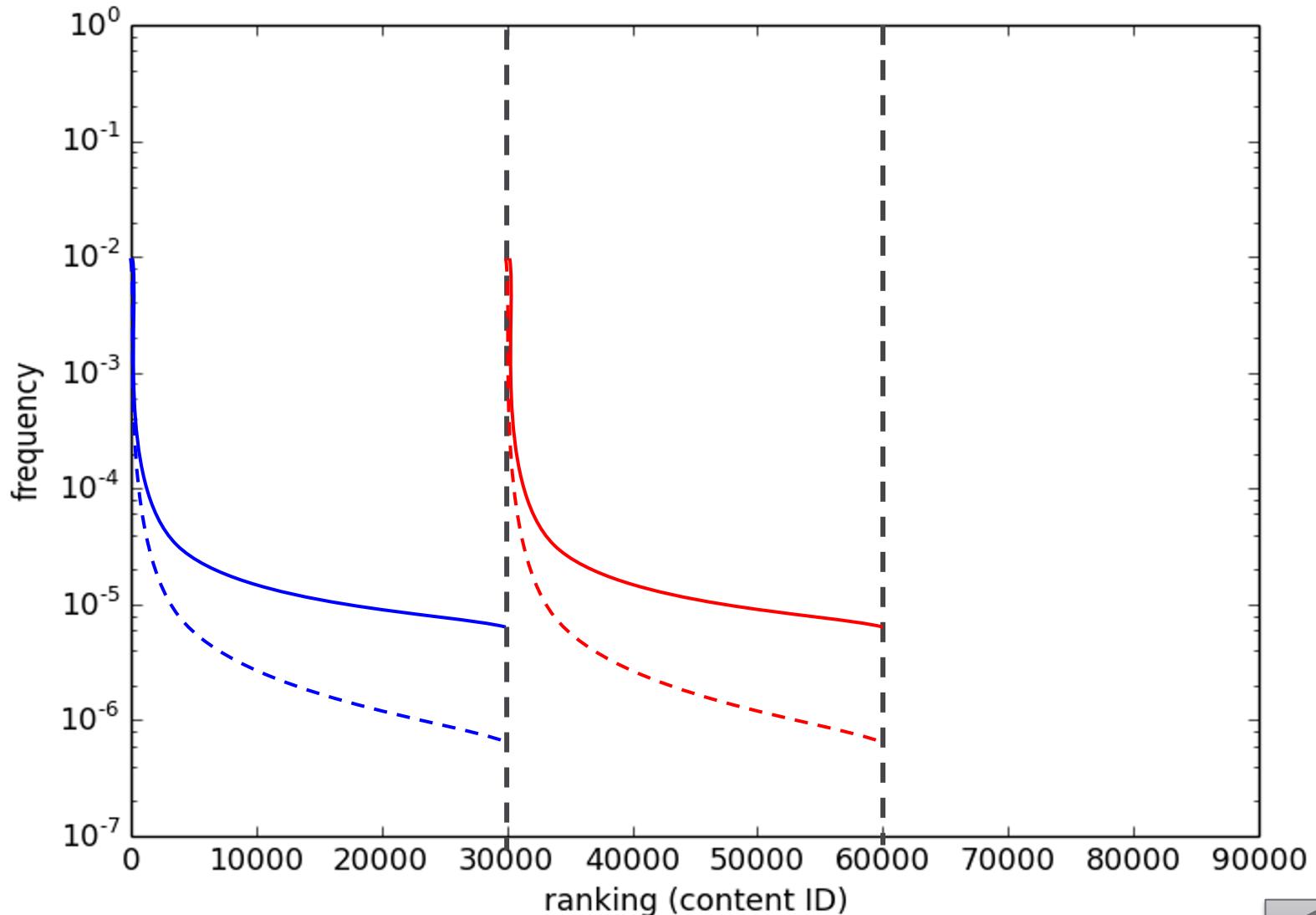
Evaluation

User Interest Distribution



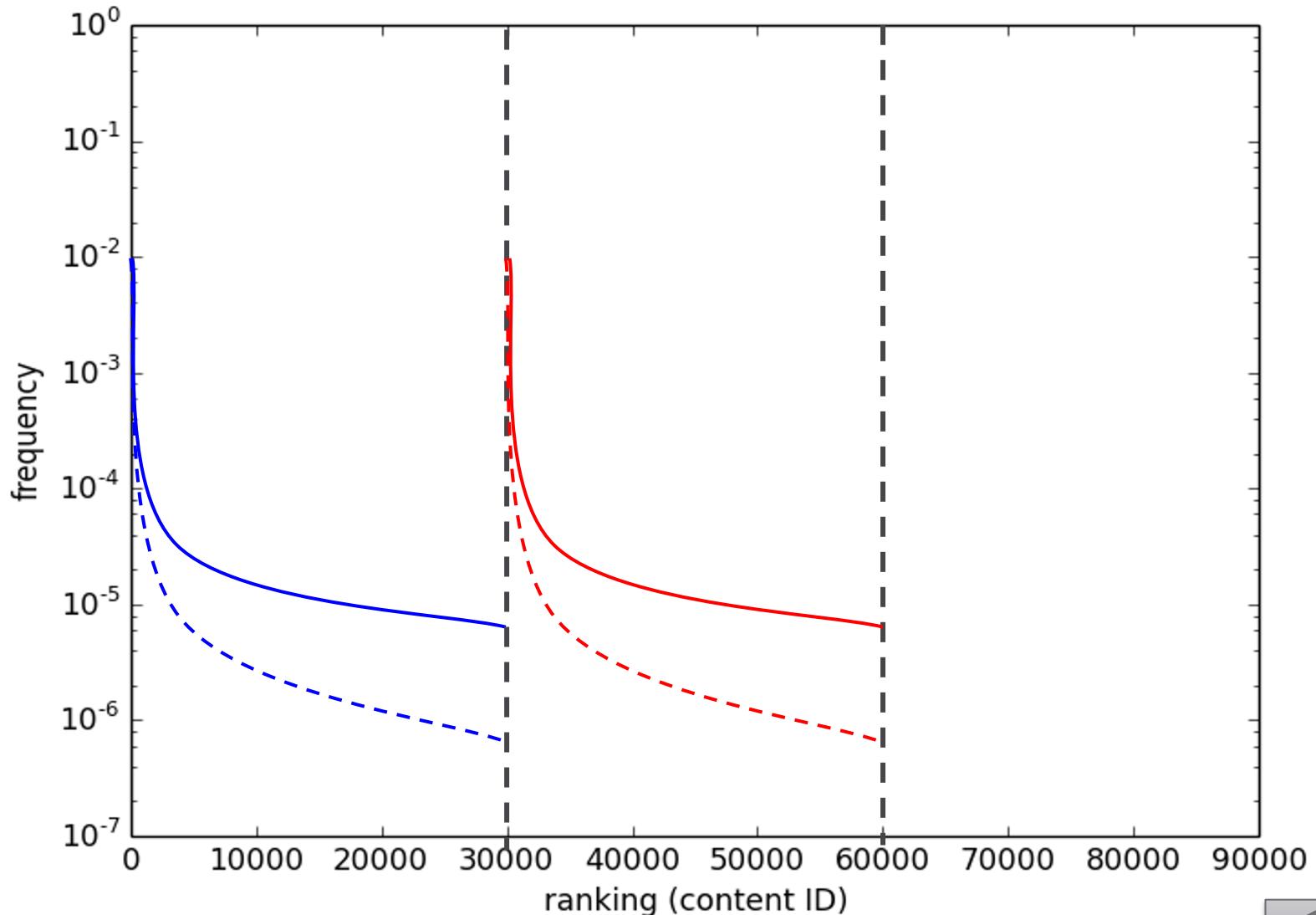
Evaluation

User Interest Distribution



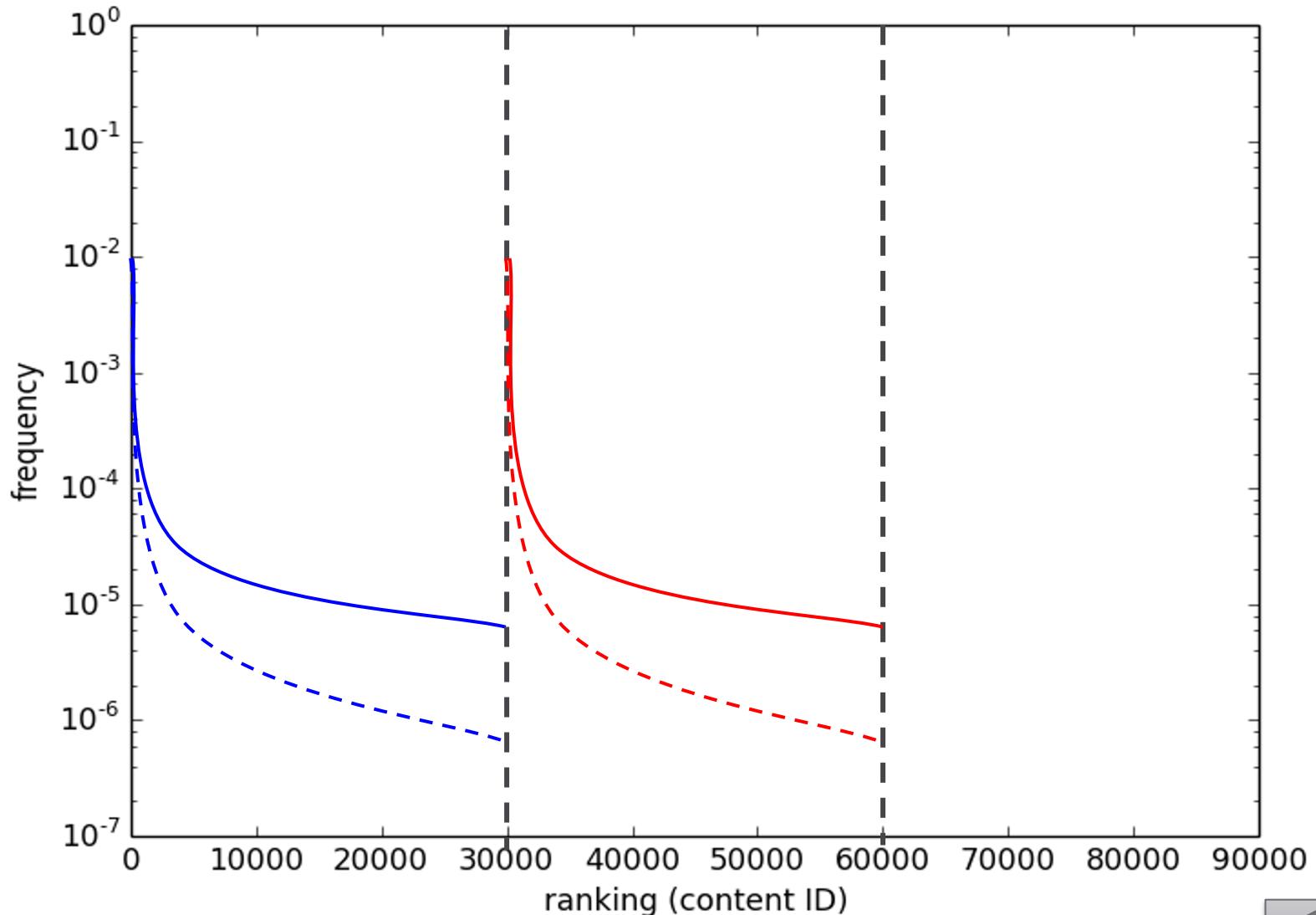
Evaluation

User Interest Distribution

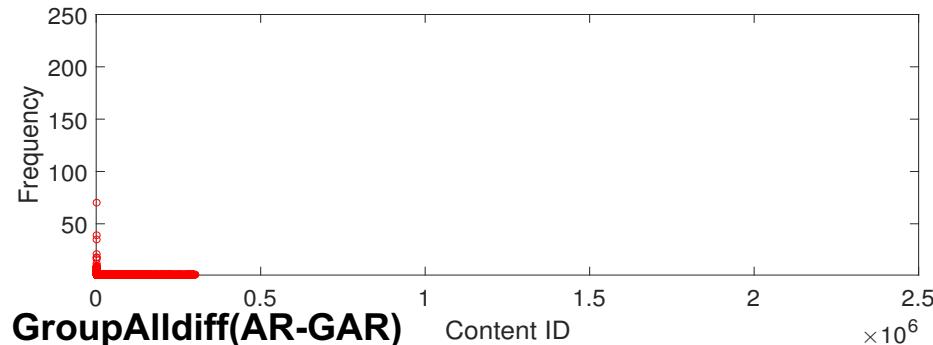
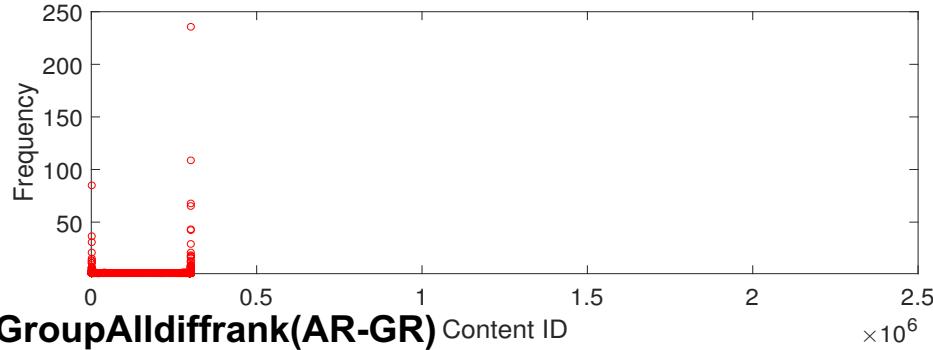
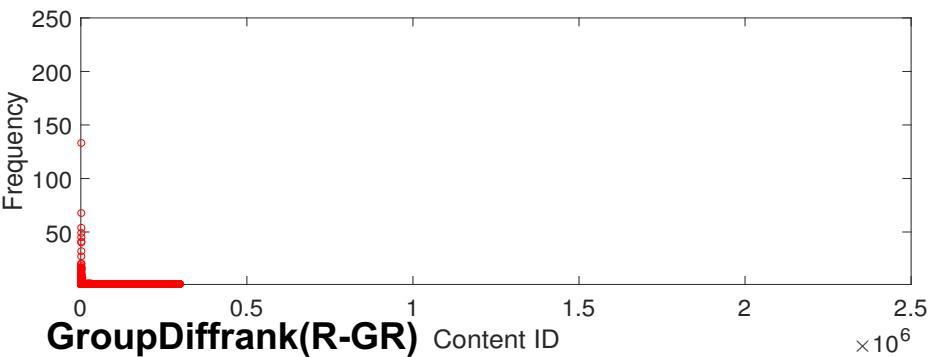
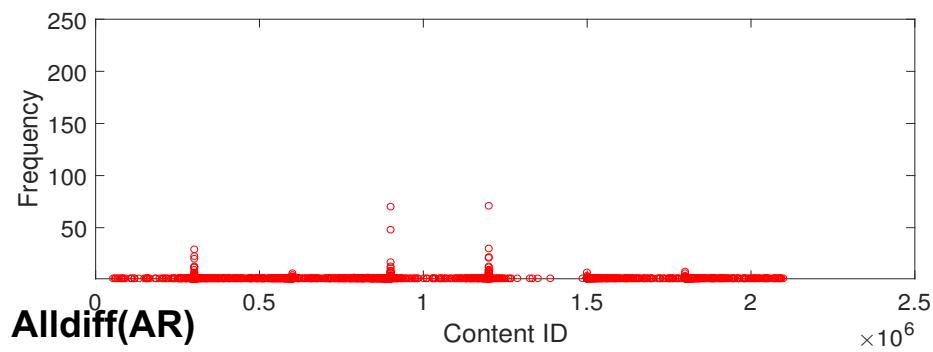
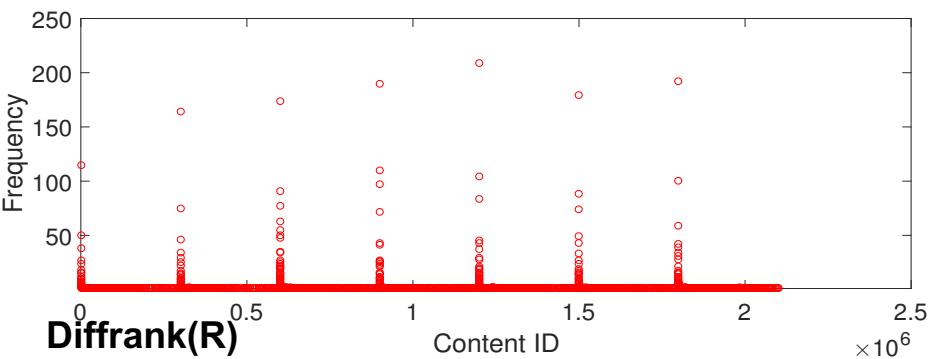


Evaluation

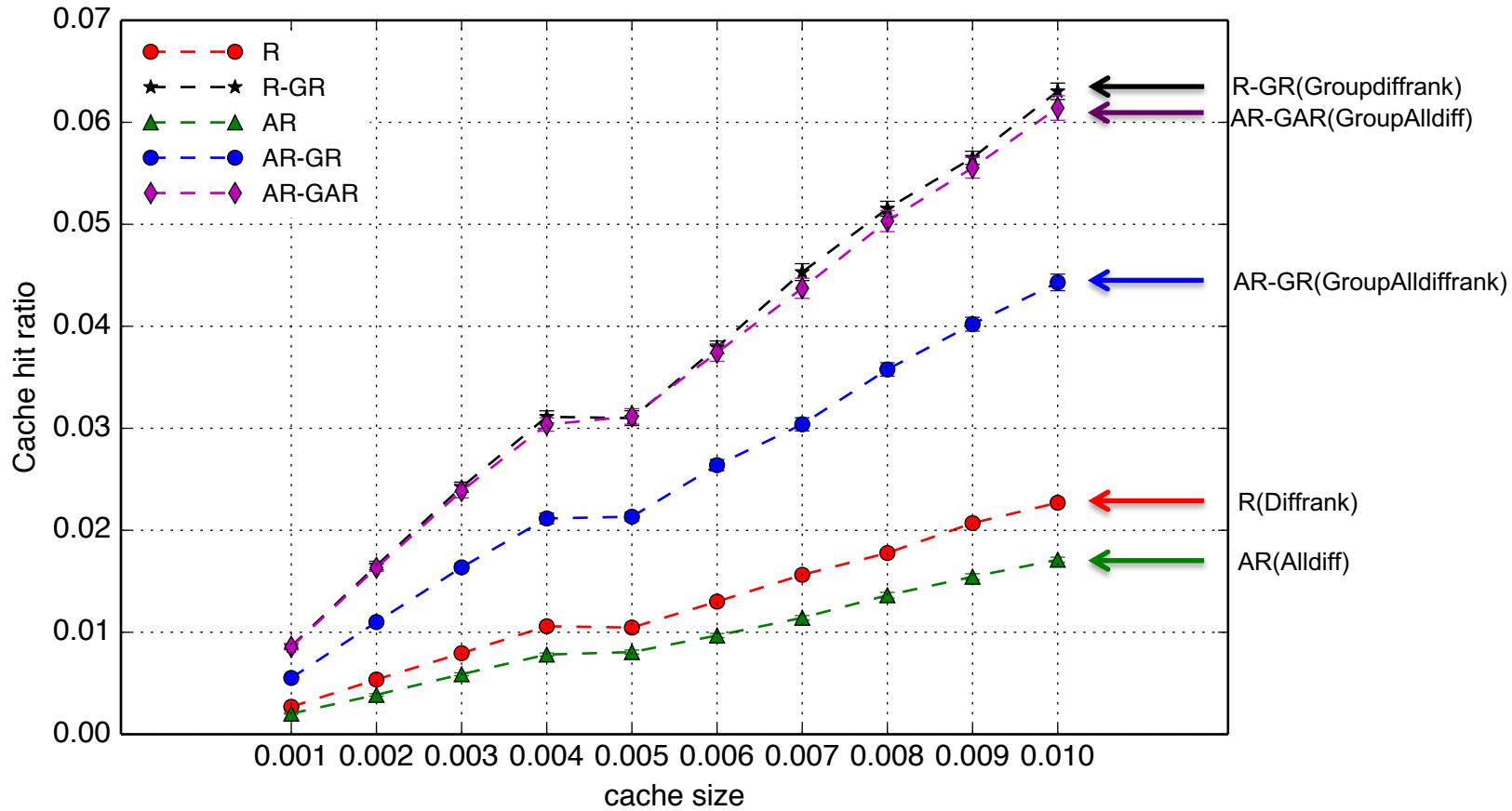
User Interest Distribution



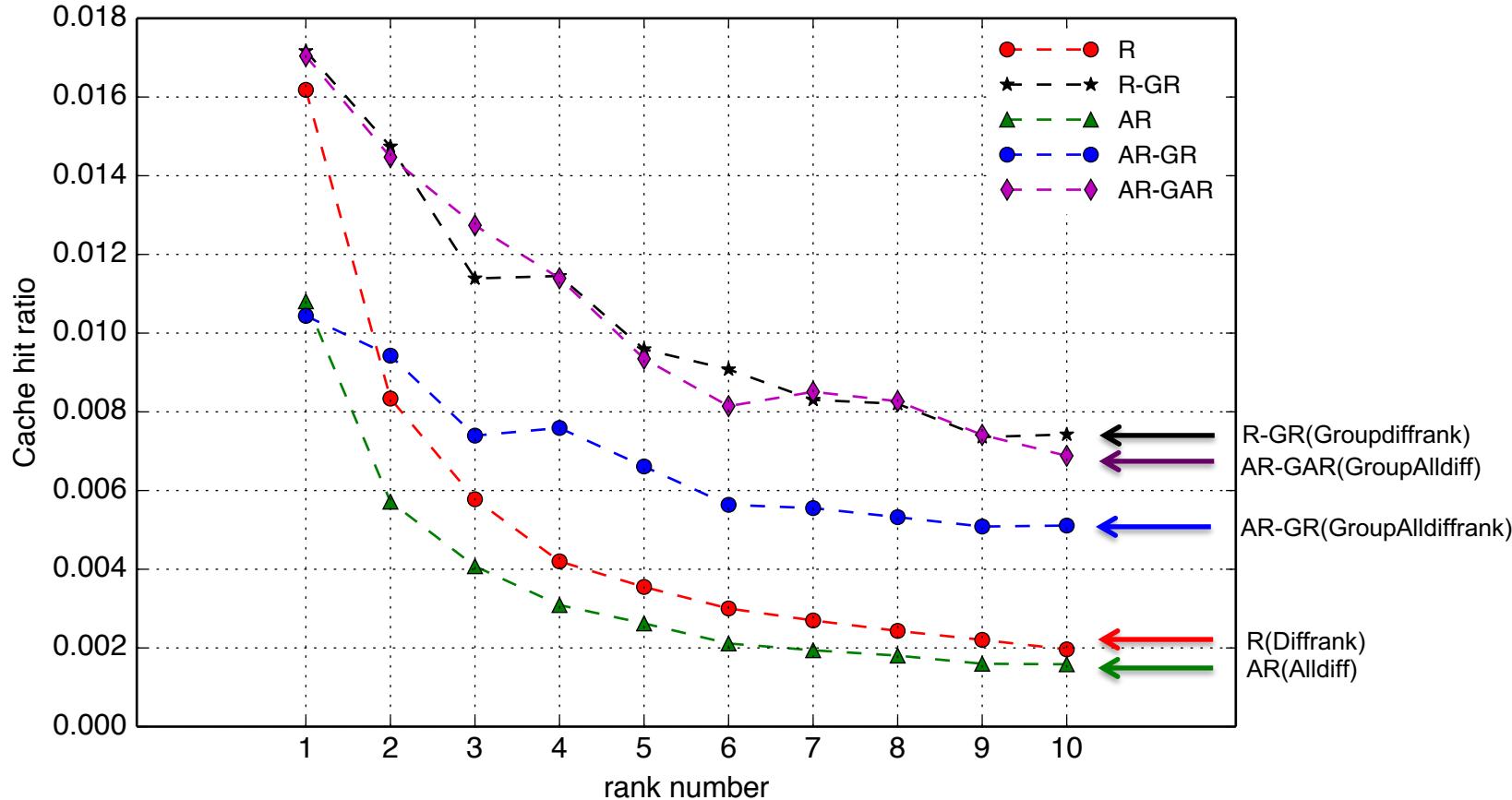
Requests seen by a single cache



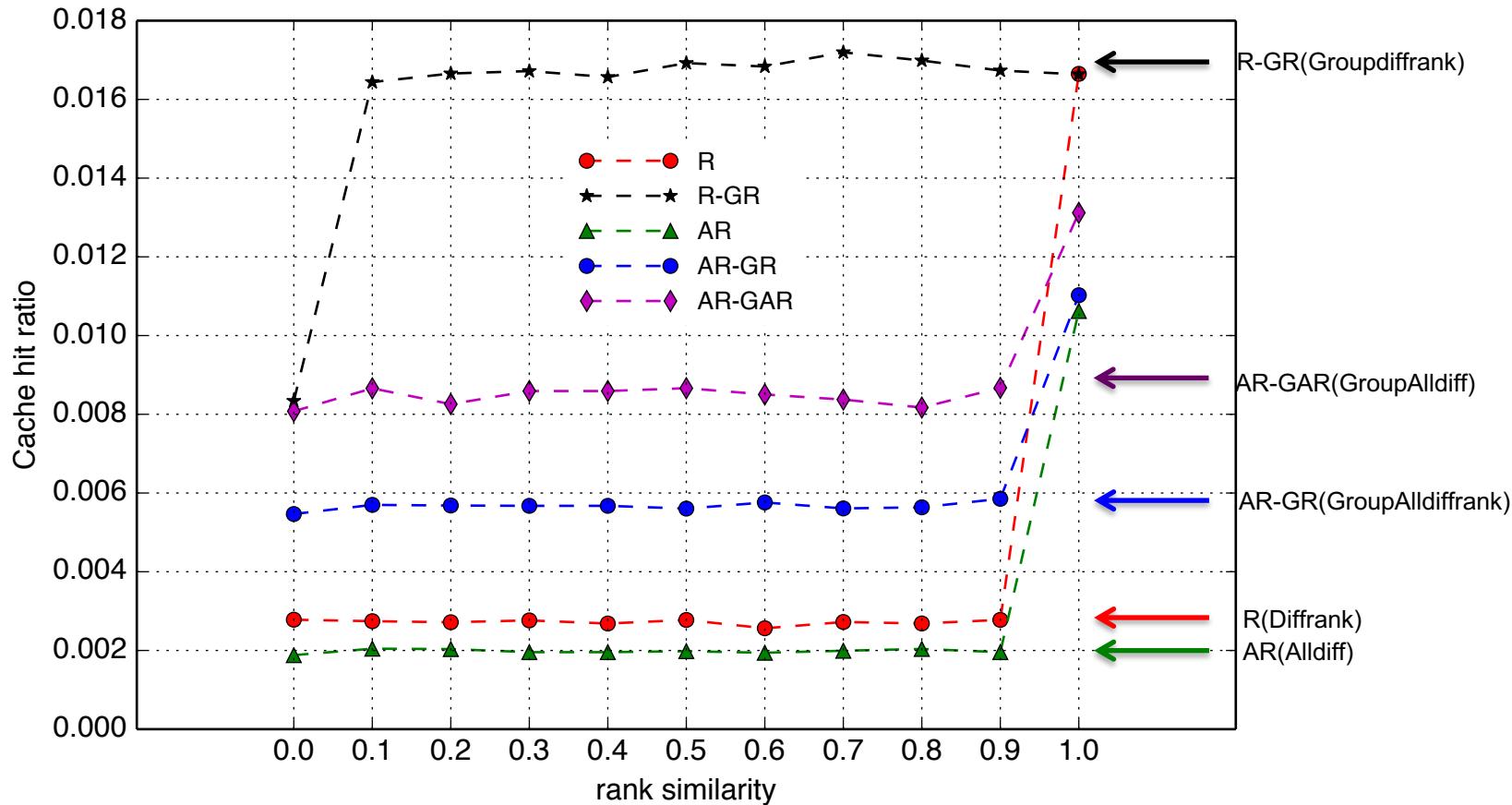
Cache hit ratio



Effect of rank number



Effect of rank similarity



Conclusion

- Result: Grouping users can improve cache hit ratio.
- Main contribution:
 - We propose a novel profiling strategy, where requests are grouped according to interest similarity.
 - Cache servers receive similar requests with higher probability. This increases cache hit rates and reduces redundant copies in caches.
 - After grouping, cached popular content can serve more users which saves cache resources and additional resources could be offered to nonmainstream users.
- Future work:
 - user clustering: scikit-learn, processing
 - optimization of retransmission delay and cache hit ratio improvement, processing.

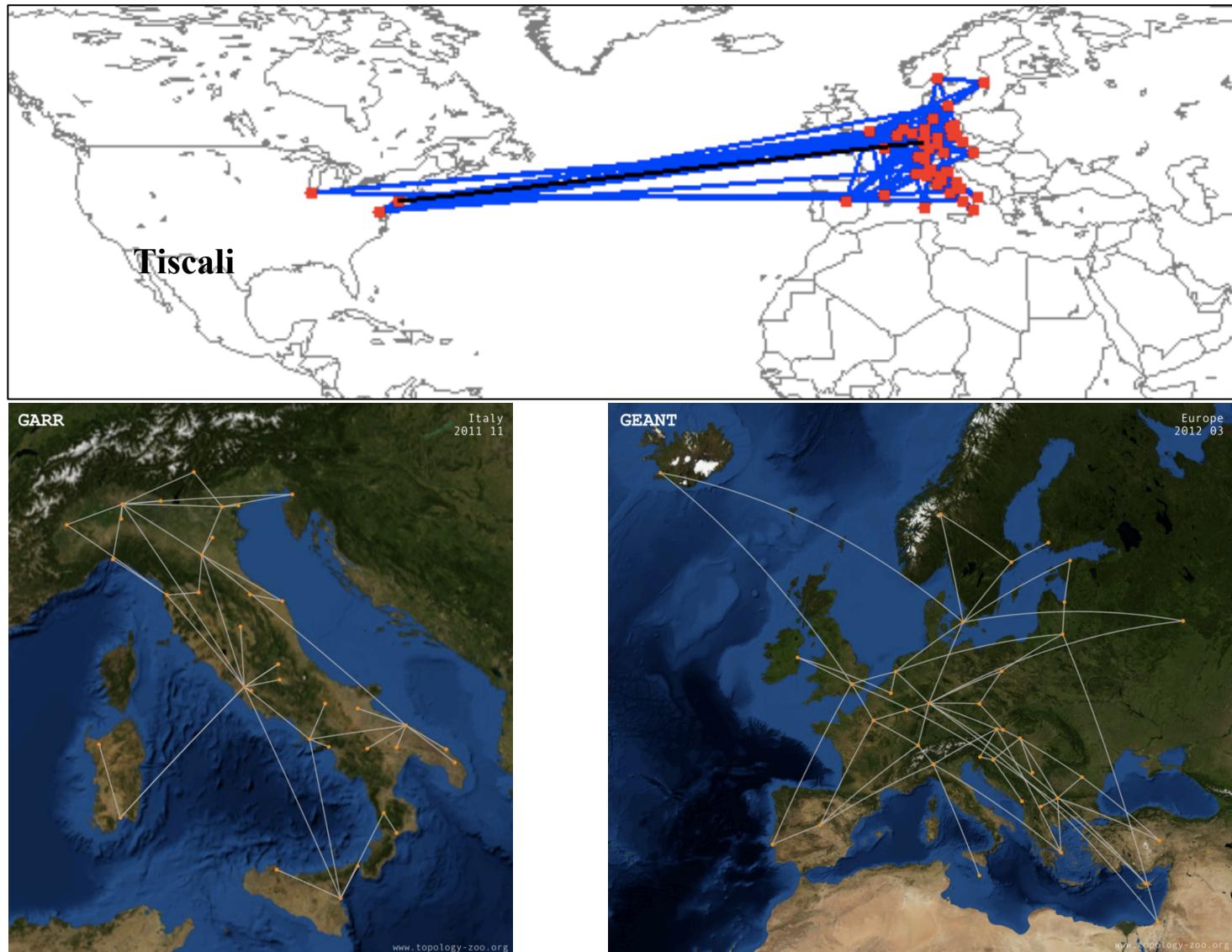
Q and A:

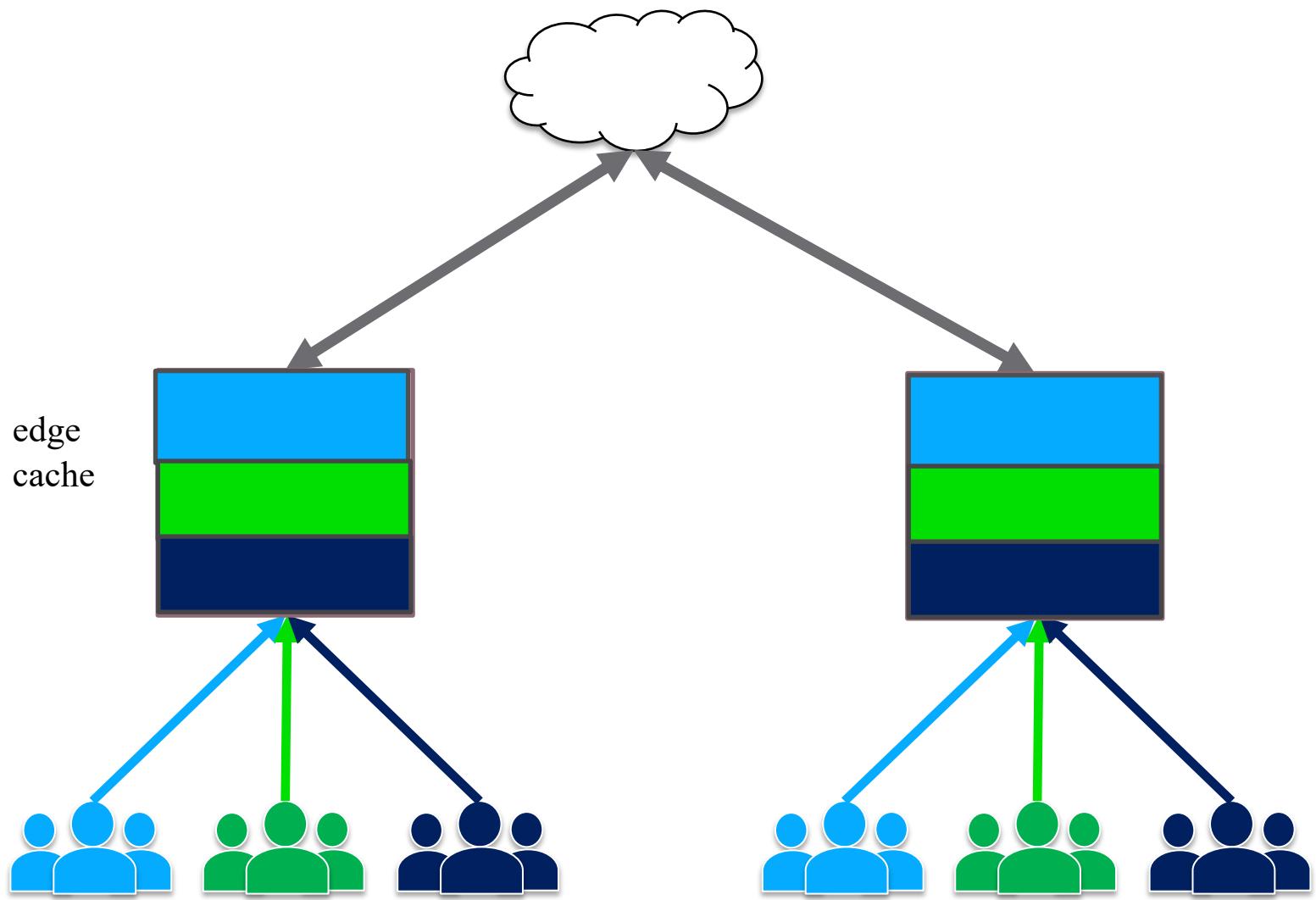
Thank you!

pengyuan.zhou@cs.helsinki.fi

Evaluation

Topology





User Interest

Recommender systems also contribute to steady user interest

- Recommender systems help users find content of interest faster
- But, researchers also find out that it may also manipulate users
 - *Adomavicius, Gediminas, et al. "Do recommender systems manipulate consumer preferences? A study of anchoring effects." Information Systems Research 24.4 (2013): 956-975.*
 - *Anchoring or focalism is a cognitive bias that describes the common human tendency to rely too heavily on the first piece of information offered (the "anchor") when making decisions*