On the Multi-Activation Oriented Design of D2D-Aided Caching Networks

Yang Cai, Andreas F. Molisch

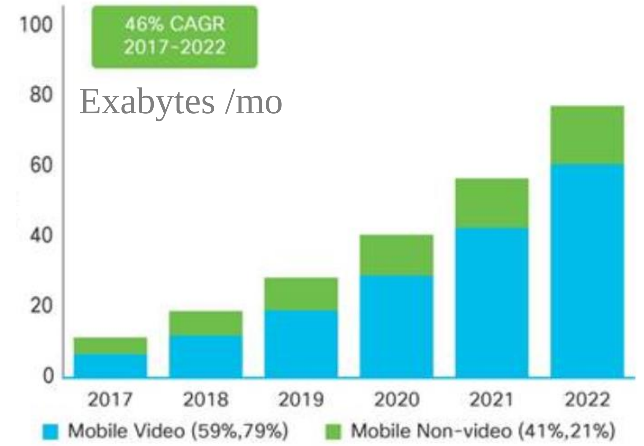
University of Southern California, CA, USA

Outline

- Background
- System Model
- Proposed Design
- Numerical Results
- Conclusions

Background

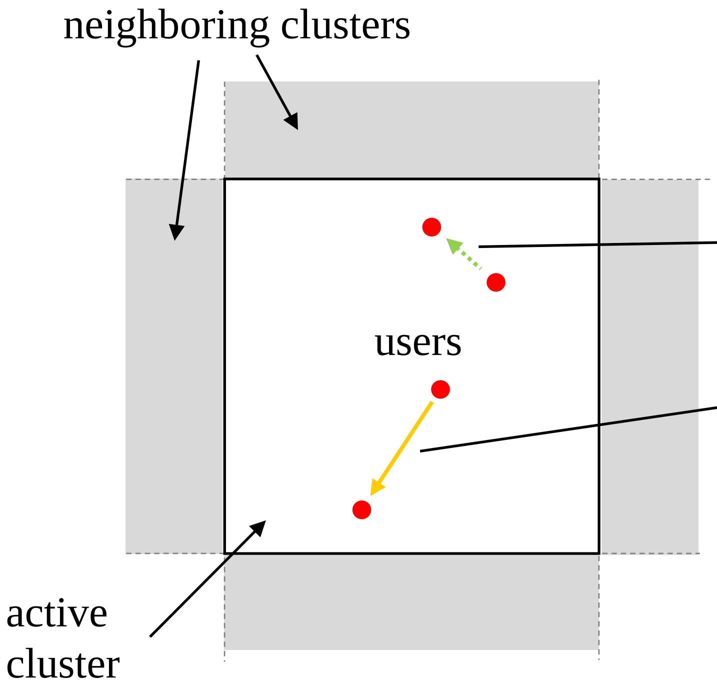
Data traffic of mobile video (source: Cisco virtual networking index)



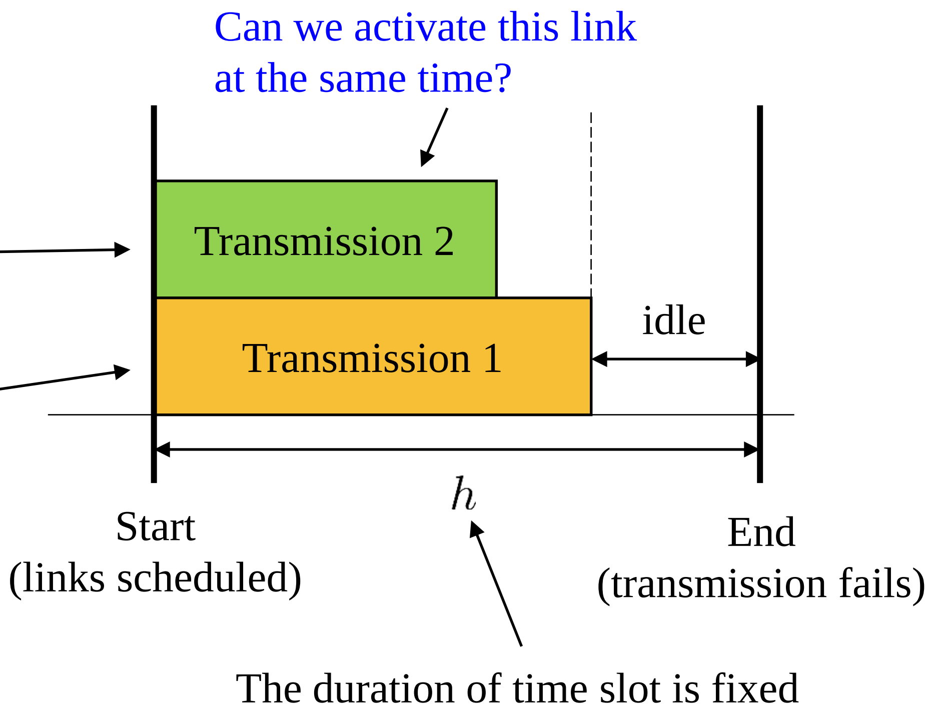
- Huge demand for video contents
- D2D-aided caching networks
  - Store the videos in the user devices
  - Transmit the videos by D2D comm.
- Related work
  - Transmission scheme (cluster model)
    - A cell is divided into several clusters
    - Users in the same cluster can cooperate
    - Only one link is activated at a time
  - Caching policy
    - Which files are cached by which user
    - Multiple metrics: throughput, hit rate, …
    - Optimum caching distribution [Ji2016]

Transmission Scheme

- Spatial
  - Cluster model



- Temporal
  - A typical time slot



Motivation

- Multiple links can be activated simultaneously within one cluster
  - Advantage: improve the spectral efficiency
- Problem 1: clash avoidance
  - Caching distribution: increase # of copies of the popular files
    - Fact: users requesting the same video might access to the same node
- Problem 2: interference management
  - Given # of activated links: link scheduling and power control
  - Determine the optimal duration of each transmission session
    - Tradeoff between more links and higher SINR

Outline

- Background
- System Model
- Proposed Design
- Numerical Results
- Conclusions

System Model

- Consider a cluster of N users with video requests
  - Library includes N_f possible videos, each of size F
- Request model
  - Poisson process with λ
  - Recent request replaces an older one
  - Preference distribution $\boldsymbol{p}_\mathrm{r} = \{p_{\mathrm{r},i} : 1 \leq i \leq N_\mathrm{f}\}$
- Caching model (probabilistic caching)
  - Caching distribution $\boldsymbol{p}_\mathrm{c} = \{p_{\mathrm{c},i} : 1 \leq i \leq N_\mathrm{f}\}$
  - Caching size for each user S

System Model

- Transmission scheme
  - Channel model
    - Pathloss $\eta(r) = 10\alpha \log_{10} r + 10 \log_{10} \eta_0$
  - Signal to interference and noise ratio (SINR) of link $e \in \mathcal{E}$
    - Decision variables: activation indicator $a_e \in \{0, 1\}$, and transmit power $p_e \in [0, P_{\max}]$

$$\gamma_e = \frac{a_e p_e r_e^{-\alpha}}{I_e + \eta_0 \sigma_e^2}$$

interference — noise power

$$\sum_{e \in \{\text{links involving node } i\}} a_e \leq 1$$

$$\gamma_e \geq a_e \gamma_0, \ \text{with } \gamma_0 \triangleq \mathrm{e}^{\frac{F \ln 2}{Bh}} - 1$$

  - Time is slotted
    - Reliability constraint: the transmission must be completed within one time slot (# of activated links is related to the duration of each time slot)

System Model

- Goal: maximize the D2D-supplied throughput

size of each video

$$T = \lim_{n \to \infty} \frac{F}{nh} \sum_{t=1}^{n} L(t)$$

$L(t)$: # of activated links in time slot t

h: duration of each time slot

  - P1: increase # of clash-free D2D requests (related to caching distribution)
  - P2: interference management (link scheduling and power control)
  - P3: tradeoff high SINR and more links (duration of each time slot)

Outline

- Background
- System Model
- Proposed Design
- Numerical Results
- Conclusions

Proposed Design P1

- Problem description
  - Input: the # of requests U (i.e., the number of active users)
    - Assumption: inactive users serve as helper, totally $V=(N-U)S$ cached videos
  - Output: caching distribution that maximizes # of clash-free D2D requests

- Expected # of clash-free D2D links Z
  - For a certain file i
    - # of requests $u_i \sim \mathcal{B}(U, \boldsymbol{p}_{\mathrm{r},i})$ (binominal dist.)
    - # of copies within the cluster $v_i \sim \mathcal{B}(V, \boldsymbol{p}_{\mathrm{c},i})$
    - # of clash-free requests $z_i = \min(u_i, v_i)$

$$\max_{\boldsymbol{p}_{\mathrm{c}} \geq 0} \; Z(\boldsymbol{p}_{\mathrm{c}}) = \sum_{i=1}^{N_{\mathrm{f}}} \mathbb{E}\{z_i\} \quad \text{(not tractable)}$$

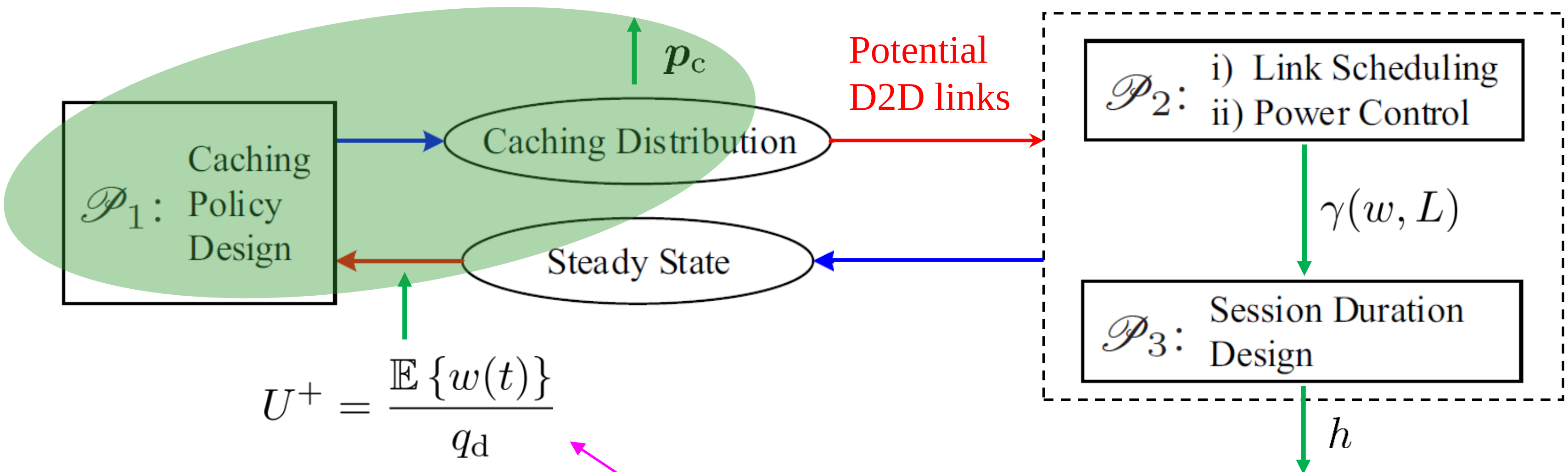
$$\text{s. t.} \; \sum_{i=1}^{N_{\mathrm{f}}} \boldsymbol{p}_{\mathrm{c},i} = 1.$$

  - Approximate form: $Z(\boldsymbol{p}_{\mathrm{c}}) \approx U - \sum_{i=1}^{N_{\mathrm{f}}} A_i \exp\left[-(1-Q_i)V\boldsymbol{p}_{\mathrm{c},i}\right]$ (concave, tractable)

    A_i, Q_i: parameters related to preference dist.

Proposed Design

- Iterative design
  - Green arrows represent the information updated in each iteration



$$\mathbb{E}\{\#\text{ of D2D requests}\} = U^{+} q_{\mathrm{d}} = \mathbb{E}\{w(t)\}$$

Can be calculated from the steady state distribution of $w(t)$

prob. for a request to be D2D supplied

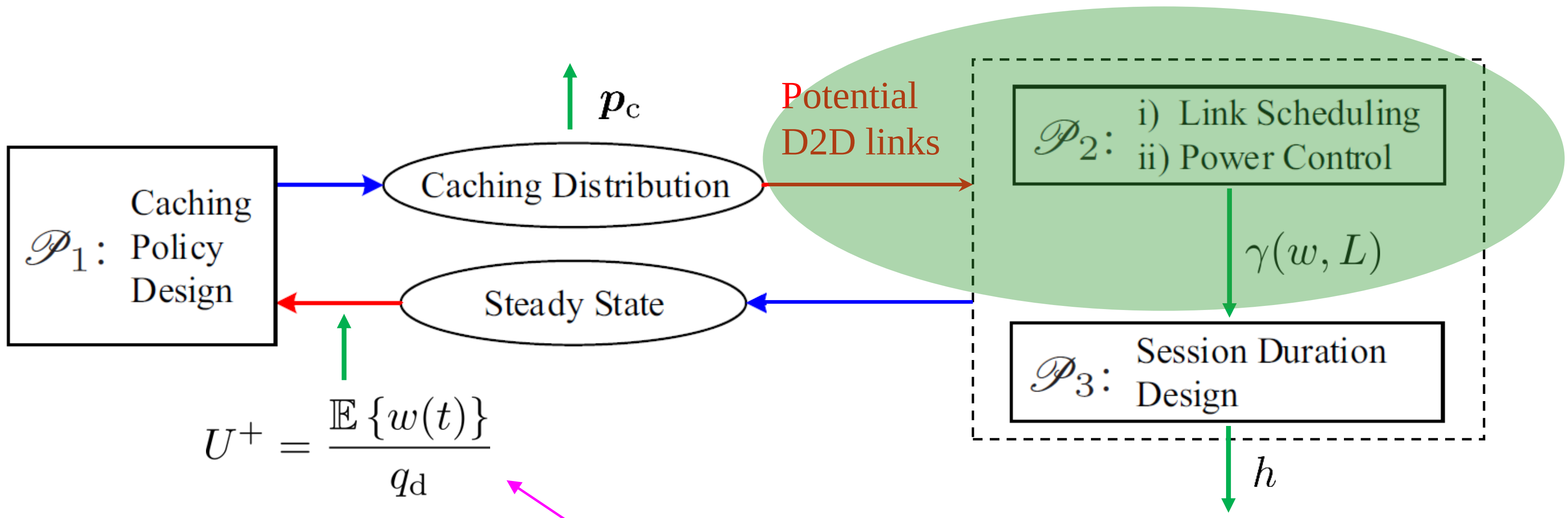
Proposed Design P2

- Transmission scheme (network operations)
  - Input: # of desired links L
  - Output: a set of L links and the corresponding transmit powers that maximize the minimum SINR (max-min-SINR)

- Link scheduling & Power control
  - In general, dealing with the two problems jointly is difficult
  - Our approach (separate)
    - Link scheduling: suppose the transmit power is $P_{\max}$ for all activated links, and find the links that max-min-SINR (greedy algorithm: select the links one by one)
    - Power control: given the scheduled links, coordinate their transmit powers to max-min-SINR (feasibility check problem for linear system, which is convex)

Proposed Design

- Iterative design
  - Green arrows represent the information updated in each iteration



$$U^{+} = \frac{\mathbb{E}\{w(t)\}}{q_{\mathrm{d}}}$$

$$\mathbb{E}\{\#\text{ of D2D requests}\} = U^{+} q_{\mathrm{d}} = \mathbb{E}\{w(t)\}$$

Can be calculated from the steady state distribution of $w(t)$

prob. for a request to be D2D supplied

Proposed Design P3

- Transmission scheme (duration of each time slot)
  - Tradeoff between more links (large h) and higher SINR (small h)
  - Require extensive simulations (dynamic network, random topology)

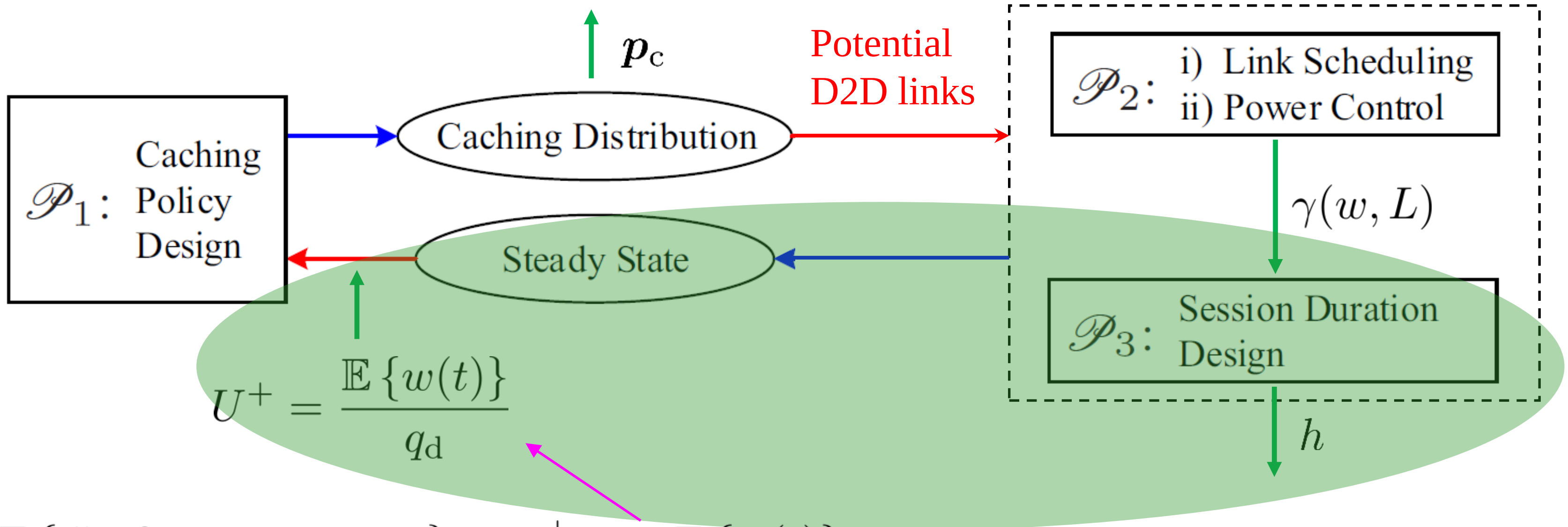
- Throughput $T = \frac{F\,\mathbb{E}\{L;h\}}{h} = \frac{F\,\mathbb{E}_w\{\mathbb{E}_L\{L|w;h\}\}}{h}$

  video size — b) w: # of all links — a) L: # of activated links

  - Approximated throughput: we can use the min-SINR $\gamma_k(w,L)$ calculated from K_0 random network topologies (one snapshot) to
    a) Approximate the conditional dist. of L (given w and h)
    b) Approximate the dist. of w over time (Markov chain)

Proposed Design

- Iterative design
  - Green arrows represent the information updated in each iteration



$$\mathbb{E}\{\#\text{ of D2D requests}\} = U^+ q_d = \mathbb{E}\{w(t)\}$$

prob. for a request to be D2D supplied

Can be calculated from the steady state distribution of $w(t)$

Outline

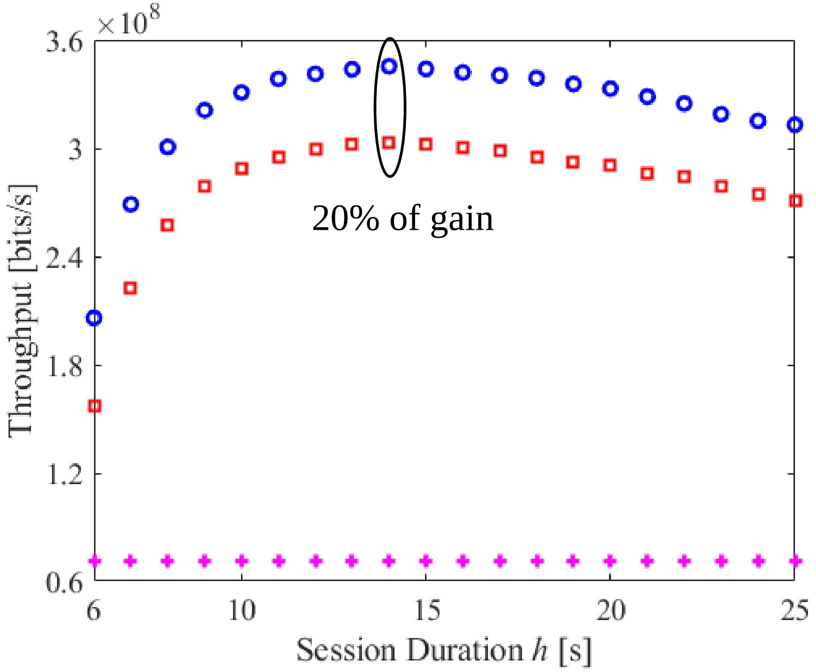
- Background
- System Model
- Proposed Design
- Numerical Results
- Conclusions

Numerical Experiments

- Configuration
  - Network setup
    - $N = 100$ users, within a cluster of size $100\ \mathrm{m} \times 100\ \mathrm{m}$
    - Speed: $1.5\,\mathrm{m/s}$, change to random direction every $10\,\mathrm{s}$, reflecting in the area
  - Video request
    - $N_{\mathrm{f}} = 500$ videos, each of size $F = 250\ \mathrm{MB}$
    - Request intensity $\lambda = 10^{-2}\,/\mathrm{s}$
    - Preference distribution (Zipf) $\boldsymbol{p}_{\mathrm{r},i} = i^{-0.6}/(\sum_{j=1}^{N_{\mathrm{f}}} j^{-0.6})$
  - Wireless transmission
    - D2D transmission: pathloss $\eta(r) = 36.8\log_{10}(r) + 37.6$, bandwidth $B = 20\ \mathrm{MHz}$, power budget $P_{\max} = 100\ \mathrm{mW}$, noise $\sigma_e^2 = -97\ \mathrm{dBm}$, antenna gain $-3.5\ \mathrm{dB}$
    - BS transmission: $N_{\mathrm{b}} = 10$ users, at a constant rate $R_{\mathrm{b}} = 200\mathrm{kb/s}$

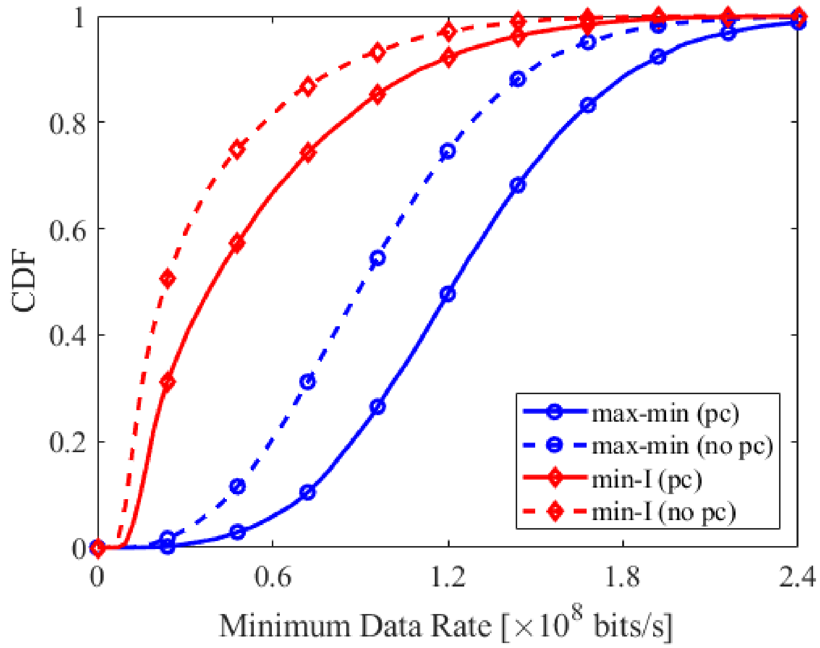
Numerical Experiments

- Gain of proposed caching distribution
  - Compared with max-hit and selfish caching



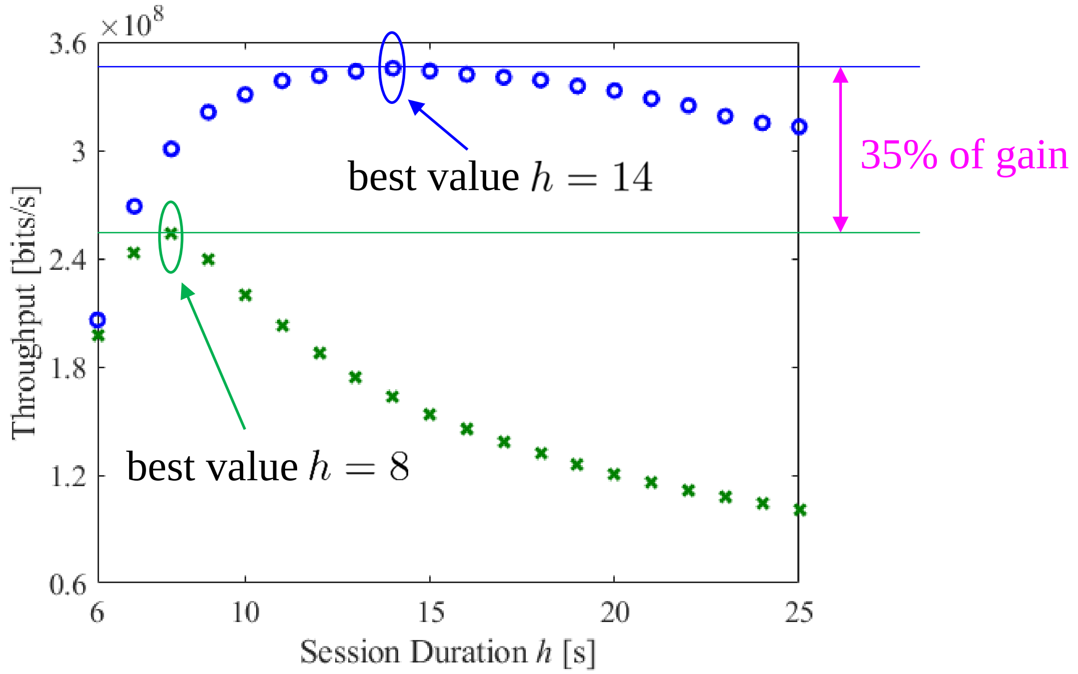
Numerical Experiments

- Gain of network operation (the CDF of the min data rate)
  - Link scheduling: max-min-SINR outperforms min-interference [Zhang2016]
  - Power control (solid line) leads to significant gain



Numerical Experiments

- Gain of multi-activation
  - Compared with TDMA
    - Less sensitive to h



Outline

- Background
- System Model
- Proposed Design
- Numerical Results
- Conclusions

Conclusions

- MA can improve the spectral efficiency, and thus the throughput

- Joint optimization of the caching policy and transmission scheme
  - Caching distribution: clash-free access to the contents
  - Interference management: the max-min-SINR criteria
  - System parameter design: tradeoff between more links and higher SINR

Thank you!