Coded Caching: Discussing Promises and Challenges

Eleftherios Lampiris
lampiris@eurecom.fr

EURECOM

December 2017

Outlook

Promises

- Coded Caching
- MISO Coded Caching
- Decentralised Approach

Outlook

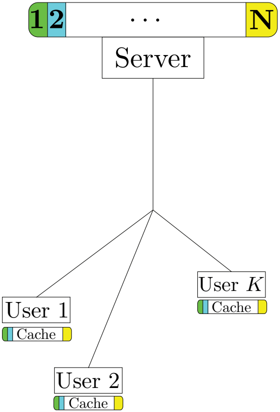
Promises

- Coded Caching
- MISO Coded Caching
- Decentralised Approach

Challenges

- Uneven Channel Strengths
- CSI Requirements
- Subpacketization issues

Coded Caching

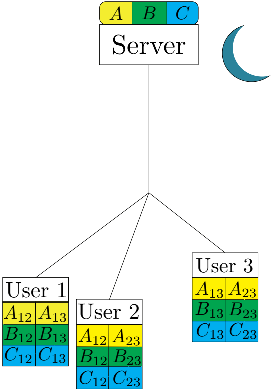


[Maddah-Ali, Niesen '12]

- N Files (Uniform Demands)
- F bits per File
- K Users
- $M \cdot F$ User Memory
- $\gamma = \frac{M}{N}$ Normalized Cache

Example 1

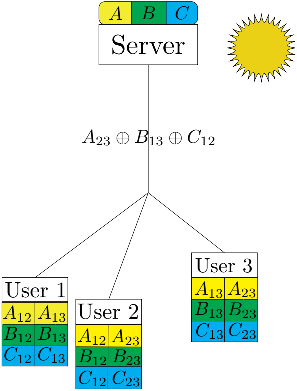
3 Users - 3 Files - $\gamma = \frac{2}{3}$



- $M = 2$
- Files are divided into 3 parts $\{12, 13, 23\}$
- Cache i is filled-up with all parts indexed with i

Example 1

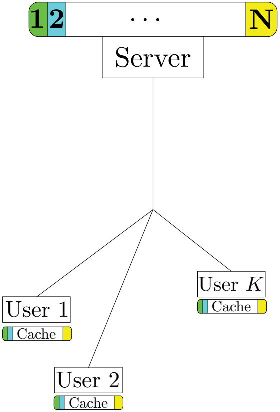
3 Users - 3 Files - $\gamma = \frac{2}{3}$



$A_{23} \oplus B_{13} \oplus C_{12}$

Single Message serves all users at the same time

Coded Caching



[Maddah-Ali, Niesen '12]

- N Files (Uniform Demands)
- F bits per File
- K Users
- $M \cdot F$ User Memory
- $\gamma = \frac{M}{N}$ Normalized Cache

Users Served Simultaneously

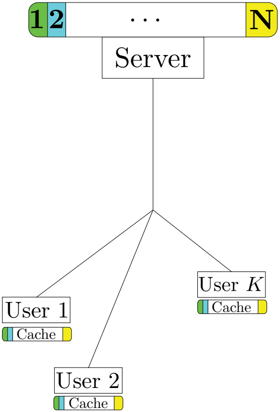
$$d_{\Sigma} = K\gamma + 1$$

Optimality

Optimal within a multiplicative factor of 2.008.

[Qian, Maddah-Ali, Avestimehr '17]

Effect of Coded Caching

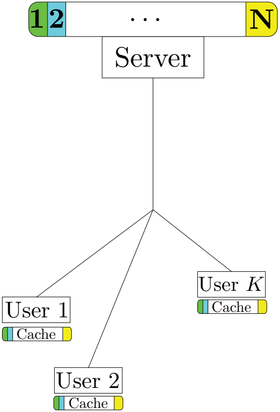


[Maddah-Ali, Niesen '12]

Time under Uniform Demand

$$T = K(1-\gamma)\frac{1}{K\gamma+1} \approx \frac{1-\gamma}{\gamma}$$

Effect of Coded Caching



[Maddah-Ali, Niesen '12]

Time under Uniform Demand

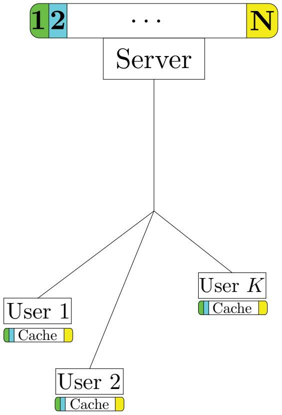
$$T = K(1-\gamma)\frac{1}{K\gamma+1} \approx \frac{1-\gamma}{\gamma}$$

Per User Rate

$$R_u = \left(\gamma + \frac{1}{K}\right) \cdot R_{\text{tot}} \approx \gamma \cdot R_{\text{tot}}$$

Each gets a γ piece of the pie!!

Effect of Coded Caching



[Maddah-Ali, Niesen '12]

Time under Uniform Demand

$$T = K(1-\gamma)\frac{1}{K\gamma+1} \approx \frac{1-\gamma}{\gamma}$$

Per User Rate

$$R_u = \left(\gamma + \frac{1}{K}\right) \cdot R_{\text{tot}} \approx \gamma \cdot R_{\text{tot}}$$

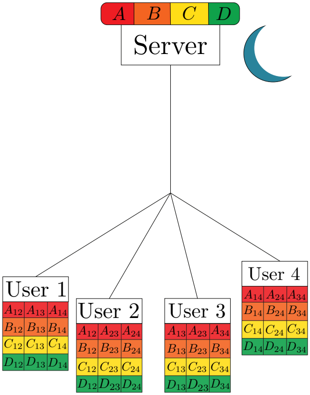
Each gets a γ piece of the pie!!

Intuition

Even unwanted packets can help reduce interference

Example 2

4 Users - 4 Files - $\gamma = \frac{2}{4}$

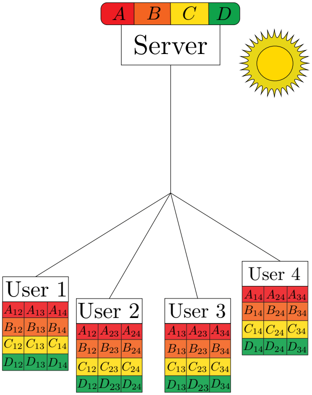


Placement Phase

- Files are divided into 6 parts $\{12, 13, 14, 23, 24, 34\}$
- Cache i is filled-up with all parts indexed with i

Example 2

4 Users - 4 Files - $\gamma = \frac{2}{4}$

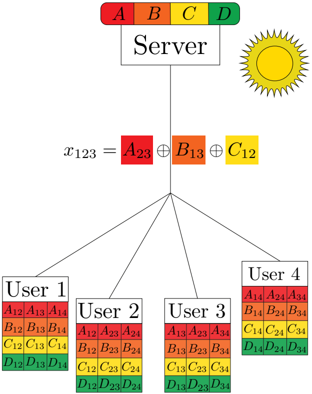


Delivery Phase

Every Message serves 3 users at the same time

Example 2

4 Users - 4 Files - $\gamma = \frac{2}{4}$



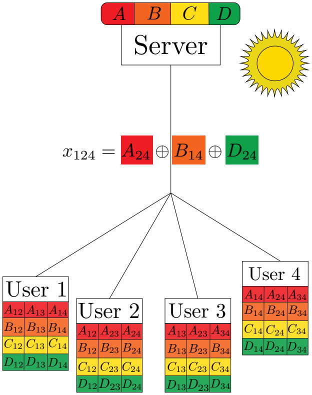
Delivery Phase

Every Message serves 3 users at the same time

- $x_{123} = A_{23} \oplus B_{13} \oplus C_{12}$

Example 2

4 Users - 4 Files - $\gamma = \frac{2}{4}$



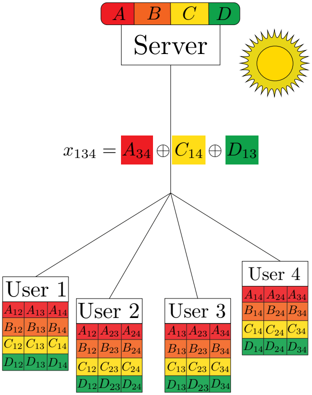
Delivery Phase

Every Message serves 3 users at the same time

- $x_{123} = A_{23} \oplus B_{13} \oplus C_{12}$
- $x_{124} = A_{24} \oplus B_{14} \oplus D_{24}$

Example 2

4 Users - 4 Files - $\gamma = \frac{2}{4}$



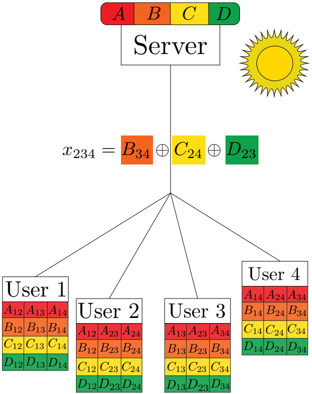
Delivery Phase

Every Message serves 3 users at the same time

- $x_{123} = A_{23} \oplus B_{13} \oplus C_{12}$
- $x_{124} = A_{24} \oplus B_{14} \oplus D_{24}$
- $x_{134} = A_{34} \oplus C_{14} \oplus D_{13}$

Example 2

4 Users - 4 Files - $\gamma = \frac{2}{4}$

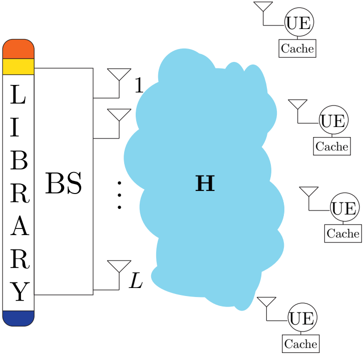


Delivery Phase

Every Message serves 3 users at the same time

- $x_{123} = A_{23} \oplus B_{13} \oplus C_{12}$
- $x_{124} = A_{24} \oplus B_{14} \oplus D_{24}$
- $x_{134} = A_{34} \oplus C_{14} \oplus D_{13}$
- $x_{234} = B_{34} \oplus C_{24} \oplus D_{23}$

MISO Coded Caching



[Shariatpanahi, Motahari, Khalaj '15 &
Naderializadeh, Maddah-Ali, Avestimehr '16]

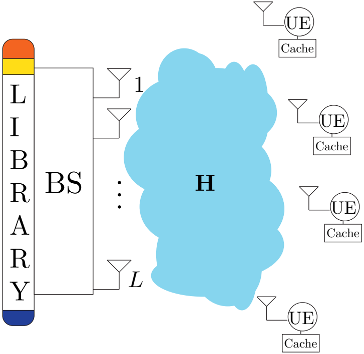
- N files
- K users
- L antennas
- γ fractional cache

Users Served

$$d_{\Sigma} = K\gamma + L \quad ^{\star}$$

$^{\star} d_{\Sigma} = \min\{K, K\gamma + L\}$
Linear One-Shot Order Optimal Gap $= 2$

MISO Coded Caching



[Shariatpanahi, Motahari, Khalaj '15 &
Naderializadeh, Maddah-Ali, Avestimehr '16]

- N files
- K users
- L antennas
- γ fractional cache

Users Served

$$d_\Sigma = K\gamma + L \quad ^\star$$

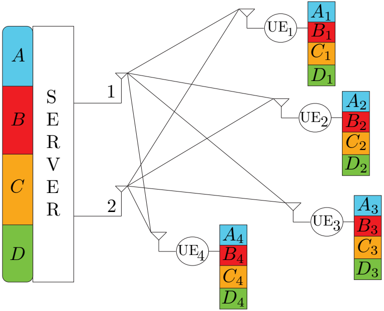
Intuition

Vector of Linear Combinations:
Cache-out some packets
Null-out some packets.

$^\star d_\Sigma = \min\{K, K\gamma + L\}$
Linear One-Shot Order Optimal Gap $= 2$

MISO CC Example

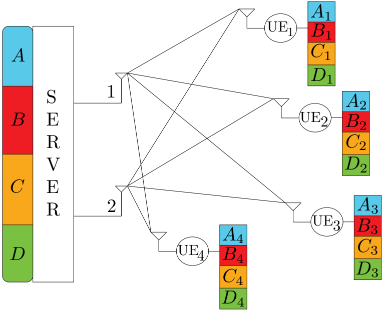
$K = 4,\ N = 4,\ L = 2,\ \gamma = \frac{1}{4}$



h_{ij} : Antenna i to User j Channel

MISO CC Example

$K = 4,\ N = 4,\ L = 2,\ \gamma = \frac{1}{4}$

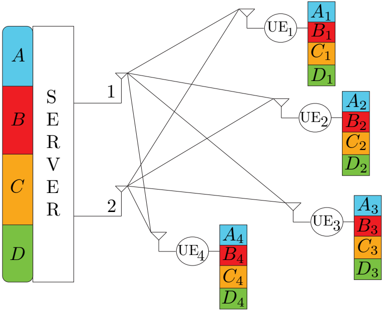


$$\underline{x}_{123} = \begin{bmatrix} h_{13}^{-1}A_2 + h_{11}^{-1}B_3 + h_{12}^{-1}C_1 \\ -h_{23}^{-1}A_2 - h_{21}^{-1}B_3 - h_{22}^{-1}C_1 \end{bmatrix}$$

h_{ij} : Antenna i to User j Channel

MISO CC Example

$K = 4,\ N = 4,\ L = 2,\ \gamma = \frac{1}{4}$



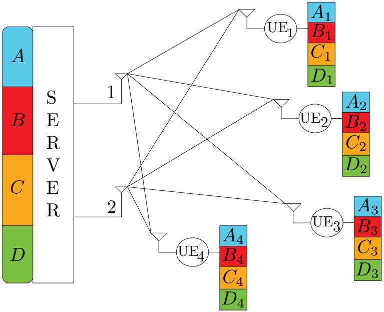
$$\underline{x}_{123} = \begin{bmatrix} h_{13}^{-1}A_2 + h_{11}^{-1}B_3 + h_{12}^{-1}C_1 \\ -h_{23}^{-1}A_2 - h_{21}^{-1}B_3 - h_{22}^{-1}C_1 \end{bmatrix}$$

$$\begin{aligned} y_1 &= A_2 + C_1(h_{11}h_{12}^{-1} - h_{21}h_{22}^{-1}) \\ y_2 &= B_3 + A_2(h_{12}h_{13}^{-1} - h_{22}h_{23}^{-1}) \\ y_3 &= C_1 + B_3(h_{13}h_{11}^{-1} - h_{23}h_{21}^{-1}) \end{aligned}$$

h_{ij} : Antenna i to User j Channel

MISO CC Example

$K = 4,\ N = 4,\ L = 2,\ \gamma = \frac{1}{4}$



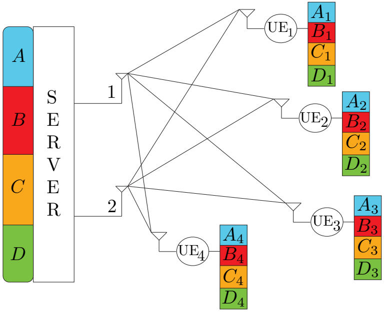
$$\underline{x}_{123} = \begin{bmatrix} h_{13}^{-1}A_2 + h_{11}^{-1}B_3 + h_{12}^{-1}C_1 \\ -h_{23}^{-1}A_2 - h_{21}^{-1}B_3 - h_{22}^{-1}C_1 \end{bmatrix}$$

$$\underline{x}_{124} = \begin{bmatrix} h_{12}^{-1}A_4 + h_{14}^{-1}B_1 + h_{11}^{-1}D_2 \\ -h_{22}^{-1}A_4 - h_{24}^{-1}B_1 - h_{21}^{-1}D_2 \end{bmatrix}$$

h_{ij} : Antenna i to User j Channel

MISO CC Example

$K = 4,\ N = 4,\ L = 2,\ \gamma = \frac{1}{4}$



h_{ij} : Antenna i to User j Channel

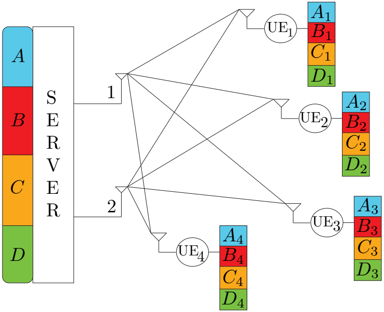
$$\underline{x}_{123} = \begin{bmatrix} h_{13}^{-1}A_2 + h_{11}^{-1}B_3 + h_{12}^{-1}C_1 \\ -h_{23}^{-1}A_2 - h_{21}^{-1}B_3 - h_{22}^{-1}C_1 \end{bmatrix}$$

$$\underline{x}_{124} = \begin{bmatrix} h_{12}^{-1}A_4 + h_{14}^{-1}B_1 + h_{11}^{-1}D_2 \\ -h_{22}^{-1}A_4 - h_{24}^{-1}B_1 - h_{21}^{-1}D_2 \end{bmatrix}$$

$$\underline{x}_{134} = \begin{bmatrix} h_{14}^{-1}A_3 + h_{11}^{-1}C_4 + h_{13}^{-1}D_1 \\ -h_{24}^{-1}A_3 - h_{21}^{-1}C_4 - h_{23}^{-1}D_1 \end{bmatrix}$$

MISO CC Example

$K = 4,\ N = 4,\ L = 2,\ \gamma = \frac{1}{4}$



h_{ij} : Antenna i to User j Channel

$$\underline{x}_{123} = \begin{bmatrix} h_{13}^{-1} A_2 + h_{11}^{-1} B_3 + h_{12}^{-1} C_1 \\ -h_{23}^{-1} A_2 - h_{21}^{-1} B_3 - h_{22}^{-1} C_1 \end{bmatrix}$$

$$\underline{x}_{124} = \begin{bmatrix} h_{12}^{-1} A_4 + h_{14}^{-1} B_1 + h_{11}^{-1} D_2 \\ -h_{22}^{-1} A_4 - h_{24}^{-1} B_1 - h_{21}^{-1} D_2 \end{bmatrix}$$

$$\underline{x}_{134} = \begin{bmatrix} h_{14}^{-1} A_3 + h_{11}^{-1} C_4 + h_{13}^{-1} D_1 \\ -h_{24}^{-1} A_3 - h_{21}^{-1} C_4 - h_{23}^{-1} D_1 \end{bmatrix}$$

$$\underline{x}_{234} = \begin{bmatrix} h_{13}^{-1} B_4 + h_{14}^{-1} C_2 + h_{12}^{-1} D_3 \\ -h_{23}^{-1} B_4 - h_{24}^{-1} C_2 - h_{22}^{-1} D_3 \end{bmatrix}$$

Decentralized Coded Caching

- Number of users showing up is unknown
- Cache uniformly at random

Result

$$T = \frac{(1-\gamma)}{\gamma}(1-(1-\gamma)^K) \quad \star$$

$\star$ Order Optimal

[Maddah-Ali, Niesen '13]

Decentralized Coded Caching

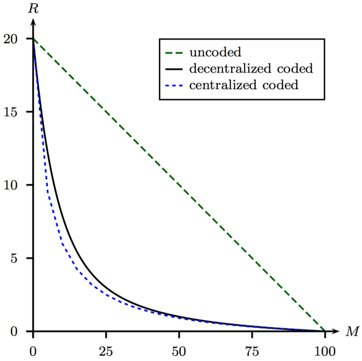
- Number of users showing up is unknown
- Cache uniformly at random

Result

$$T = \frac{(1-\gamma)}{\gamma}(1-(1-\gamma)^K) \quad \star$$

$\star$ Order Optimal

[Maddah-Ali, Niesen '13]



[source: Maddah-Ali, Niesen '13]

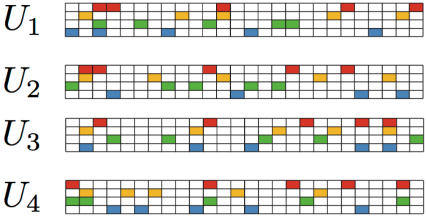
Decentralized Coded Caching

Example

Algorithm

- Start from higher order cliques and move to lower order
- Greedy approach

Placement



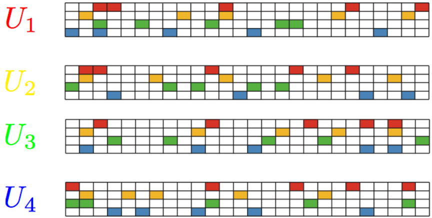
Decentralized Coded Caching

Example

Algorithm

- Start from higher order cliques and move to lower order
- Greedy approach

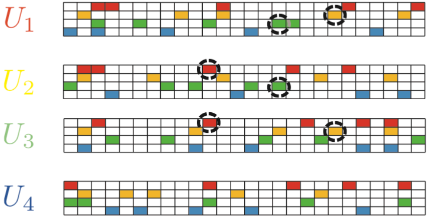
Requests



Decentralized Coded Caching

Example

Clique of 3

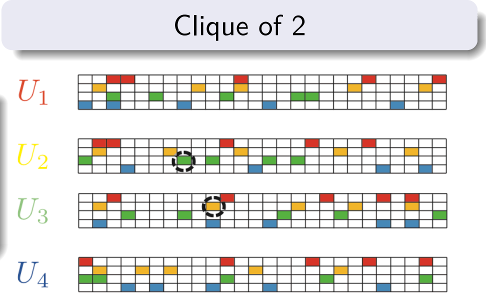


Algorithm

- Start from higher order cliques and move to lower order
- Greedy approach

Decentralized Coded Caching

Example



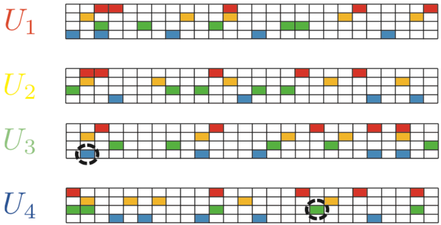
Algorithm

- Start from higher order cliques and move to lower order
- Greedy approach

Decentralized Coded Caching

Example

Clique of 2



Algorithm

- Start from higher order cliques and move to lower order
- Greedy approach

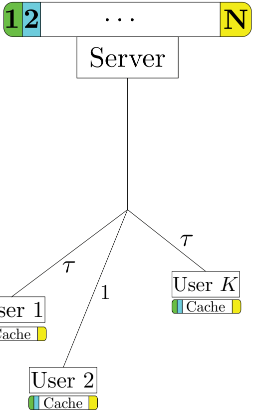
Challenges

Compromised gains due to

- Uneven Channel Strengths
- High CSIT requirements
- Extreme Subpacketization

Implementation Challenges

1) Uneven Channel Strengths

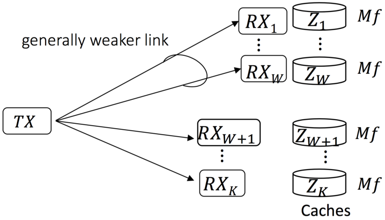


$\tau \leq 1$

- Worst-user effect
- Common reality in Wireless

Implementation Challenges

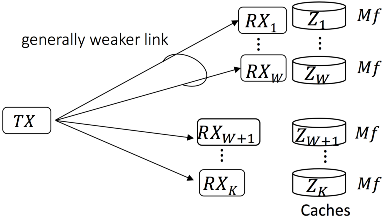
1) Uneven Channel Strengths: SISO with Topology



[Zhang, Elia '16
source: Zhang, Elia '16]

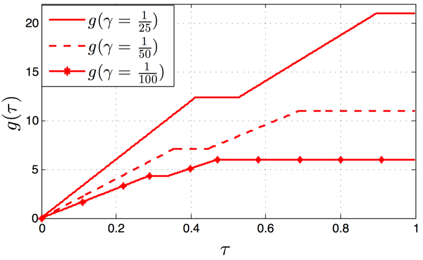
- W users have link capacity $\tau \leq 1$
- $K - W$ users have link capacity 1

- W users have link capacity $\tau \leq 1$
- $K - W$ users have link capacity 1
- Gains persist for some values



[Zhang, Elia '16
source: Zhang, Elia '16]

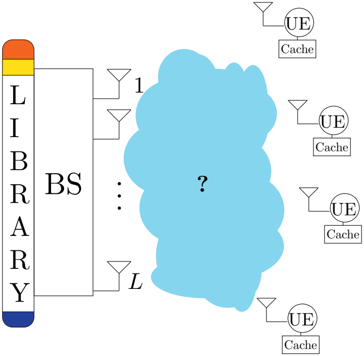
$$K = 500, W = 50$$

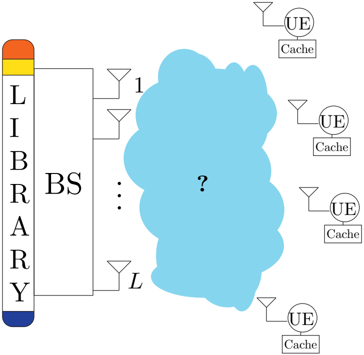


[source: Zhang, Elia '16]

Implementation Challenges

2) Too much CSI needed for MISO Coded Caching



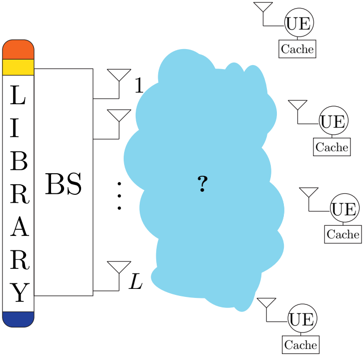


RECALL

$$y_1 = A_2 + C_1(h_{11}h_{12}^{-1} - h_{21}h_{22}^{-1})$$
$$y_2 = B_3 + A_2(h_{12}h_{13}^{-1} - h_{22}h_{23}^{-1})$$
$$y_3 = C_1 + B_3(h_{13}h_{11}^{-1} - h_{23}h_{21}^{-1})$$

Requires CSIR

Implementation Challenges
2) Too much CSI needed for MISO Coded Caching

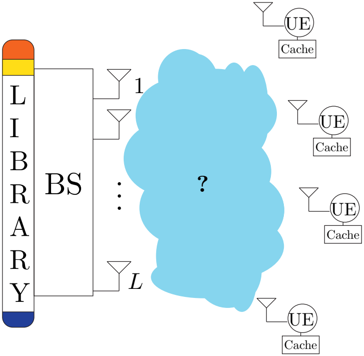


CSIT/CSIR required :
$(K\gamma + L)$ channel vectors

MISO CC requires a lot of training for high gains

Implementation Challenges

2) Too much CSI needed for MISO Coded Caching



CSIT/CSIR required :
$(K\gamma + L)$ channel vectors

MISO CC requires a lot of training for high gains

Take Home Message

Going from DoF
$K\gamma + 1 \rightarrow K\gamma + L$
Requires Massive CSI

Implementation Challenges

3) Subpacketization: Finite Length Analysis

First result discussing the astronomical file size requirement

Results

- Gain≈ 2 for Decentralized MN if $|F| \leq (\frac{e}{\gamma})^{K\gamma}$
- For a gain of g under any decentralized scheme $|F| = \mathcal{O}(\left(\frac{1}{\gamma}\right)^{g})$

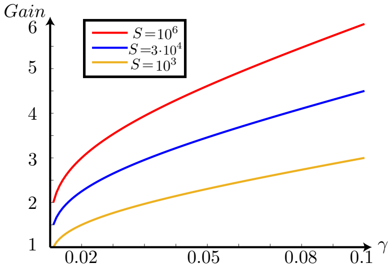
[Shanmugam, Ji, Tulino, Llorca, Dimakis '15]

First result discussing the astronomical file size requirement

Results

- Gain≈ 2 for Decentralized MN if $|F| \leq (\frac{e}{\gamma})^{K\gamma}$
- For a gain of g under any decentralized scheme $|F| = \mathcal{O}(\left(\frac{1}{\gamma}\right)^{g})$

[Shanmugam, Ji, Tulino, Llorca, Dimakis '15]

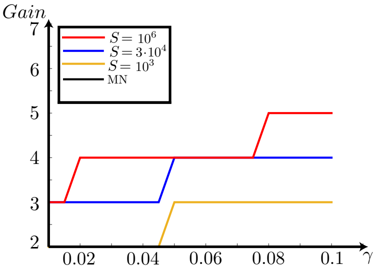


Take home message

Decentralized CC is limited by a gain of 6

Original Subpacketization

$$S_{MN} = \binom{K}{K\gamma}$$



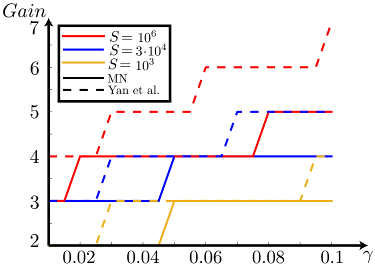
Original Subpacketization

$$S_{MN} = \binom{K}{K\gamma}$$

Reduced Subpacketization

$$S_2 = \left(\frac{1}{\gamma}\right)^{K\gamma-1}$$

[Yan, Cheng, Tang, Chen '15 & Tang, Ramamoorthy '16]



Implementation Challenges

3) Subpacketization: Other Schemes

Linear Subpacketization is Possible

- There exists a caching scheme with linear subpacketization and polynomial delivery time, $|F| = \mathcal{O}(K)$.
- In reality, $K \to \infty \Rightarrow F \to \infty$

[Shanmugam, Tulino, Dimakis '17]

Centralized CC: A hypergraph approach

- There exist no constant rate caching schemes with $|F| = \mathcal{O}(K)$
- Tradeoff between performance and subpacketization
- Schemes require $K > \frac{4}{\gamma^2}$ for $g \geq 2$

[Shangguan, Yiwei Zhang, Gennian Ge '16]

General Conclusions

Promises

- Coded Caching *in theory* can severely reduce traffic
- MISO CC can pave the way to supplement MISO gains with caching.

General Conclusions

Promises

- Coded Caching *in theory* can severely reduce traffic
- MISO CC can pave the way to supplement MISO gains with caching.

Hot Questions

- Achieve individual "speeds" without compromizing gains?
- Can feedback in MISO CC be separated from Caching Gains?
- High gains in Centralized/Decentralized?

Thanks!