Content Distribution for Mobile Infostation Networks

Wing Ho Andy Yuen
Roy D. Yates
Siun-Chuon Mau

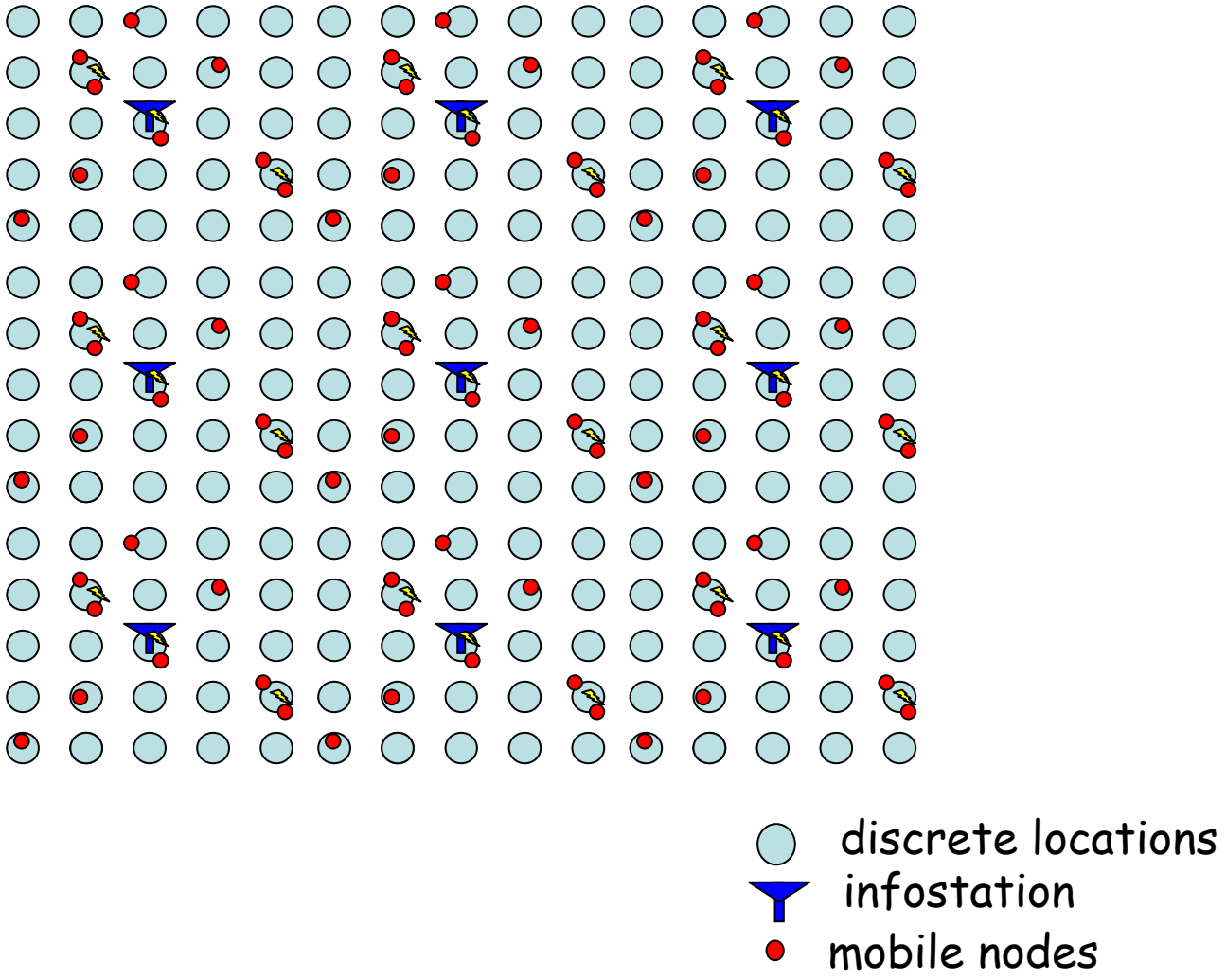
Mobile Infostations

- Multihop networks have low capacity
- New architectures (Mobile Infostations)
  - (Atay&Rose, Rutgers), (Papadopouli&Schulzrinne, Columbia), (Grossglauser&Tse, UCB)
  - Exploit node mobility; transmissions occur between nearest neighbors
- Common assumption: nodes cooperate to relay packets to the destination
  - Unrealistic to assume nodes will always forward packets for others
- Social contract for content distribution: Bilateral file exchanges occur only when both nodes retrieve a file that they want

Content Distribution

- Applications
  - Movie distribution network
  - CNN news network with periodic news updates
- Network Architecture
  - A node near an infostation downloads files
  - Two nodes near each other can exchange files
  - Data dissemination by all nodes at all locations
- Issues
  - Can everyone get a fair data rate?
  - Can everyone download the movie in a timely manner?
  - How does performance depend on node density, node mobility, and content size?

A Mobile Infostation Network

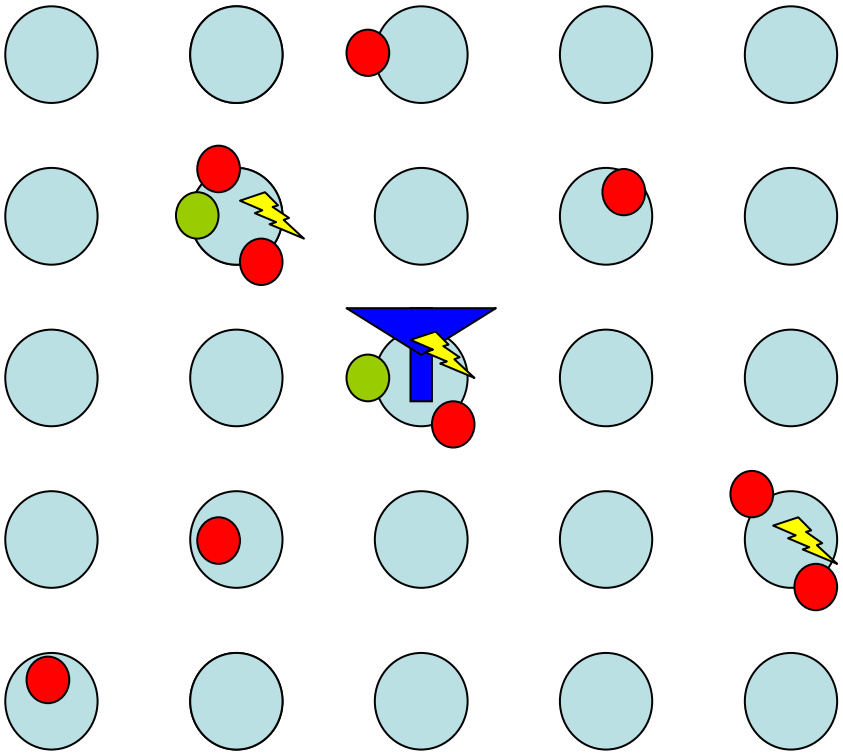


Popular Movie Downloading

- N: number of nodes in an area
- K: number of files cached at the infostation
- L: number of locations per infostation
- Bandwidth and Transmission Concurrency constraints:
  - one random node is picked to download 2 files from the infostation
  - two random nodes are picked to exchange 1 file in each direction at each location
- What does a user download?
  - Random algorithm: a node randomly selects a file it does not have

System Model

L=25
Locations



discrete mobile locations

infostation

node

Performance metrics

- T_1: time when 80% of nodes get all files
  - a measure of **Networking effect**
- T_2: time when all nodes get 80% of the files
  - a measure of **Fairness**
- $T_1 << T_2$ implies the network performance is unfair

- T_3: time when all nodes get all files
- T_4: time when a random node get all files
- Throughput Capacity $C_i = K/E[T_i]$ i=1,2,3,4
  - where K is the number of files
  - Unit: number of files per node per unit time

Factors Affecting Transmission

- Fraction of transmit mobiles $\beta(N)$
  - L locations $\Rightarrow$ L possible simultaneous transmissions
    - At a location, only 2 mobiles can transmit
    - At the infostation, only 1 mobile can download
  - Optimum node density is 1.7 nodes/location,
    - Contention success probability is 0.6
- Probability of File Exchange P_X
  - When K is large, probability of file exchange is 1 except one of the nodes has (1) nothing or (2) everything

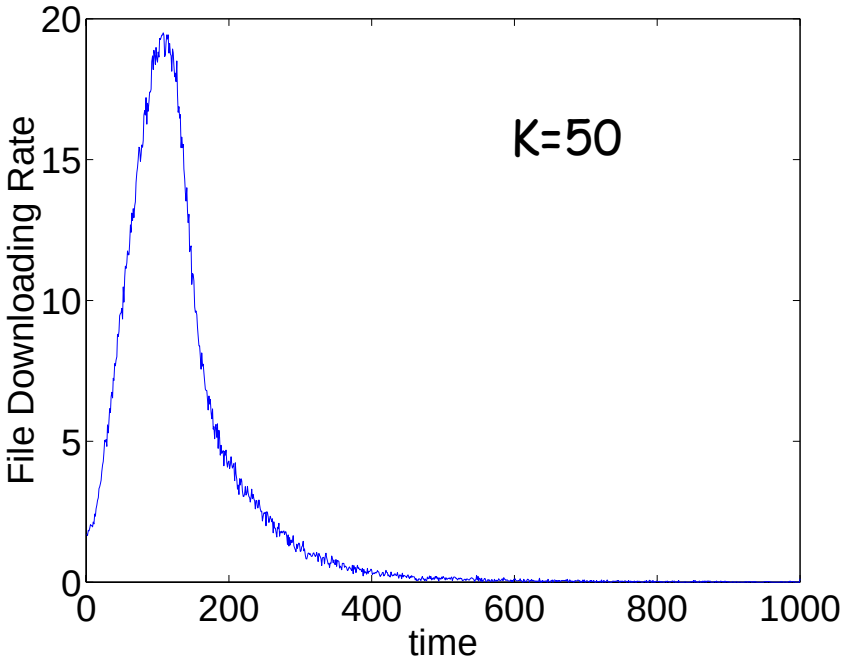
Throughput Capacity

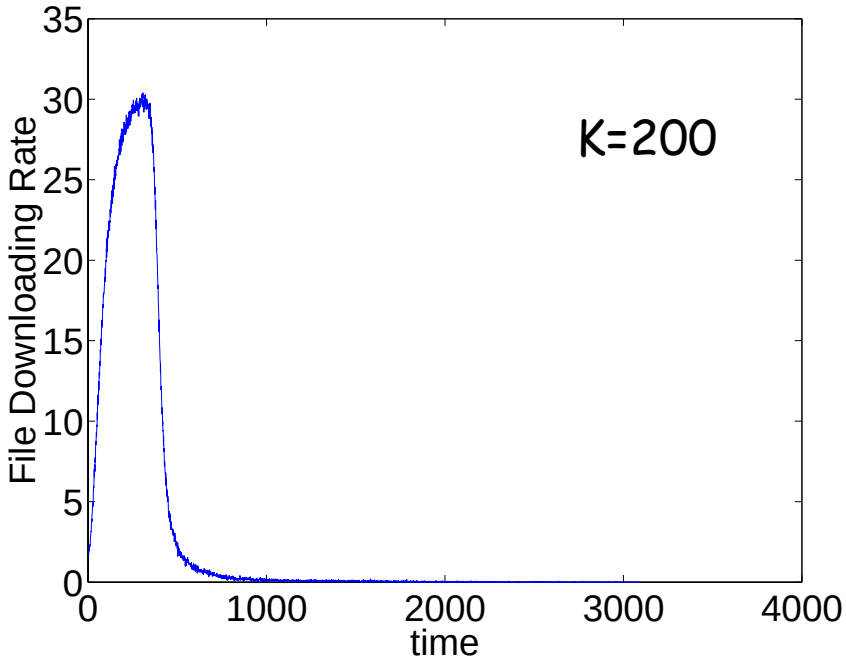
- Assumptions:
  - memoryless mobility
  - K is large
- Throughput capacity depends on node mobility and node density only
- When node density and mobility are high, throughput capacity depends on node density only

$$\lim_{K\to\infty} C_4 = \lambda_0 \beta(N) \approx \beta(N)$$

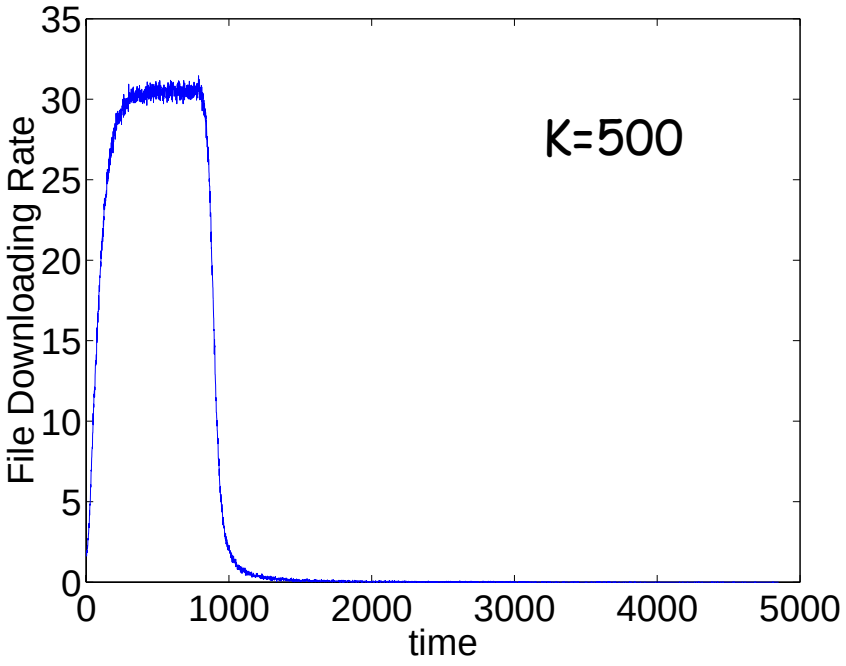
$$\lim_{K\to\infty} T_4 = \frac{K}{\beta(N)}$$

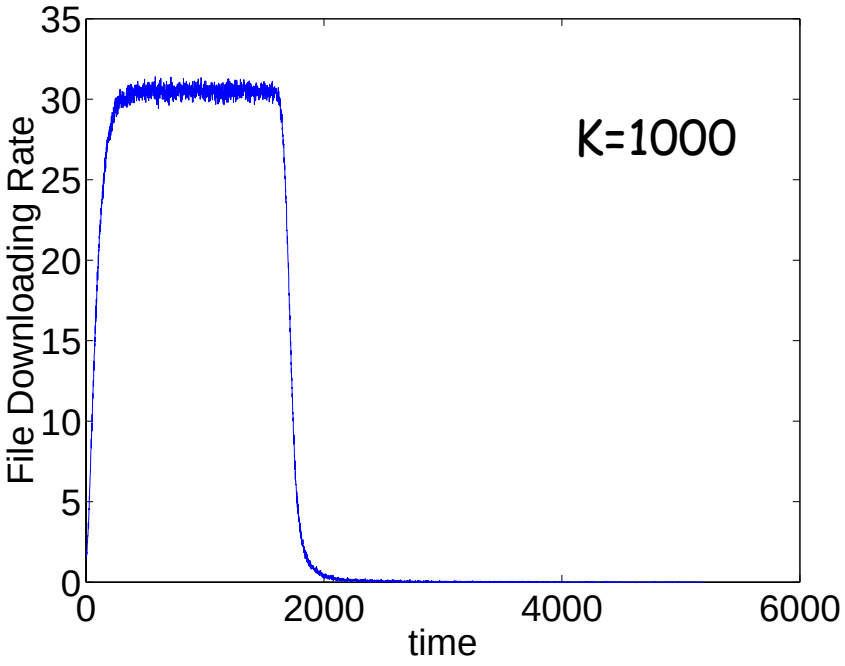
File Downloading Rate vs. time



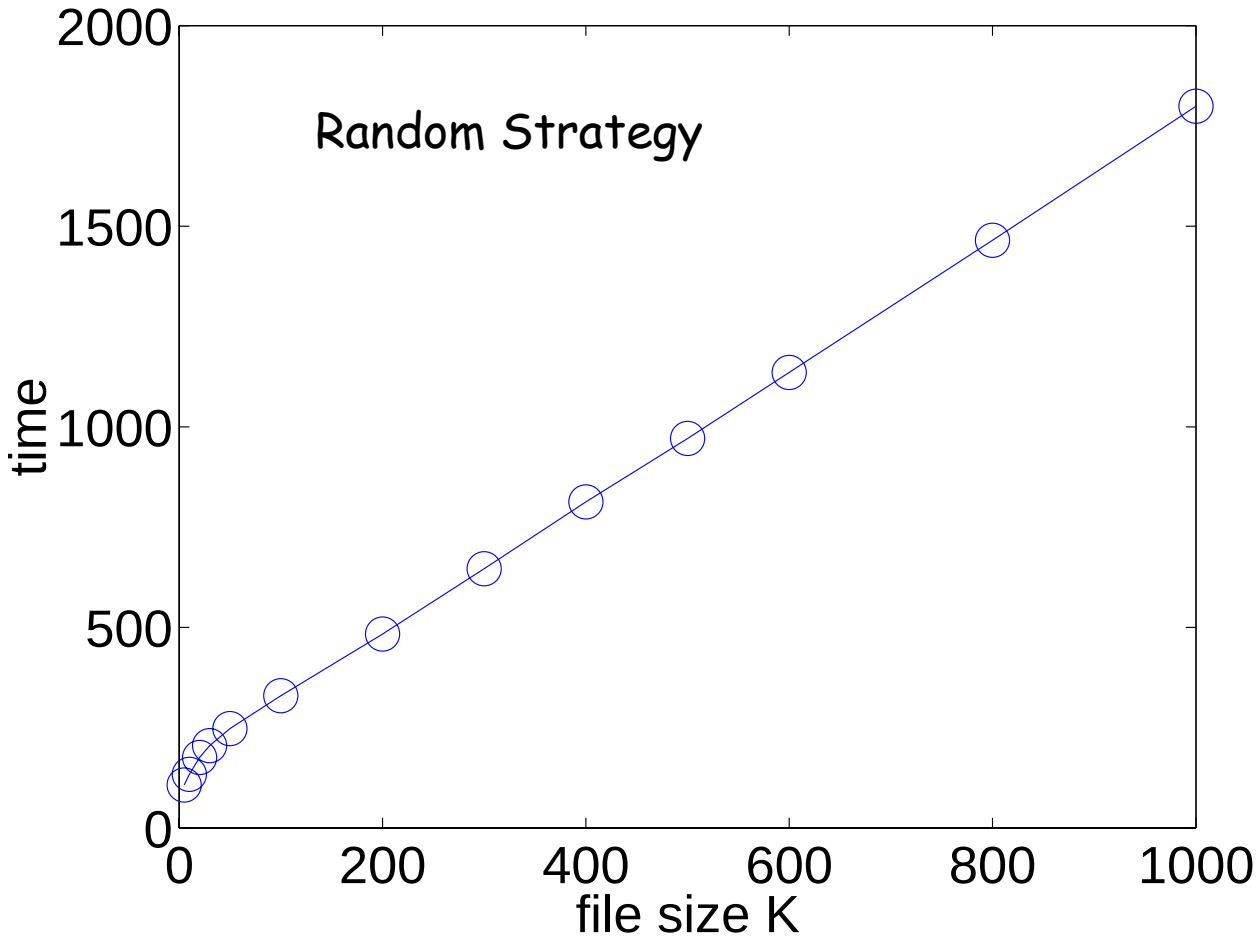


File Downloading Rate vs. time

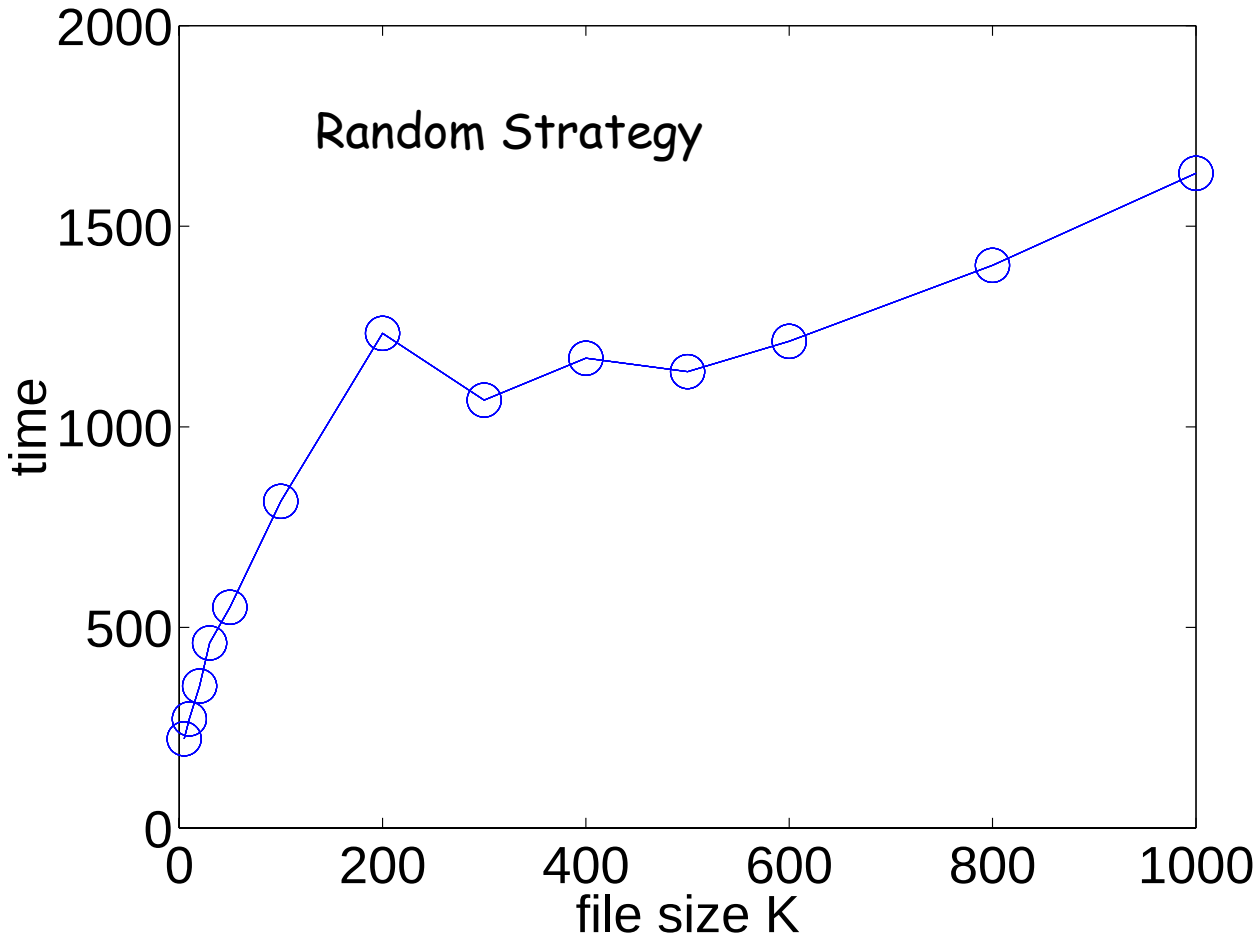




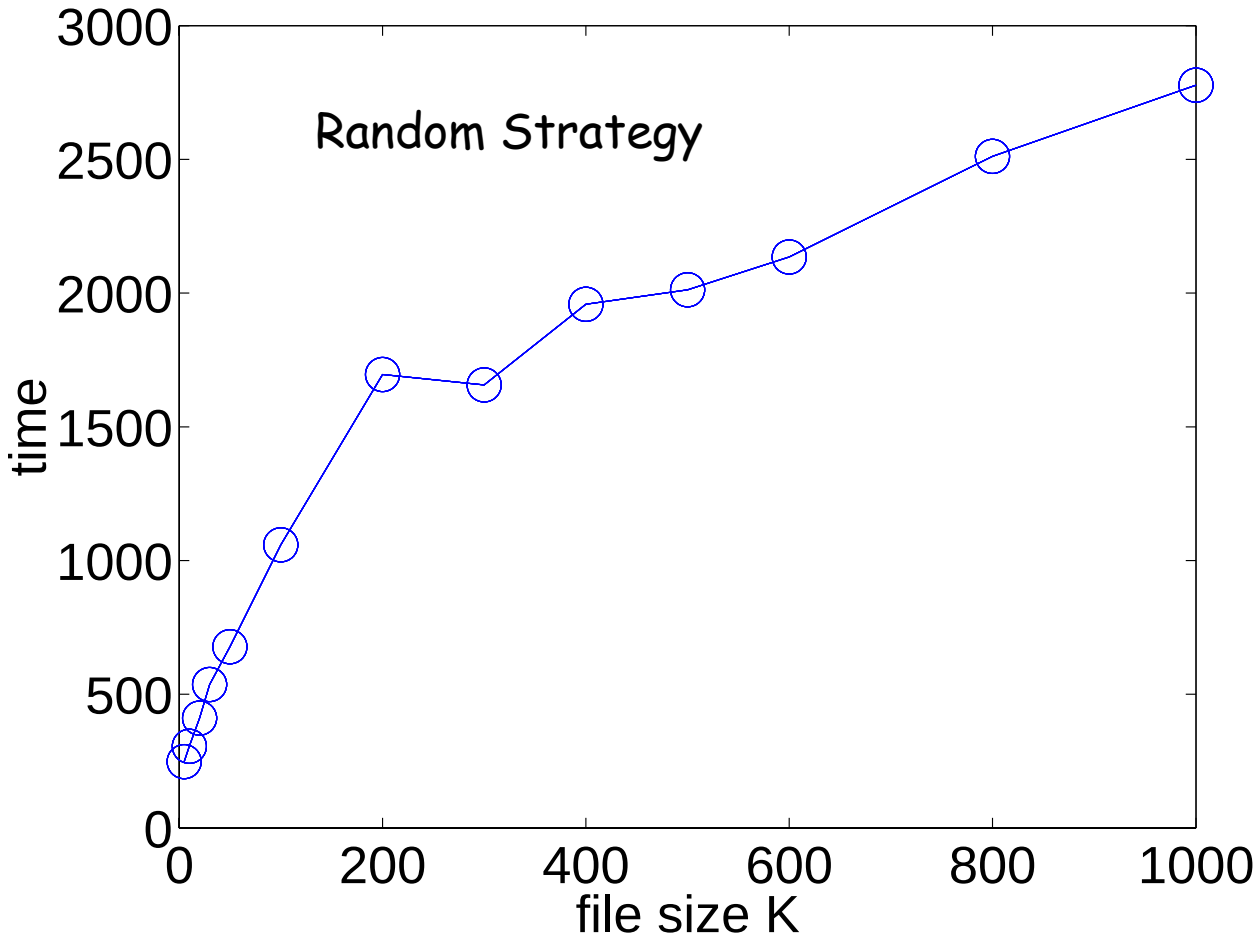
T_1 vs. number of files K



T_2 vs. number of files K



T_3 vs. number of files K



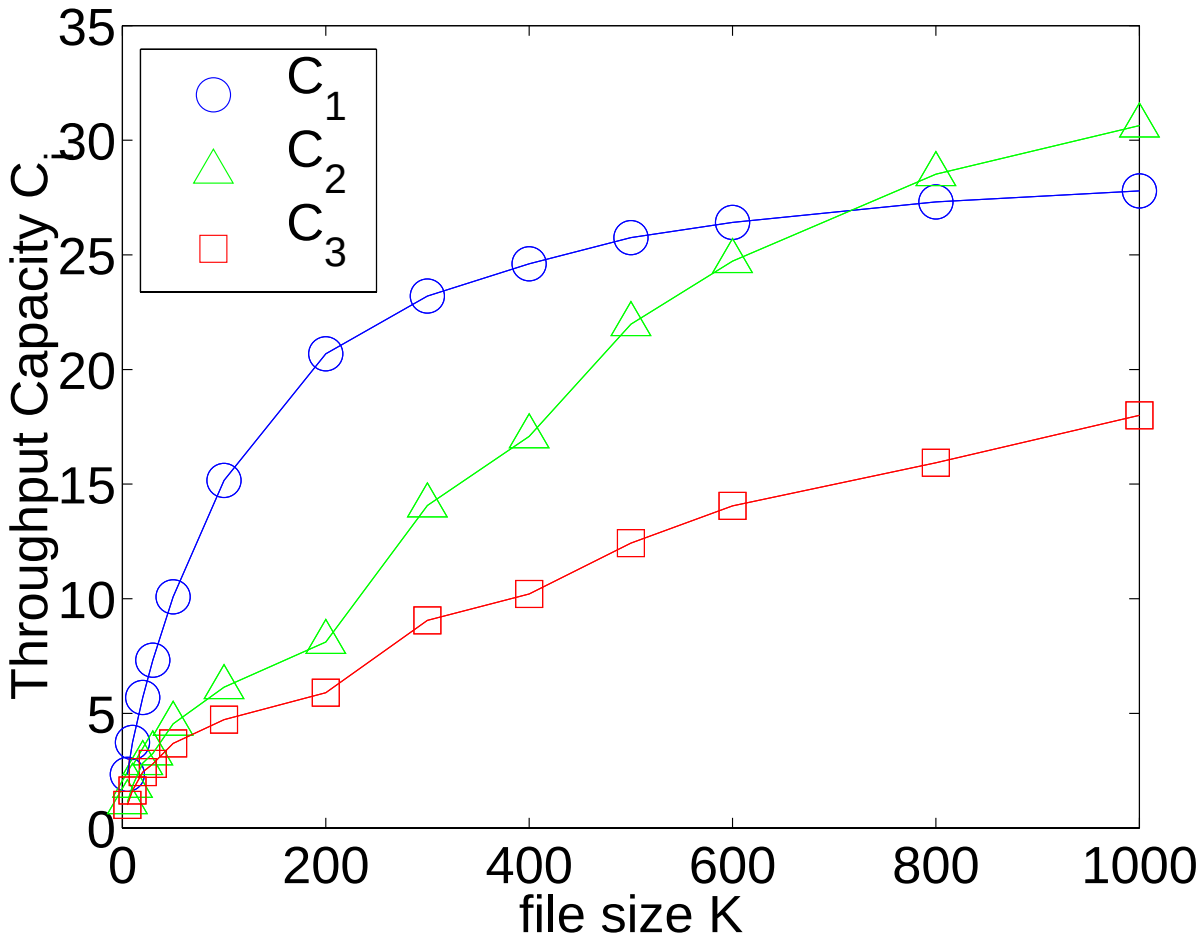
Observations

- Asymptotic performance for large K
  - Throughput capacity is constant
  - Time all nodes get all files (simulation) equals Time a random node get all files (analysis)
  - Networking is fair

$$\lim_{K\to\infty} E[T_3] = \lim_{K\to\infty} E[T_4]$$

$$C_3^{\infty} = C_4^{\infty}$$

C_i vs. number of files K



Data Diversity for Ad Hoc Networks

- Diversity exists when there is variation and selection
- Data diversity exploits the variation of individual preferences by having a large data set for selection
- Mechanism: when K is large, efficiency of file exchange increases
- Lesson: With multiple information dissemination services, content distribution is more efficient if data is inter-exchangeable

Extensions

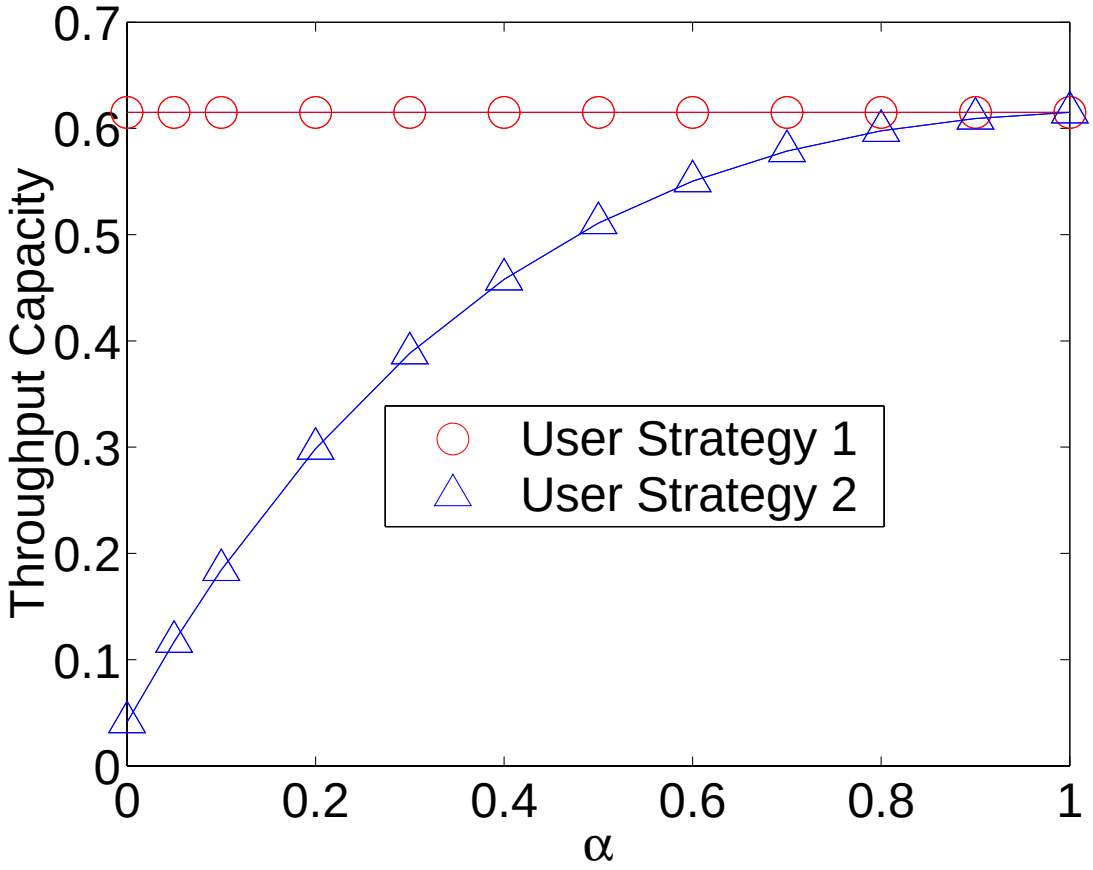
- Multiple Show Downloading
  - Example: TV shows, multiple movies
  - Multiple shows cached in the infostation
    - K files in total
  - Each node interested in αK files, each file wanted by αN users
  - Users have overlapping interests
  - User Strategy?
    - Download items of genuine interest
    - Download what I don't have
      - popular files I am not interested in might be good for exchange

User Strategy

- Download what I am genuinely interested:
  - Each file has αN networking agents for dissemination
  - Throughput capacity is smaller than for popular movie d/l
- Download what I don't have:
  - Each file has N networking agents for dissemination
  - Throughput capacity is the same as popular movie d/l
  - minimize my d/l time by helping the network in data dissemination!
  - Exploitation of multiuser diversity
- Lesson:
  - 1000 files at the infostation; I am interested in only 50 files.
  - To expedite the d/l of my 50 files, I enlarge my preference list and d/l whenever possible

Throughput Capacity vs. α

Strategy 1: d/l what I don't have
Strategy 2: d/l what I am interested in



Conclusion

- Random strategy has near optimal performance
- Network performance affected by node mobility, density and the number of files in the system K
- When K is large,
  - Throughput capacity of the network increases to C^{∞} (data diversity)
  - Networking time and capacity for different nodes are similar (Fairness)
- In content distribution, both multiuser and data diversity can be exploited