Mobility-Aware Caching for Content-Centric Wireless Networks

Jun ZHANG

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Outline

Introduction

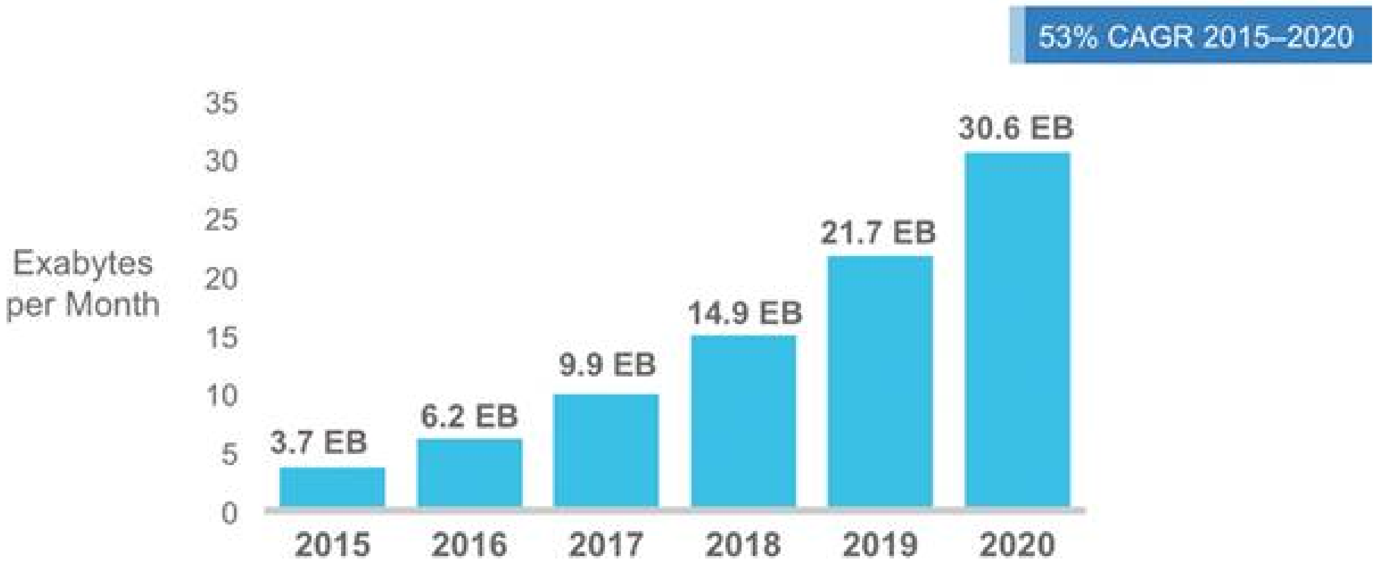
Exploiting User Mobility in Cache-Enabled Content-Centric Wireless Networks (CCWNs)

Mobility-Aware Caching Content Placement

Conclusions

Grand Challenge for Wireless Networks

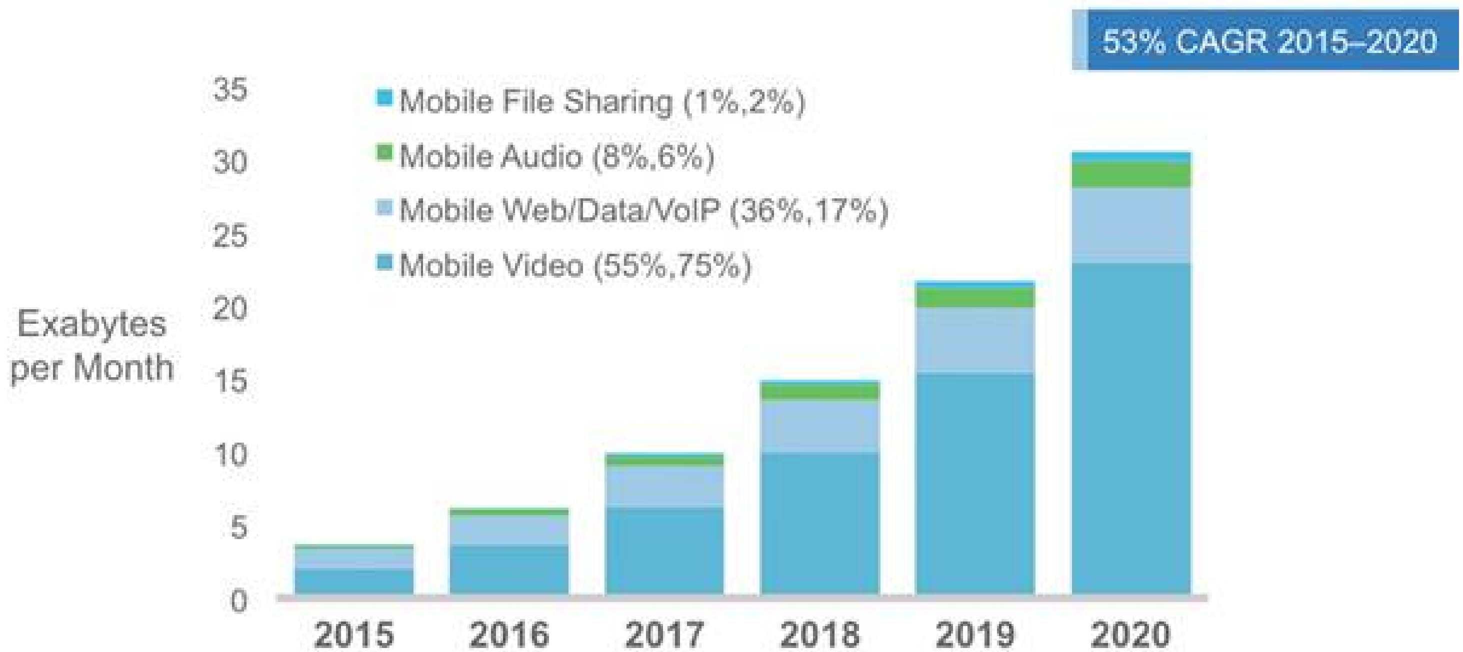
Exponential Global Mobile Traffic Growth



Source: Cisco VNI Mobile, 2016

1 EB (Exabyte) = 10^{18} B
= 1 billion GB

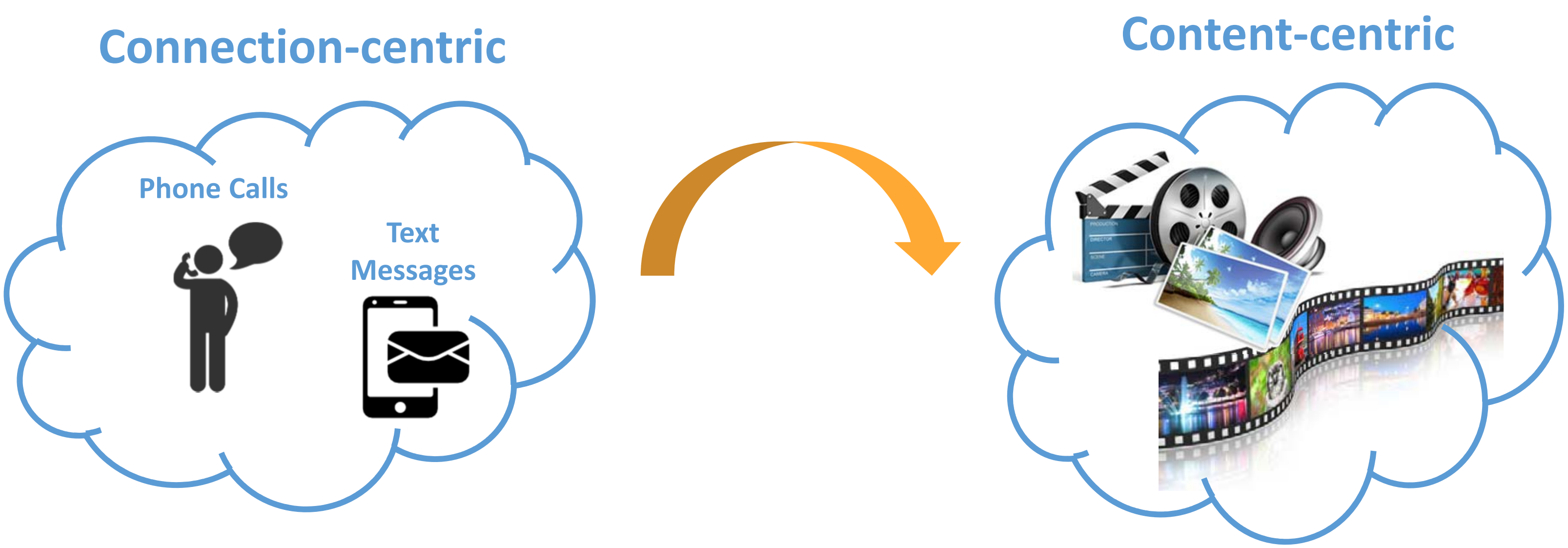
A Closer Look



Source: Cisco VNI Mobile, 2016

Mobile Video Will Generate Three-Quarters of Mobile Data Traffic by 2020

“Connection-Centric” to “Content-Centric”



<u>Opportunities</u>

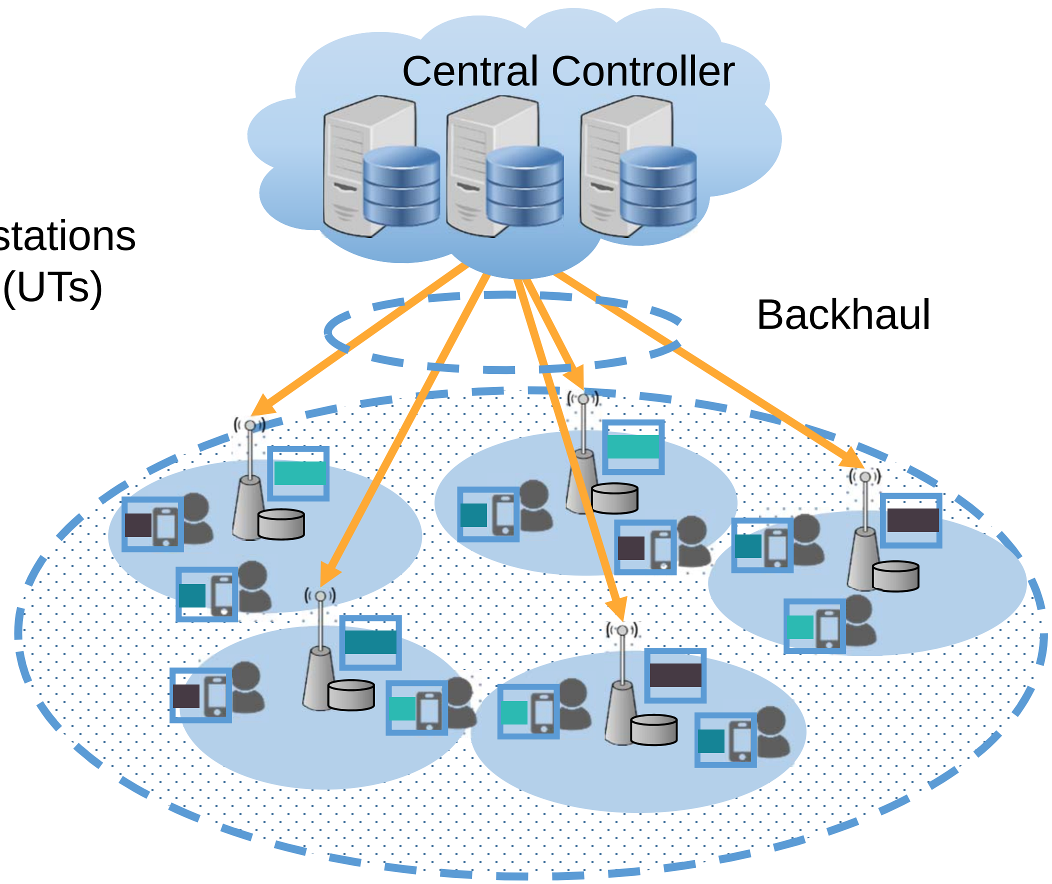
- Predictive demand
- Reusable content
- Delay tolerant, variable quality

Cache-Enabled Content-Centric Wireless Networks (CCWNs)

Abundant caching at the wireless edge, i.e., base stations (BSs) and user terminals (UTs)

Caching popular contents at the wireless edge

- Reduce the demand of backhaul links
- Lower delay
- Enable cooperation
- Improve energy efficiency
- ……



Cache – An Old Idea Reborn

Virtual memory hierarchies in CPU design

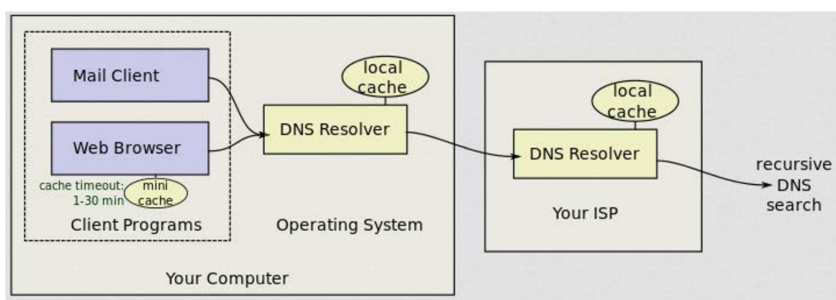
- The origin of cache

Web caching for content delivery networks (CDNs)

- To reduce traffic load, reduce delay
- Examples: Netflix, Akamai

Inquiry caching in domain name systems (DNS)

- To reduce delay and DNS server load



[ISIT 15 Tutorial, Maddah-Ali, Niesen]

Research Problems in Wireless Caching

Backhaul-Cache Tradeoff [Peng, Z, Song, Khaled (‘16)]

Cache Content Placement [Peng, Shen, Z, Khaled (‘15)]
[Liu, Bai, Z, Khaled (‘16)]

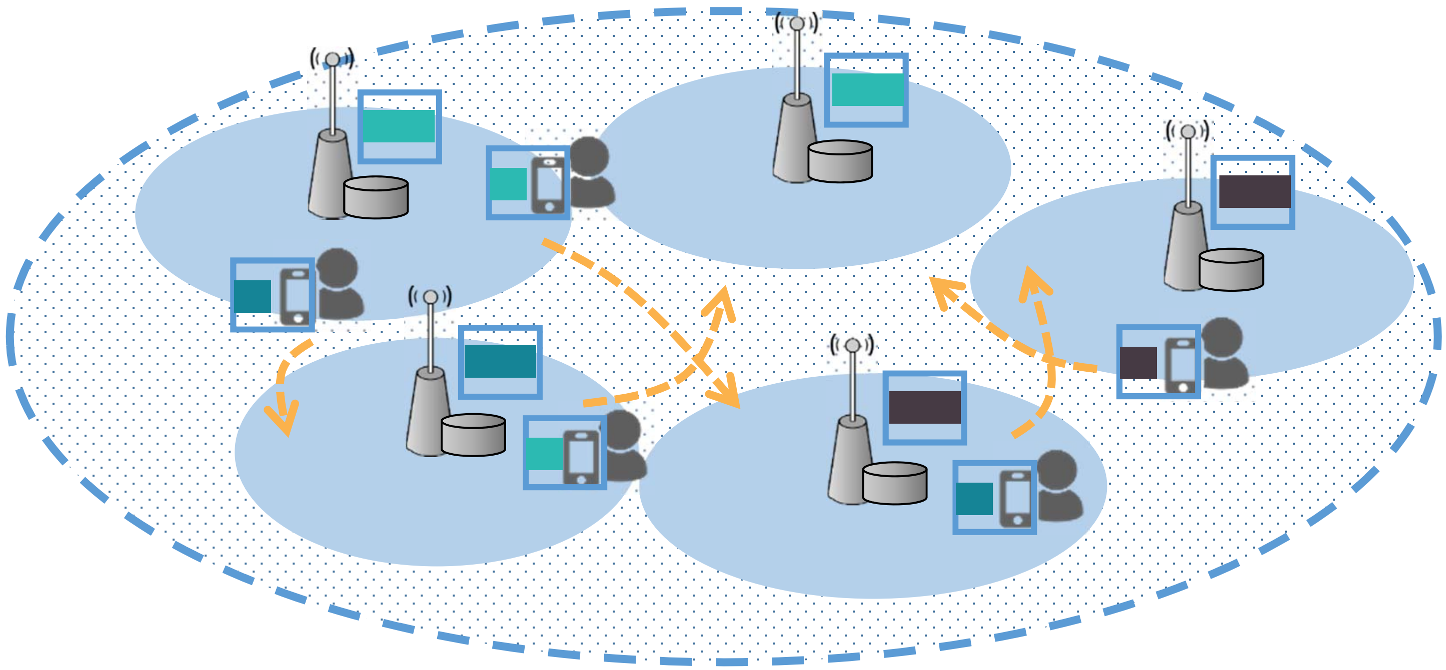
Caching Size Allocation [Peng, Z, Song, Khaled (‘16)]

Joint Data Assignment and Beamforming
[Peng, Shen, Z, Khaled (‘14)]

Other Issues:
Caching Content Update, Coded Caching…

Common Assumption: Fixed network topology

Uniqueness in CCWNs: User Mobility

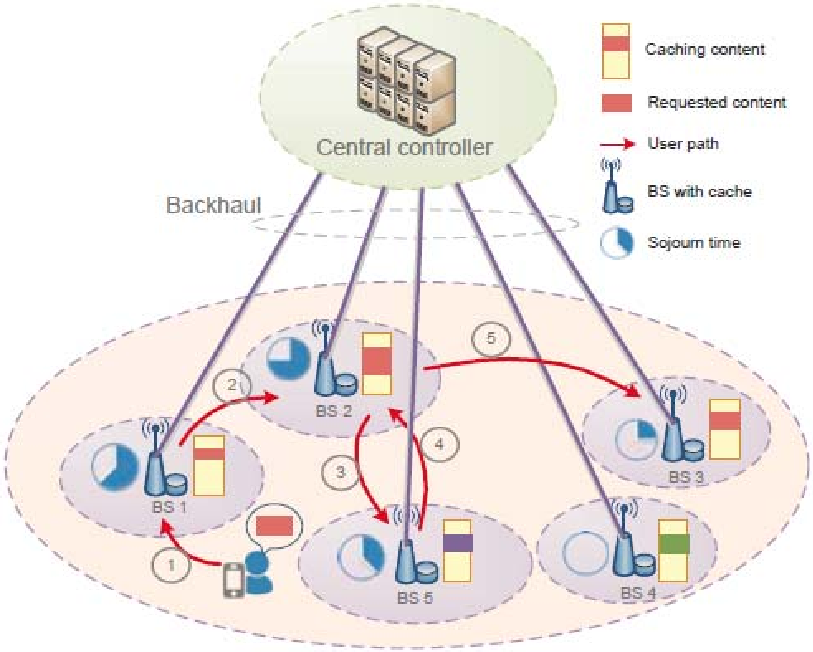


No fixed transmission link — Additional difficulties

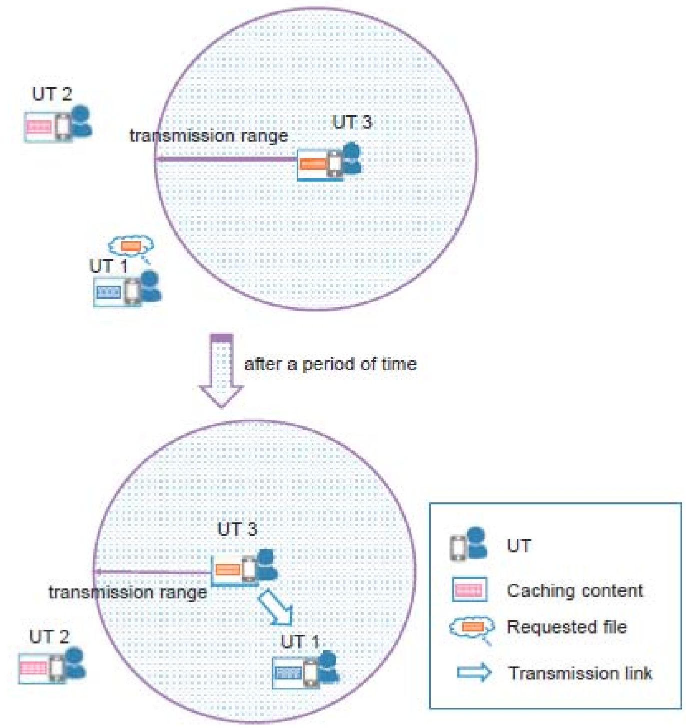
New opportunities — More potential transmission links

Local caching → **Global benefit**

Mobility-Aware Caching

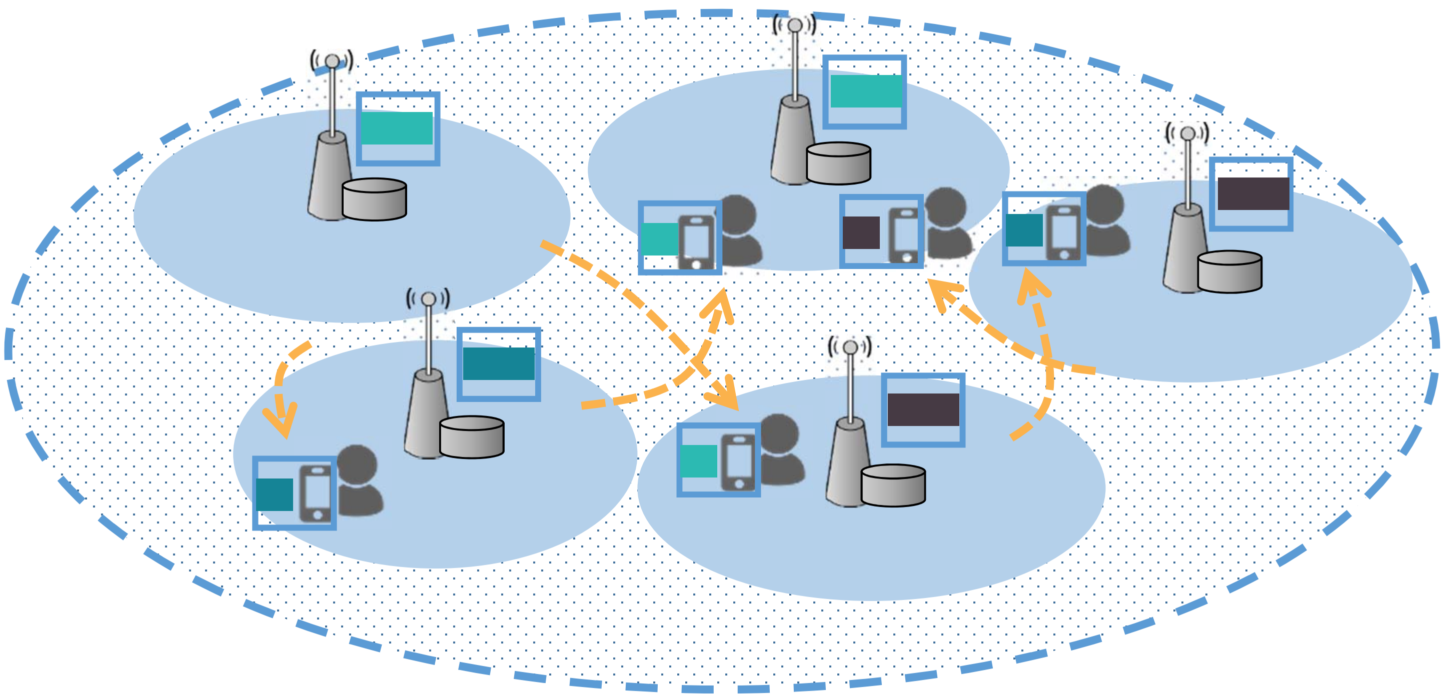


Caching at BSs



Caching at UTs

Exploiting User Mobility in Cache-Enabled CCWNs

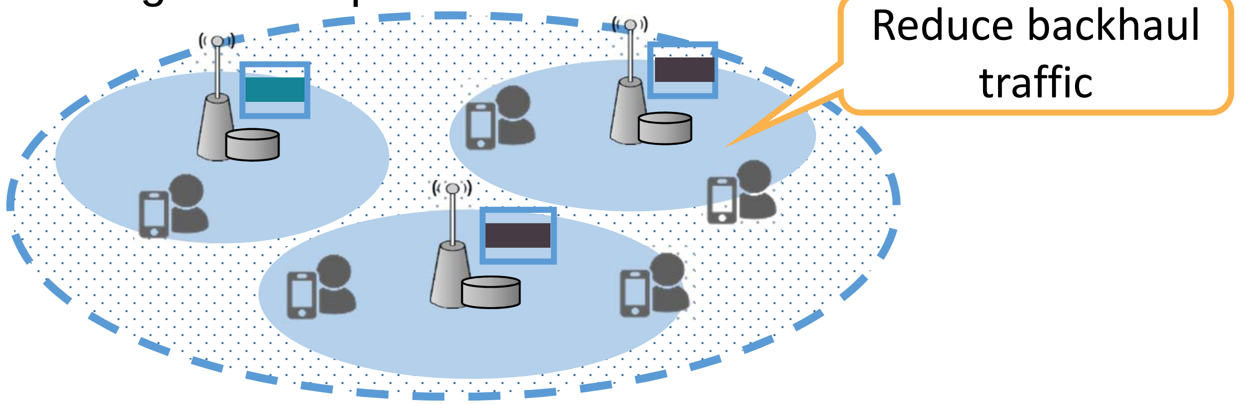


[Ref] R. Wang, X. Peng, J. Zhang, and K. B. Letaief, "Mobility-aware caching for content-centric wireless networks: Modeling and methodology," IEEE Commun. Mag., to appear.

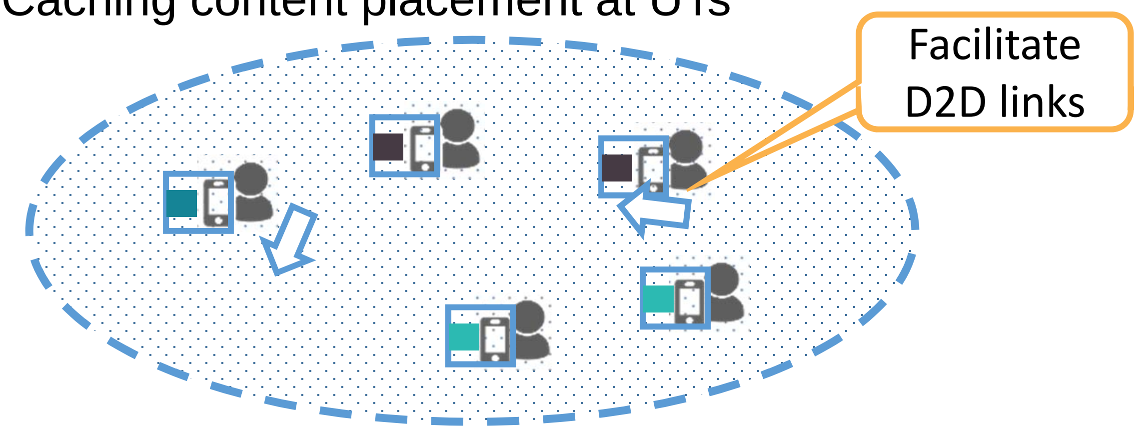
Key Design Problems of Caching in CCWNs

Caching content placement

Caching content placement at BSs



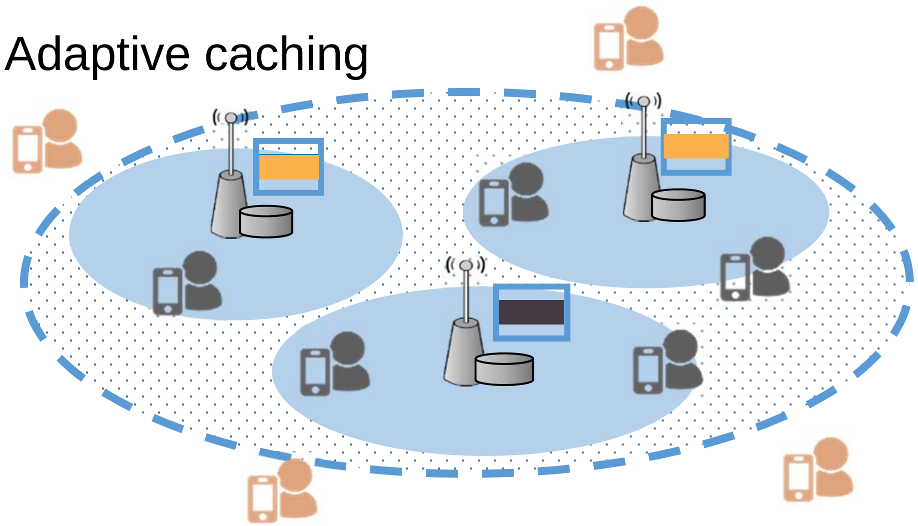
Caching content placement at UTs



Key Design Problems of Caching in CCWNs

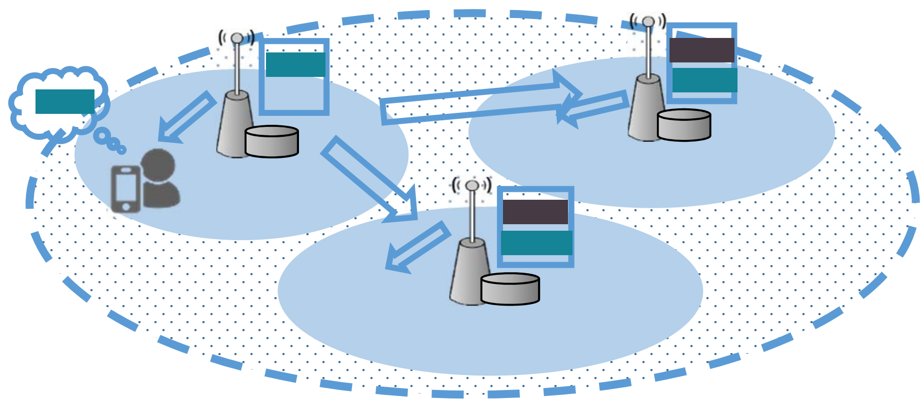
Caching content update

Adaptive caching



Replace outdated content

Proactive caching

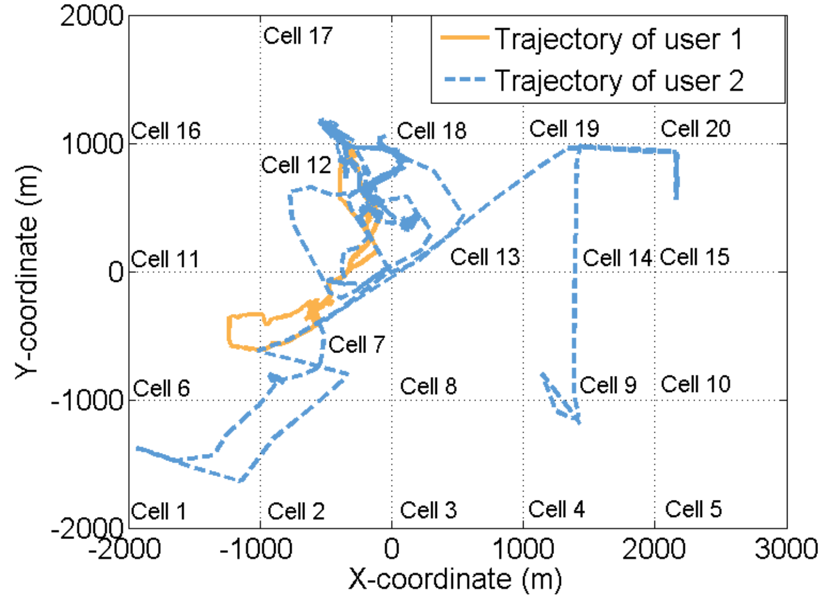


Pre-fetching the requested content

Modeling User Mobility Patterns – Spatial Properties

Spatial Properties
related to the
physical locations

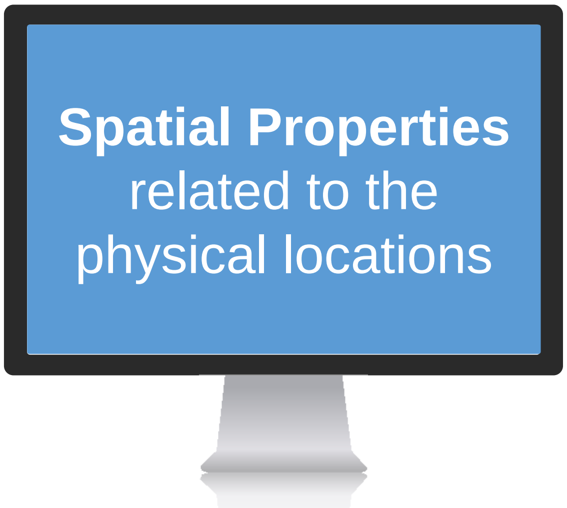
User trajectory: moving path

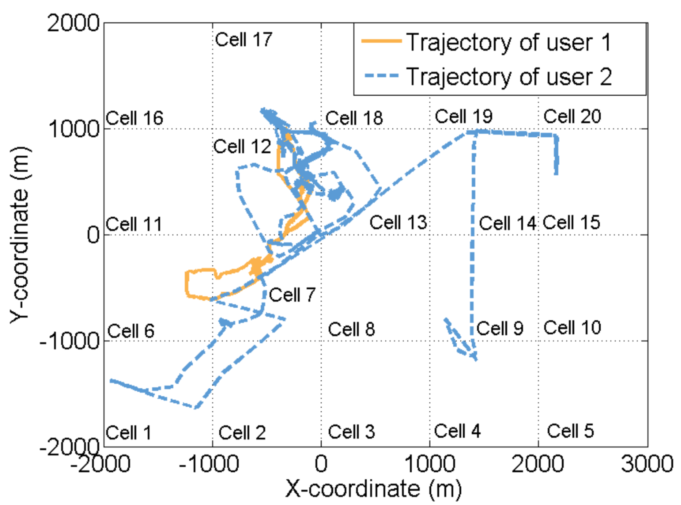


Related Model

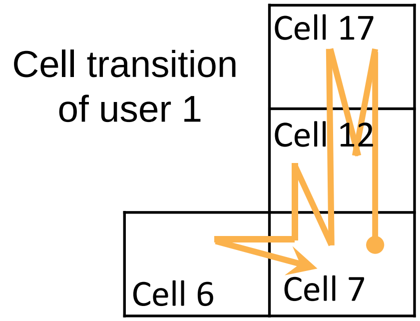
Random waypoint model [Bettstetter '04]

Modeling User Mobility Patterns – Spatial Properties





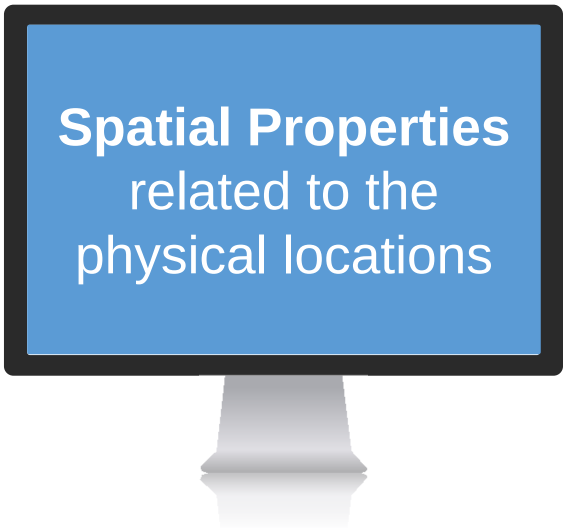
<u>Cell transition</u>: move from one cell to another



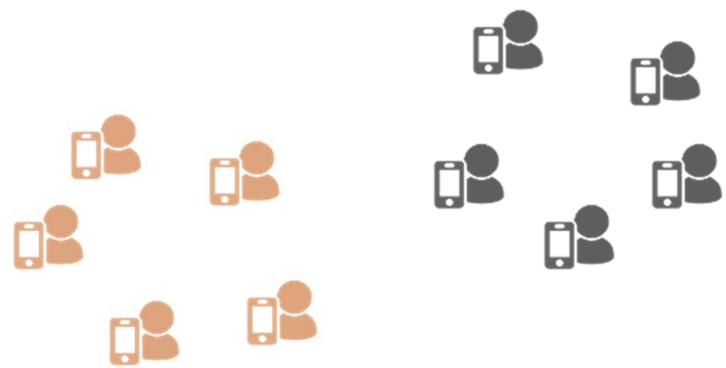
Related Model

Markov chain model [Lee '06]

Modeling User Mobility Patterns – Spatial Properties



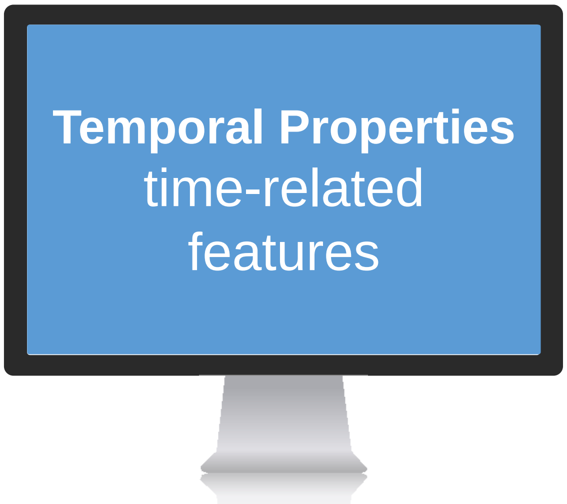
Social group: may move together
E.g., tour group, schoolmates



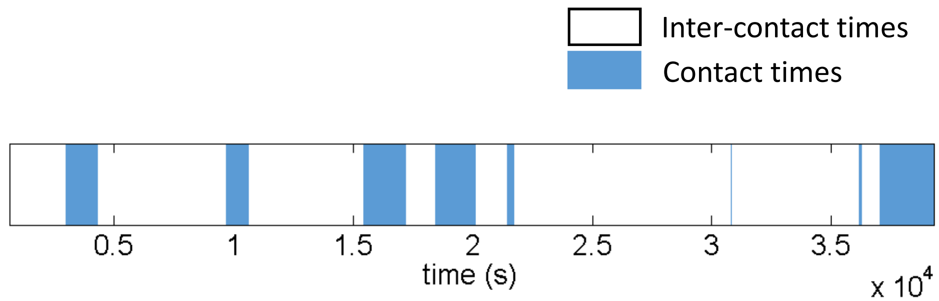
Related Model

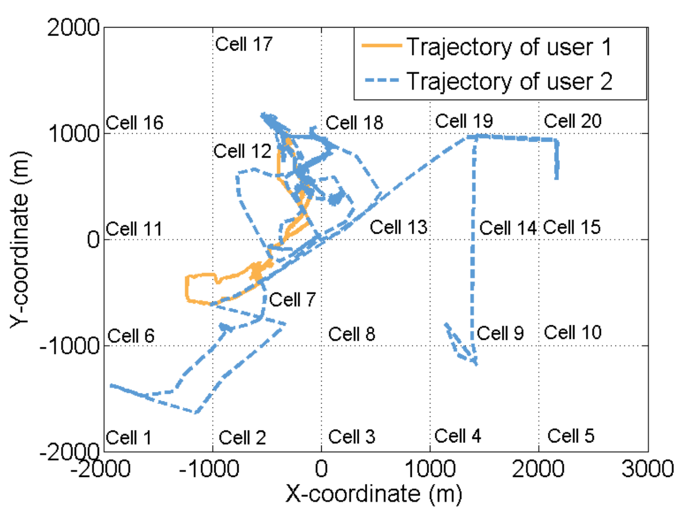
Detecting user mobile groups and characterization of group mobility properties [Nunes ’15]

Modeling User Mobility Patterns – Temporal Properties



User inter-contact time: frequency and duration that two users are connected

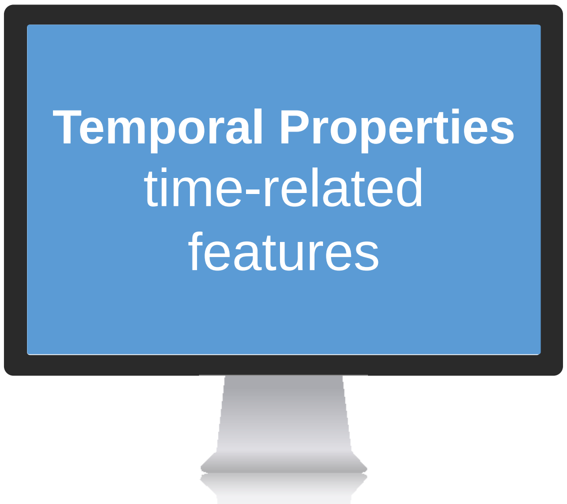




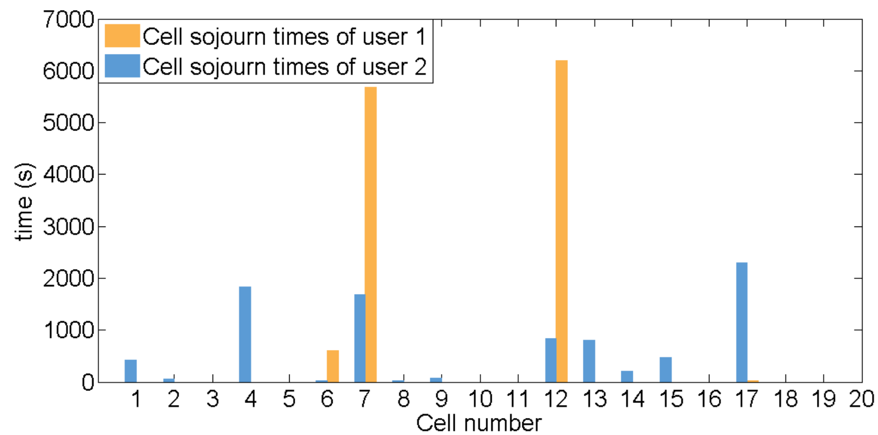
Related Model

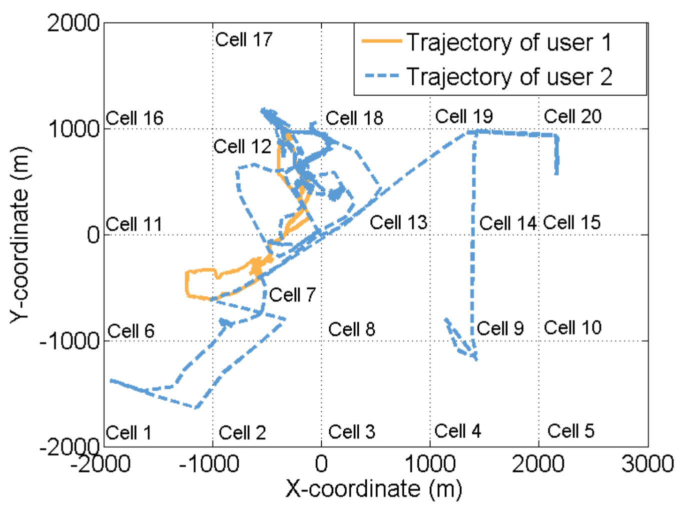
A Poisson Process to model the arrival of contact times [Conan '08]

Modeling User Mobility Patterns – Temporal Properties



<u>Cell sojourn time</u>:
time duration of a user stay in a given cell

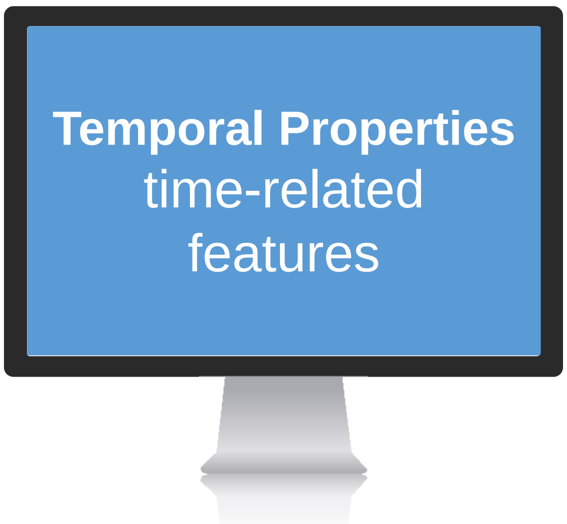




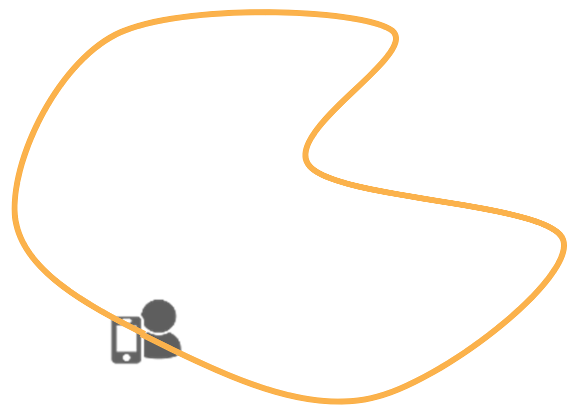
Related Model

An approach to obtain the sojourn time distributions [Lee '06]

Modeling User Mobility Patterns – Temporal Properties



Return time:
the time to return to a previous visited location



Related Model

Distribution of the return time was measured [Gonzales '08]

Exploiting Mobility for Caching in CCWNs

Mobility-aware caching content placement at BSs

- User trajectory ⇨ Serving BSs, Transmit distance
- Cell transition ⇨ Serving BSs
- Cell sojourn time ⇨ Serving time of each BS

Mobility-aware caching content placement at UTs

- Social group ⇨ More opportunities to establish D2D links
- User inter-contact time ⇨ Frequency and duration to establish D2D links

Exploiting Mobility for Caching in CCWNs

Mobility-aware adaptive caching

Social group

Return time

Mobility-aware proactive caching

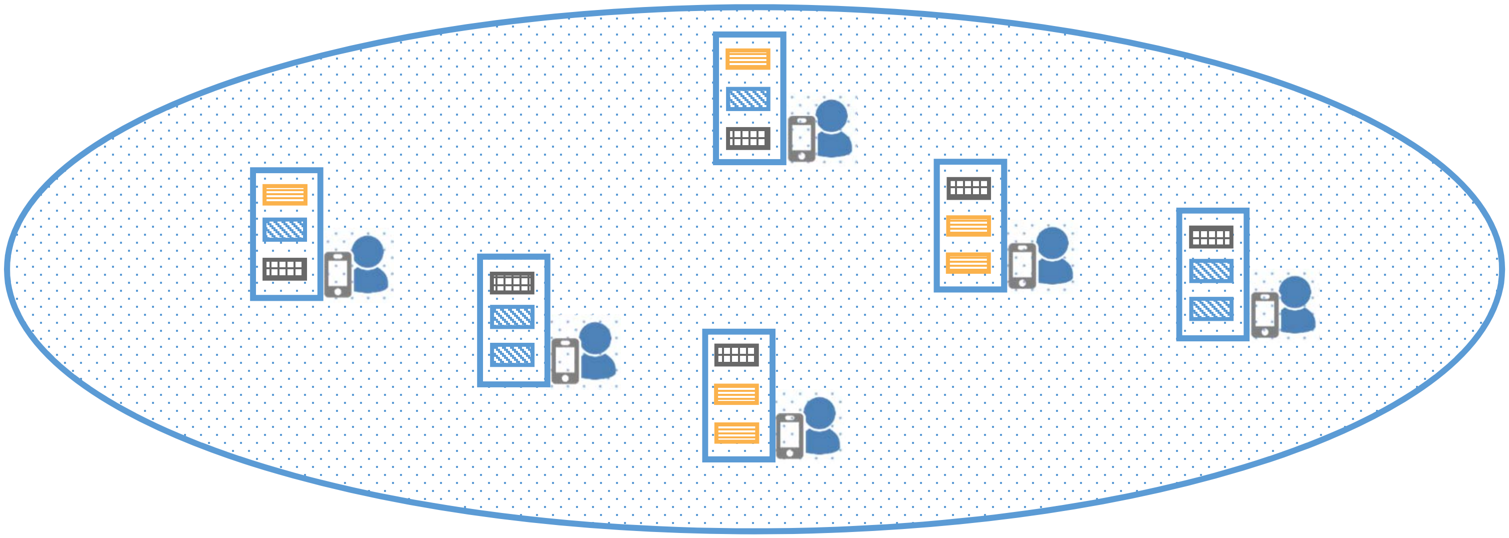
User trajectory

Cell transition

⇨ Predict the future serving BSs

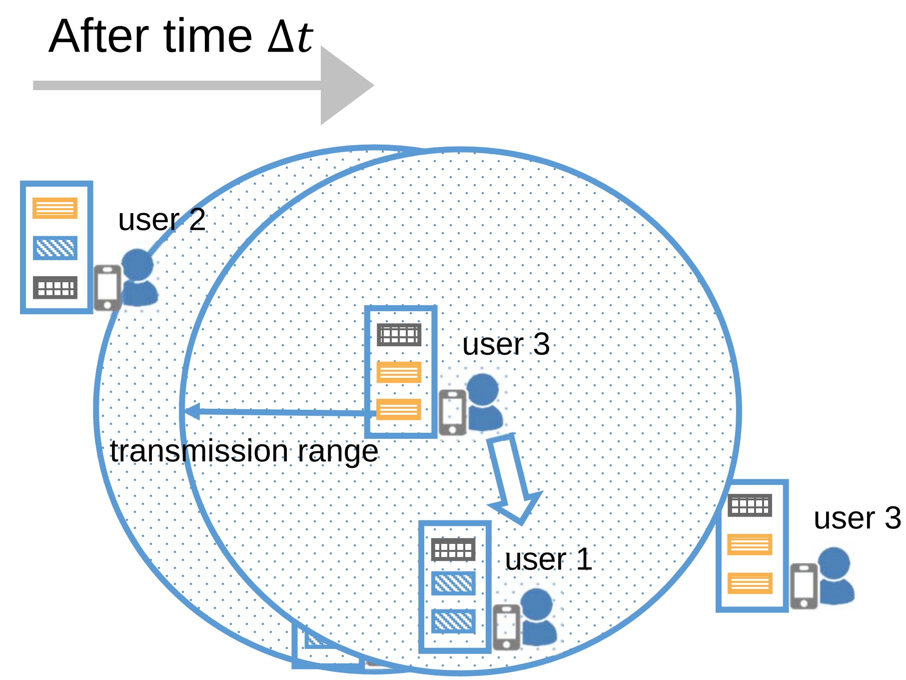
Case Study: Mobility-Aware Caching Content Placement at UTs

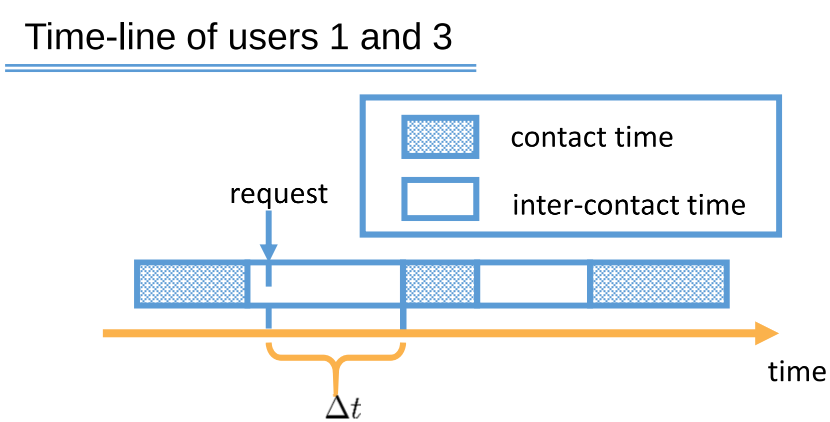
User inter-contact time ⇨ Caching content placement at UTs

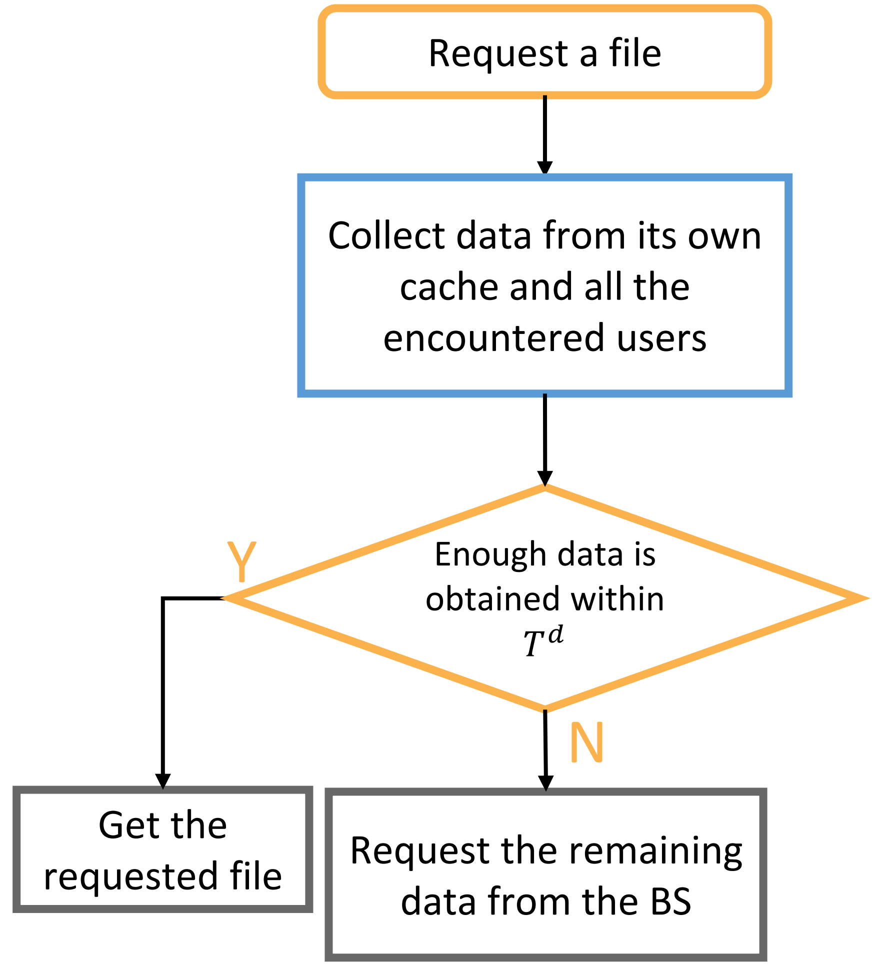


[Ref] R. Wang, J. Zhang, S.H. Song, and K. B. Letaief, "Mobility-aware caching in D2D networks," submitted to IEEE Trans. Wireless Commun., Jun. 2016. Available at http://arxiv.org/abs/1606.05282.

System Model





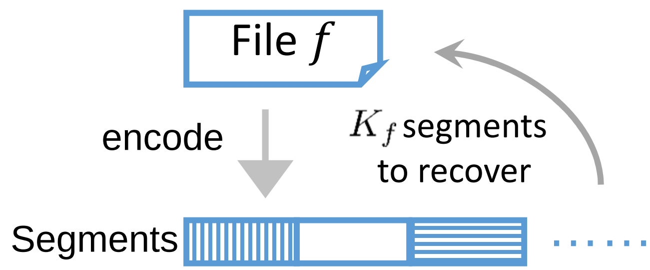


System Model

Caching strategy model

N_{file} files: $\mathcal{F} = \{1, 2, ..., N_{file}\}$

File f is requested with probability p_f

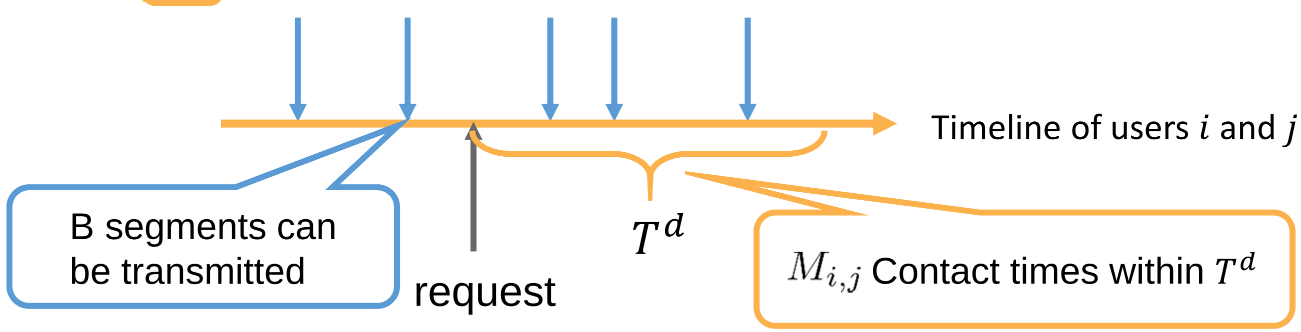


N_u mobile users: $\mathcal{D} = \{1, 2, ..., N_u\}$

- Each user can cache C segments
- User i caches $x_{i,f}$ segments of file f

User mobility model

Arrivals of contact times: Poisson Process with intensity $\lambda_{i,j}$



Problem Formulation

$$\max_{X} \quad \frac{1}{N_u} \sum_{i \in \mathcal{D}} \sum_{f \in \mathcal{F}} \frac{p_f}{K_f} \left\{ \mathbb{E} \left[\min \left(\sum_{j \in \mathcal{D}} \min(BM_{i,j}, x_{j,f}), K_f \right) \right] \right\}$$

Objective:

Data offloading ratio= $\dfrac{\text{data dilivered via D2D links}}{\text{requested data}}$

- Higher spatial efficiency
- Reduce backhaul burden

$$\text{s.t.} \quad \sum_{f \in \mathcal{F}} x_{j,f} \leqslant C, \forall j \in \mathcal{D}$$

Finite caching storage

$$x_{j,f} \in \mathbb{N}, \forall j \in \mathcal{D} \text{ and } f \in \mathcal{F}$$

Each segment is either fully stored or not stored

Main Challenges and Solutions

Evaluating the complicated objective

Pdf of $\sum_{j \in \mathcal{D}} \min(BM_{i,j}, x_{j,f})$

↓

Divide and conquer algorithm

Mixed integer non-linear programming (MINLP) problem

Algorithm	Performance	Complexity
Dynamic programming	Optimal	Exponential (much better than exhaustive search)
Greedy (Submodular maximization)	At least 50% of the optimal	Polynomial

Simulation Results

Caching strategies

Optimal Mobility-aware caching strategy:
The proposed DP optimal algorithm

Greedy Mobility-aware caching strategy:
The proposed polynomial time algorithm

Random caching strategy:
the probabilities of each user to cache segments of different files are proportional to the file request probabilities.

Popular caching strategy:
each user device stores the segments of the most popular files

File request probability

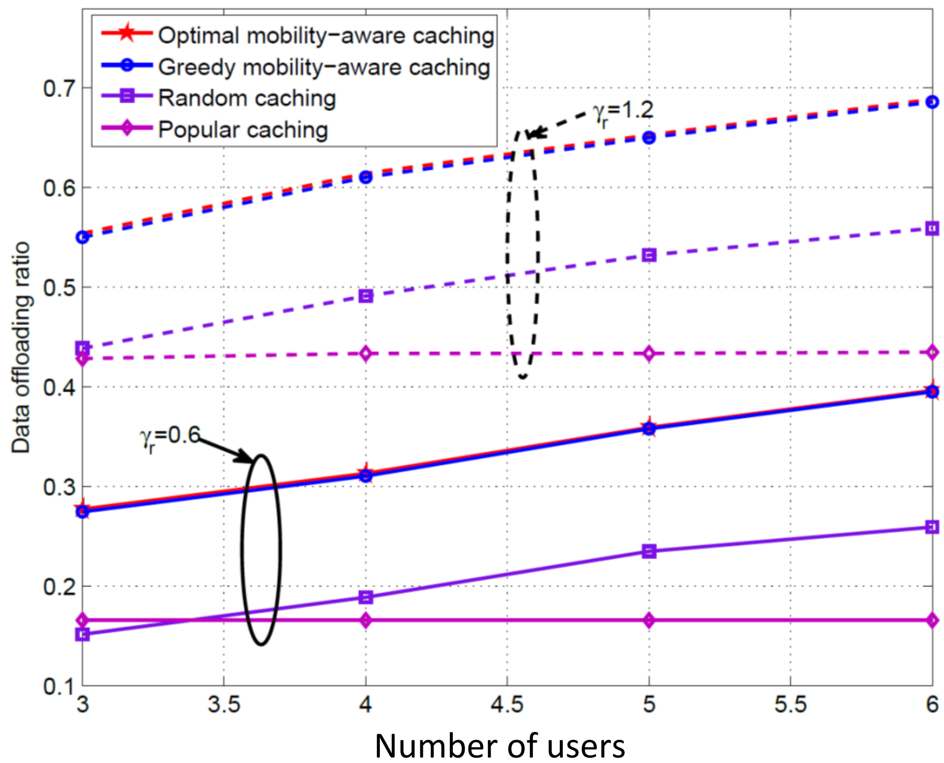
Zipf distribution with parameter γ_r

$$p_f = \frac{f^{-\gamma_r}}{\sum\limits_{i \in \mathcal{F}} i^{-\gamma_r}}$$

Simulation Results

Performance of the greedy caching algorithm is very close to the optimal one

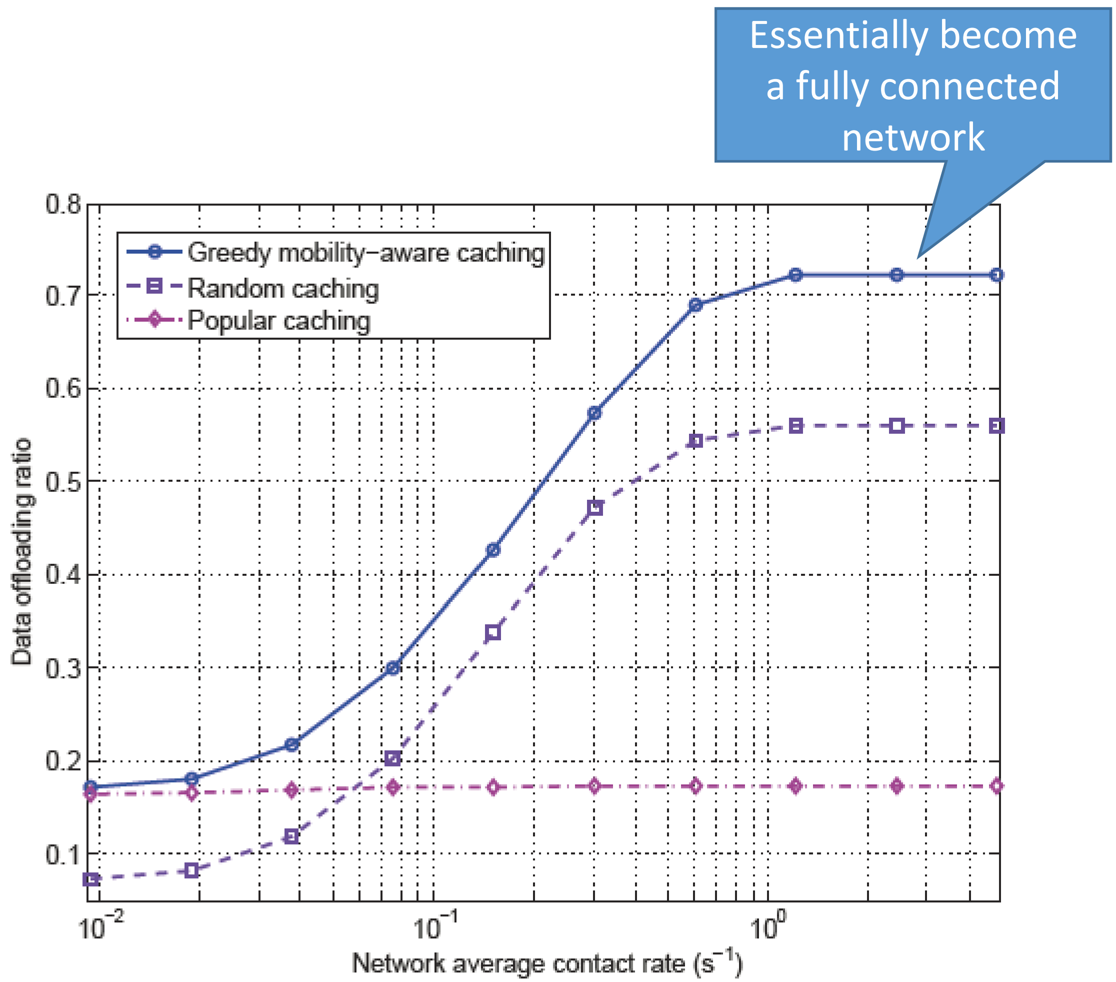
Mobility-aware caching outperforms both random and popular caching



Number of files = 50, T^d = 600 s,
Number of segments can be transmitted in on contact time = 1,
Number of segments to recover one file: randomly distributed in [1,3].

Simulation Results

It is critical to exploit mobility information

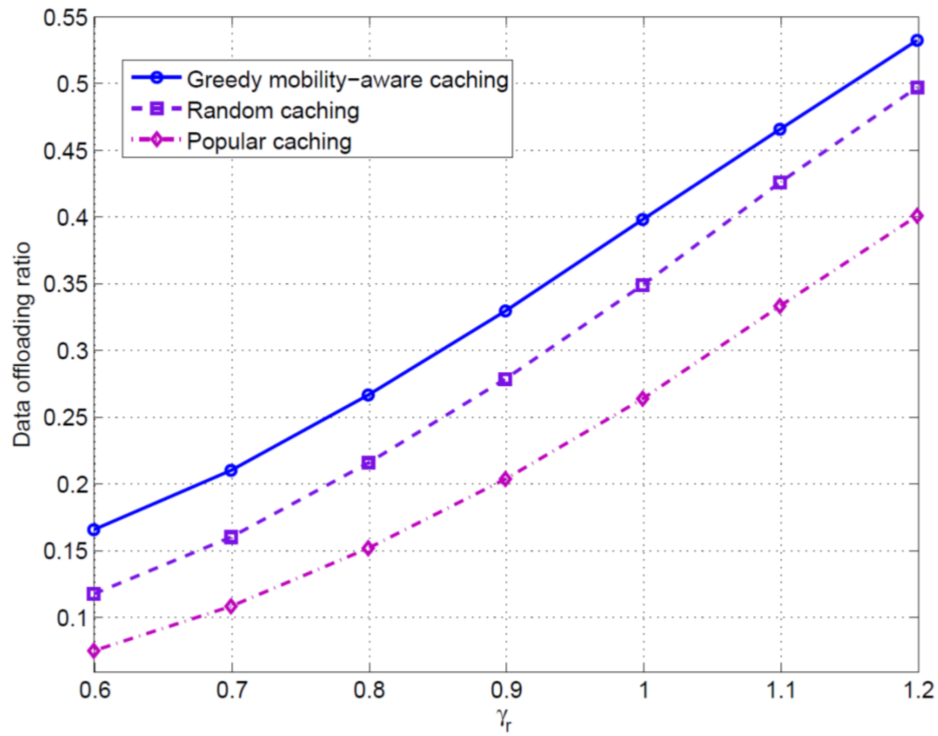


Simulation Results

Real-life data set collected in INFOCOM 2006 [Chaintreau ’07] of 78 students

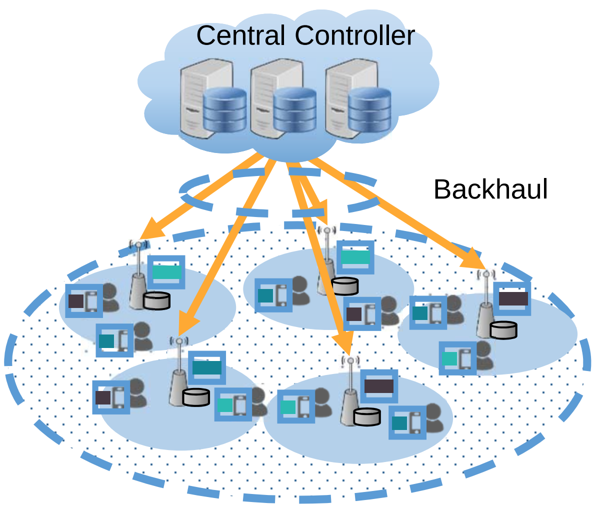
Use the daytime data during the first day to design the mobility-aware caching strategy

Show the performance during the daytime in the second day



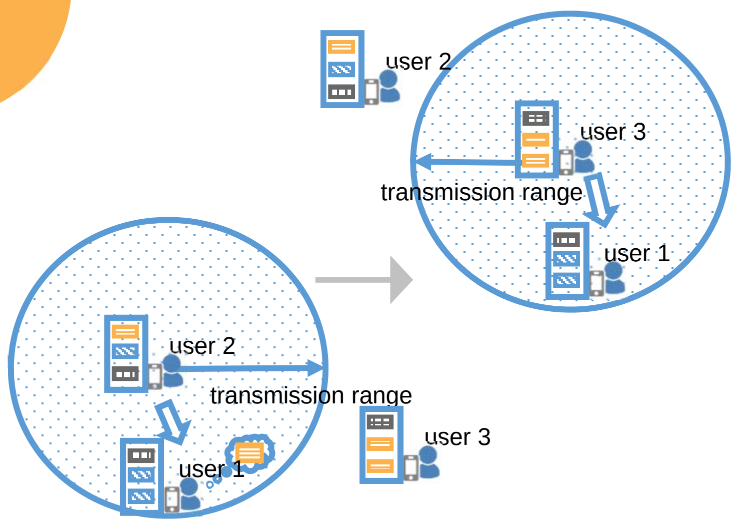
Number of files = 500, T^d = 600 s,
Number of segments can be transmitted in on contact time = 1,
Number of segments to recover one file: randomly distributed in [1,5],
Number of segments cached at each device = 10.

Conclusions



User Mobility Information
-- Valuable to account for
Mobility-Aware Caching
-- Effective for CCWNs

Lots of Interesting Directions
- Big data analytics for caching
- Privacy issues in obtaining mobility information
- Caching with dynamic user storage
- …



References

Main references

- R. Wang, X. Peng, **J. Zhang**, and K. B. Letaief, "Mobility-aware caching for content-centric wireless networks: Modeling and methodology," *IEEE Commun. Mag.*, to appear.
- R. Wang, **J. Zhang**, S.H. Song, and K. B. Letaief, "Mobility-aware caching in D2D networks," submitted to *IEEE Trans. Wireless Commun.*, Jun. 2016. Available at http://arxiv.org/abs/1606.05282.

Other references

- J. Liu, B. Bai, **J. Zhang**, and K. B. Letaief, "Content caching at the wireless network edge: A distributed algorithm via brief propagation," *IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016. (**Best Paper Award**)
- X. Peng, **J. Zhang**, S.H. Song, and K. B. Letaief, "Cache size allocation in backhaul limited wireless networks," *IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016.
- X. Peng, J.-C. Shen, **J. Zhang**, and K. B. Letaief, "Backhaul-aware caching placement for wireless networks," in *Proc. IEEE Globecom*, San Diego, CA, Dec. 2015.
- X. Peng, J.-C. Shen, **J. Zhang**, and K. B. Letaief, "Joint data assignment and beamforming for backhaul limited caching networks," in *Proc. IEEE Int. Symp. on Personal Indoor and Mobile Radio Comm. (PIMRC)*, Washington, DC, Sept. 2014. (**Best Paper Award**)

For more information: http://www.ece.ust.hk/~eejzhang/

Or **Email**: eejzhang@ust.hk

香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

Thank You!