Combining D2D and content caching for mobile network offload

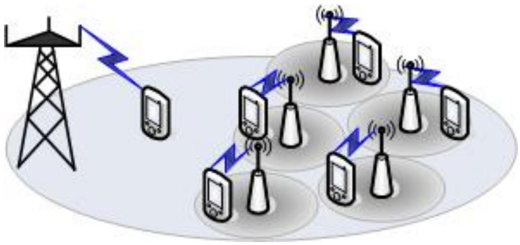
Salah Eddine Elayoubi – Orange Labs & IRT SystemX

Joint work with Antonia Masucci and Berna Sayrac

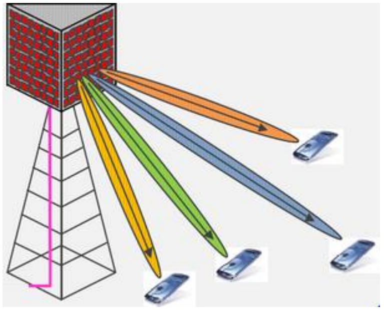
07/09/2015

Key 5G technologies

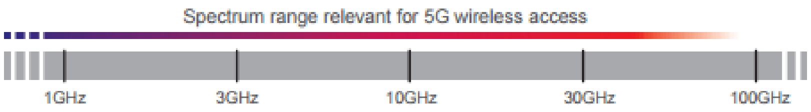
- Ultra-dense networks



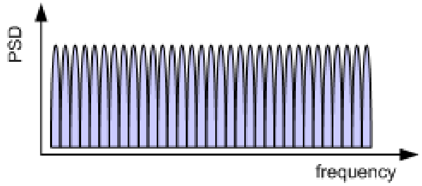
- Massive MU-MIMO



- Usage of high spectrum bands

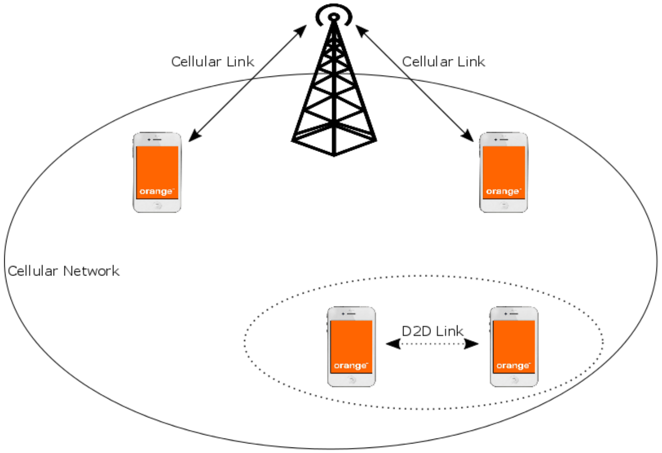


- With new waveforms?



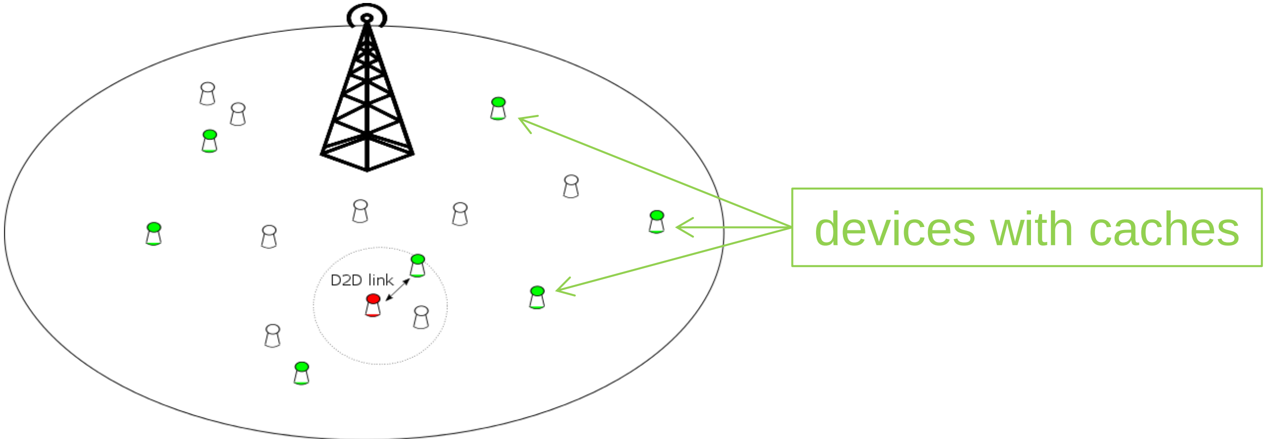
- The focus of this work is on D2D
  - short range communications
    - devices act as access points
  - cheap densification
    - no deployment and backhaul costs
- Key questions:
  - which traffic can be offloaded?
  - How to allocate resources to D2D?

- D2D communications



Our focus: traffic offload with D2D

- LTE-A (Rel.13) use cases for D2D:
  - Proximity-based services
  - Public safety services
- Our objective: offload the macro network
  - LTE-A services generate low traffic and are for specific scenarios
  - Use D2D for serving high bandwidth consuming traffic
- Video traffic is the best candidate: accounts for 40-60% of overall traffic
  - Difficulty: videos are not generated locally
  - Retrieved from distant servers in the Internet, or from caches in the core network
- Solution: caching in devices
  - All devices or a set of them act as helpers that cache popular contents
  - Users requesting a video retrieve it from a nearby device if available in its cache



Outline

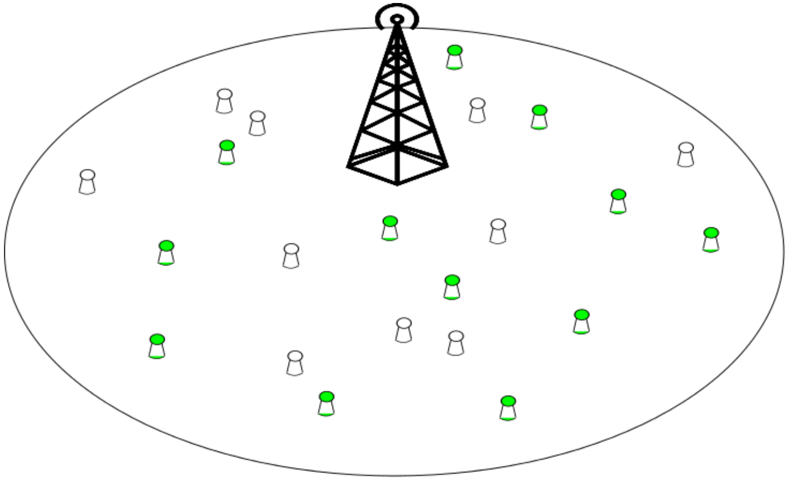
- System design principles
  - Caching policies
  - Resource allocation
- Model and results
- Conclusions

Design of the caching mechanism

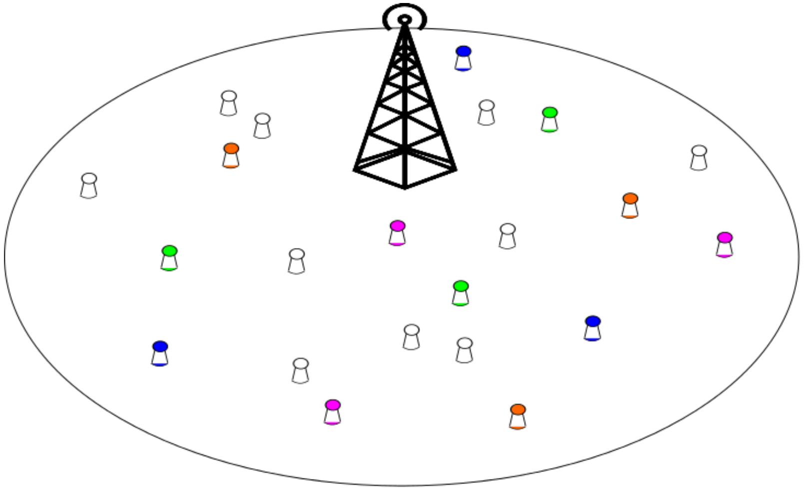
- Classical reactive caching methods (LRU and LFU)
  - not expected to be efficient as local demand is low
  - the cache does not fill sufficiently fast compared to popularity change (temporal locality)
- Pre-fetching strategies are the most viable:
  - popular contents are stored in caches
  - renewed periodically when the popularity distribution changes

- **Homogeneous bundle**

  most popular K contents in each cache

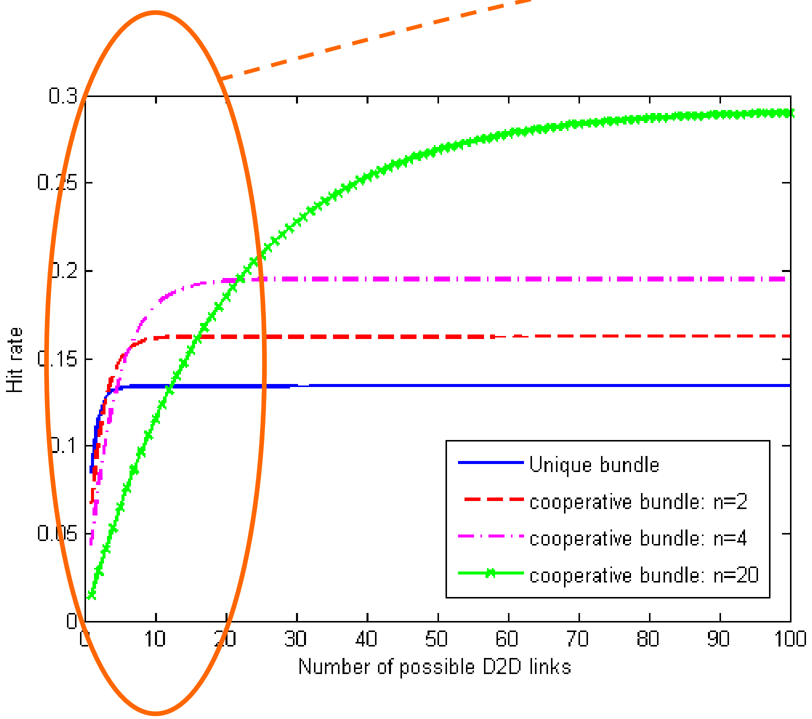


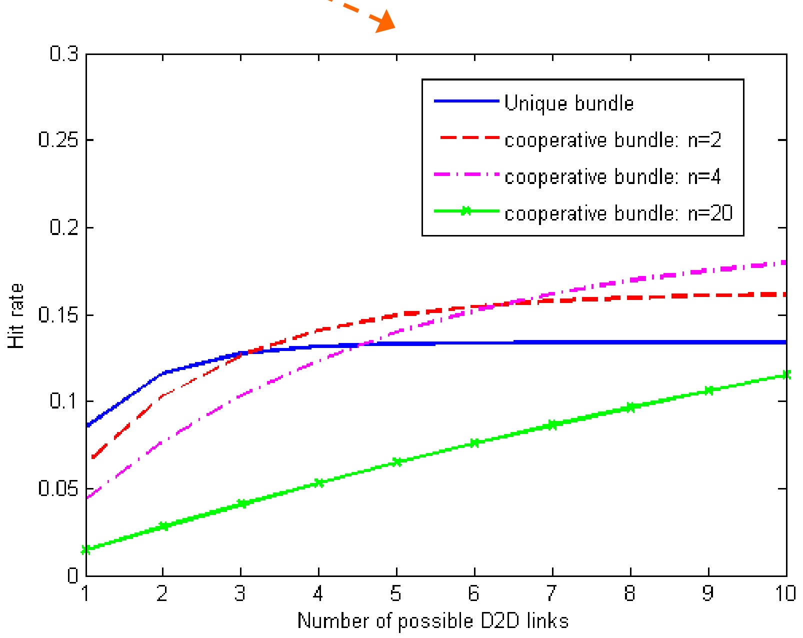
- **Cooperative bundle**
  - users with cache divided in n subclasses
  - subclass *i* caches contents numbered *i,i+n,i+2n,…*



Traffic offloaded by D2D communication

Cache size	200 contents
Catalog size	10^6 contents
Zipf(α)	α=0,8



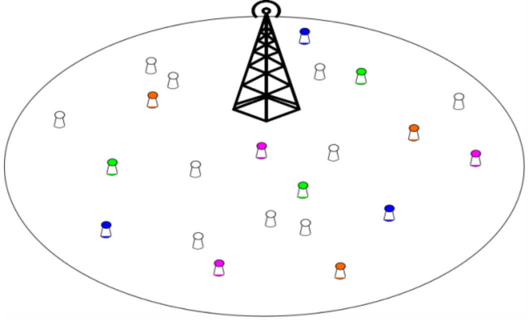


- Cooperative bundle preferred in dense D2D network
- If only a few devices are seen by each user, cache most popular files everywhere

Outline

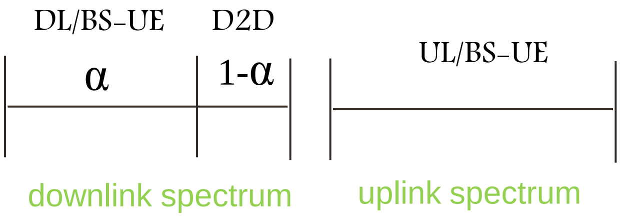
- System design principles
  - Caching policies
  - Resource allocation
- Model and results
- Conclusions

Operator-controlled D2D

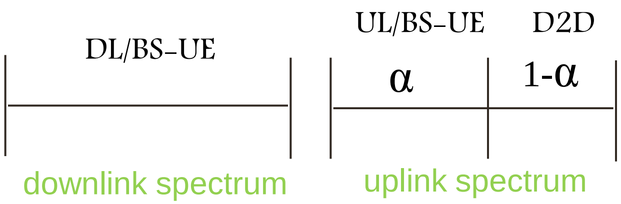


- There is no dedicated D2D spectrum
  - D2D communications are not “for free”
- D2D communications offload the downlink traffic
  - but the operator can allocate spectrum from its downlink or uplink pools

- Using downlink spectrum



- Using uplink spectrum



- The device asks his neighbors for the content
  - broadcast an interest packet (in ICN-like architectures)
- The base station allocates resources to the D2D links
  - no interference between D2D links in the coverage of one base station
  - interference may happen between D2D links in adjacent cells

Outline

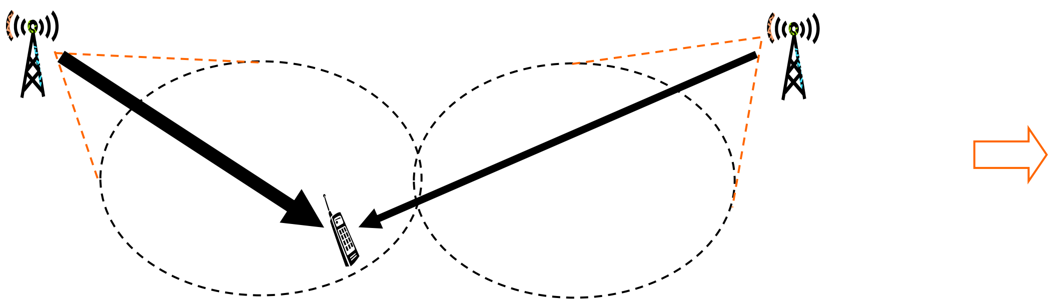
- System design principles

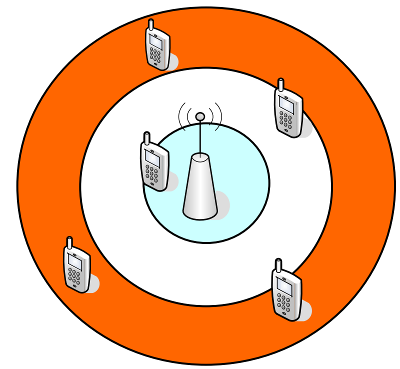
- Model and results
  - Baseline model
  - Queuing model for D2D
  - Extension to multiple-hop D2D

- Conclusions

Baseline model: macro-to-device

- Low rate at cell edge: low received signal and large interference
- Cell divided into zones of radio conditions



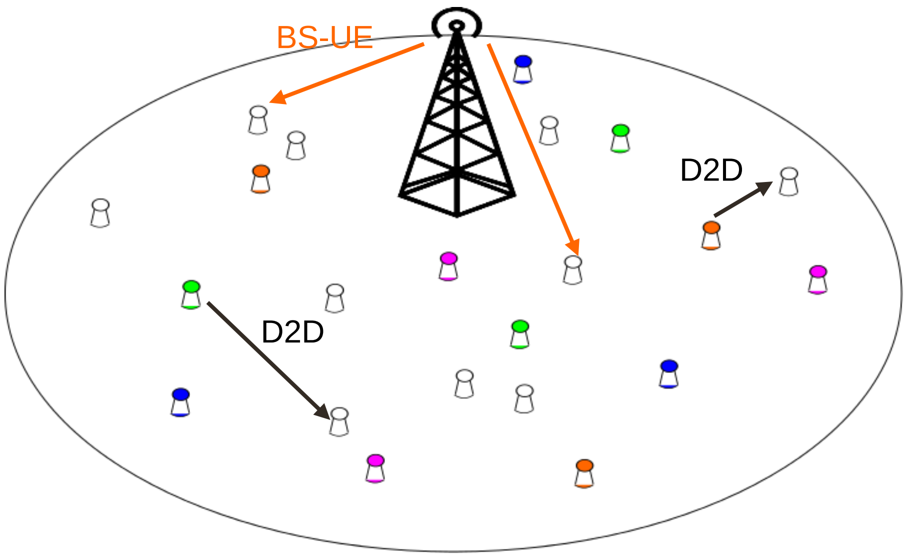


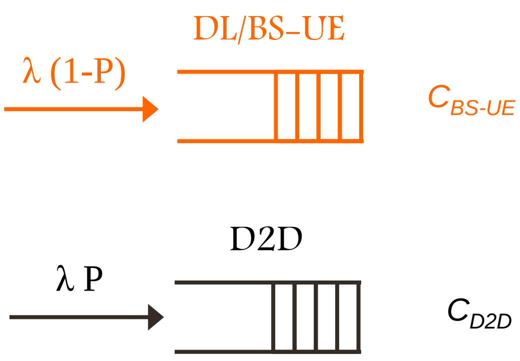
- Poisson rate λ_k for class *k* (flow/sec), each of size *V* (Kbit/flow)
- Traffic load generated by class *k* users is $\rho_k = \dfrac{\lambda_k V}{T_k}$
- Cell load is computed by $\rho = \sum_k \dfrac{\lambda_k V}{T_k} = \dfrac{\lambda V}{C_{BS-UE}}$; $\quad C_{BS-UE} = \left(\sum_k \dfrac{\lambda_k / \lambda}{T_k} \right)^{-1}$
- Cell is modeled as a Processor Sharing queue of capacity $C_{BS\text{-}UE}$ equal to the harmonic average of rates (Bonald et al., 2003)
- Stability condition: $\lambda V < C_{BS-UE}$

Outline

- System design principles
- Model and results
  - Baseline model
  - Queuing model for D2D
  - Extension to multiple-hop D2D
- Conclusions

D2D queuing model



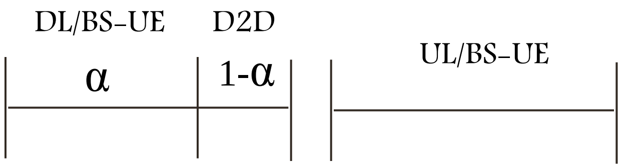


- BS-UE links still modeled as a PS queue on BS-UE spectrum
- D2D links are orthogonal, as resource allocation is centralized
  - the aggregation of D2D links modeled as a PS queue
- D2D rates depend on distances between devices
  - Capacity of D2D queue is the harmonic average of D2D rates
- D2D traffic volume depends on *P*, probability of using D2D
  - *P* depends on caching scheme

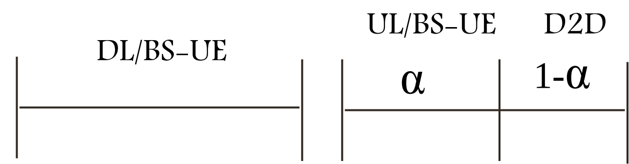
Optimal D2D resource allocation

Resource allocation:

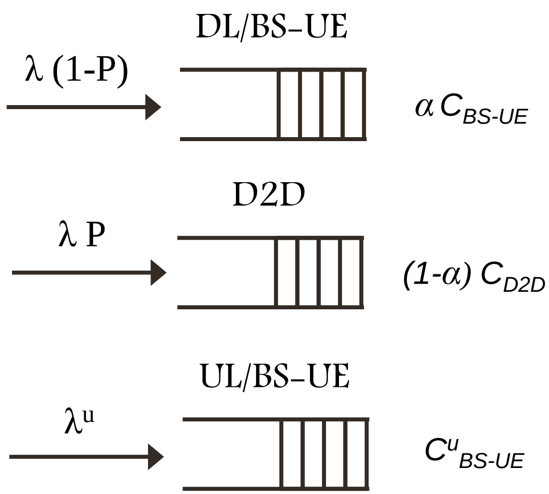
- Using downlink spectrum

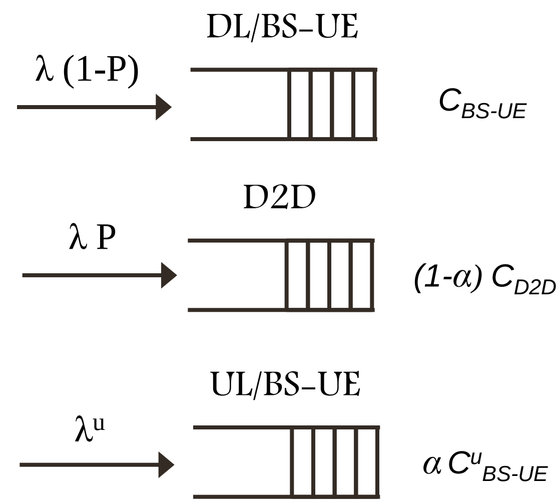


- Using uplink spectrum



Queuing model:





Stability condition:

Downlink spectrum:

$$\lambda V < \min\left(C_{BS-UE}\frac{\alpha}{1-P}, C_{D2D}\frac{1-\alpha}{P}\right)$$

$$\lambda^u V < C^u_{BS-UE}$$

Uplink spectrum:

$$\lambda V < \frac{C_{BS-UE}}{1-P}$$

$$\lambda^u V < \min\left(\alpha C^u_{BS-UE}, C_{D2D}\frac{1-\alpha}{P}\right)$$

Optimal allocation:

Downlink spectrum:

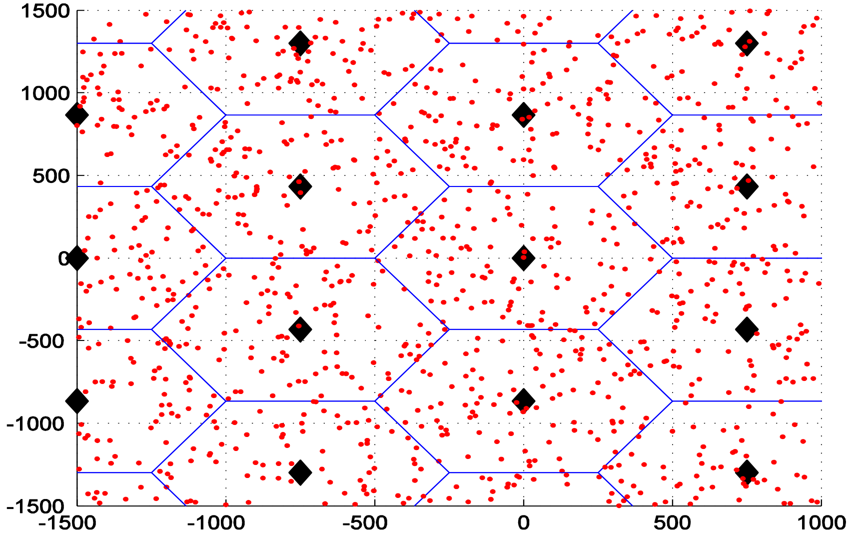
$$\alpha^* = \frac{C_{D2D}/P}{C_{D2D}/P + C_{BS-UE}/(1-P)}$$

Uplink spectrum:

$$\alpha^* = \frac{C_{D2D}/P}{C_{D2D}/P + C^u_{BS-UE}/r^u}; \quad r^u = \frac{\lambda^u}{\lambda}$$

Simulation setting: Urban Outdoor

- **Cellular layout**



- Compute SNRs for BS-UE and D2D links
- Deduce rates and average them for capacity

- **Simulation parameters**

Cell Radius	500 m
Macro BS pathloss	ITU Model
D2D pathloss	Winner Model
Tx/Rx antennas	MIMO 2x2
Carrier frequency	2 GHz
Bandwidth	10 MHz
Thermal Noise	-174 dBm
Noise Figure	9 dBm

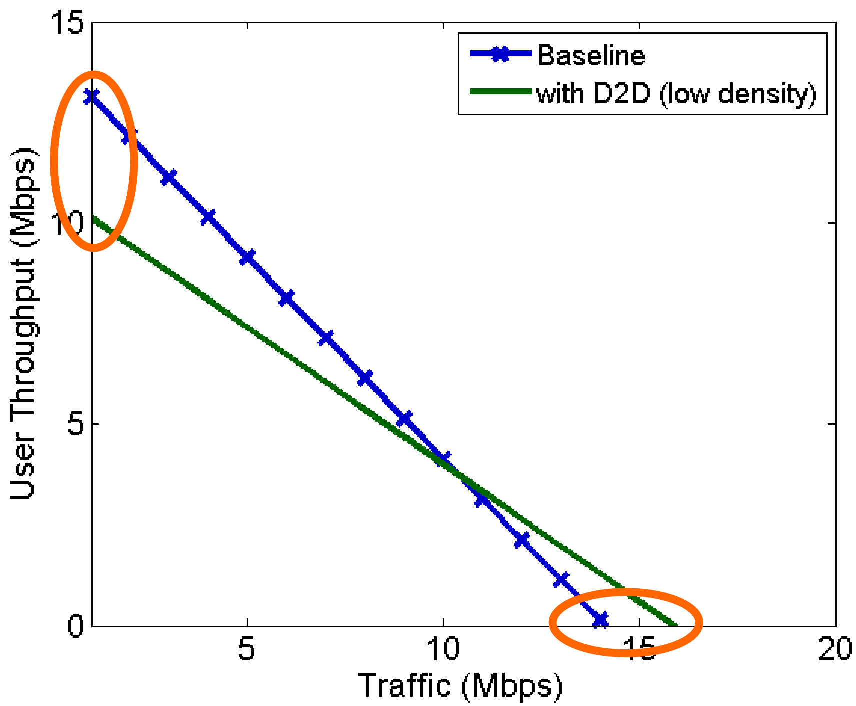
- **Queue Capacities, for a 20MHZ allocation**

BS-UE	14,1 Mbps
UE-BS	10,8 Mbps
D2D	31,3 Mbps

D2D on the downlink spectrum

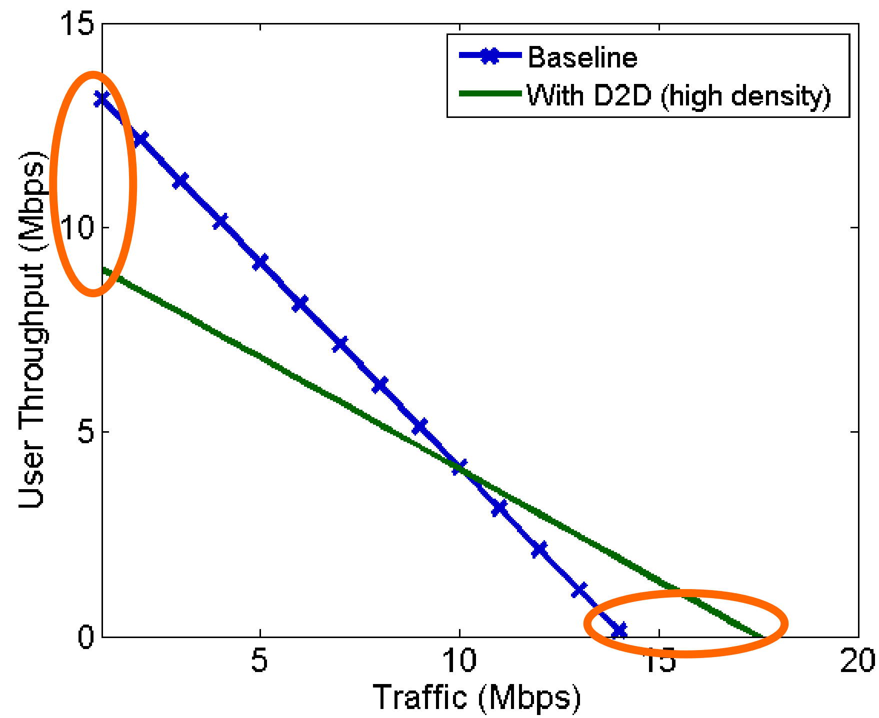
5 devices in proximity; hit rate P=0,2

Optimal resource allocation: α*=0,9



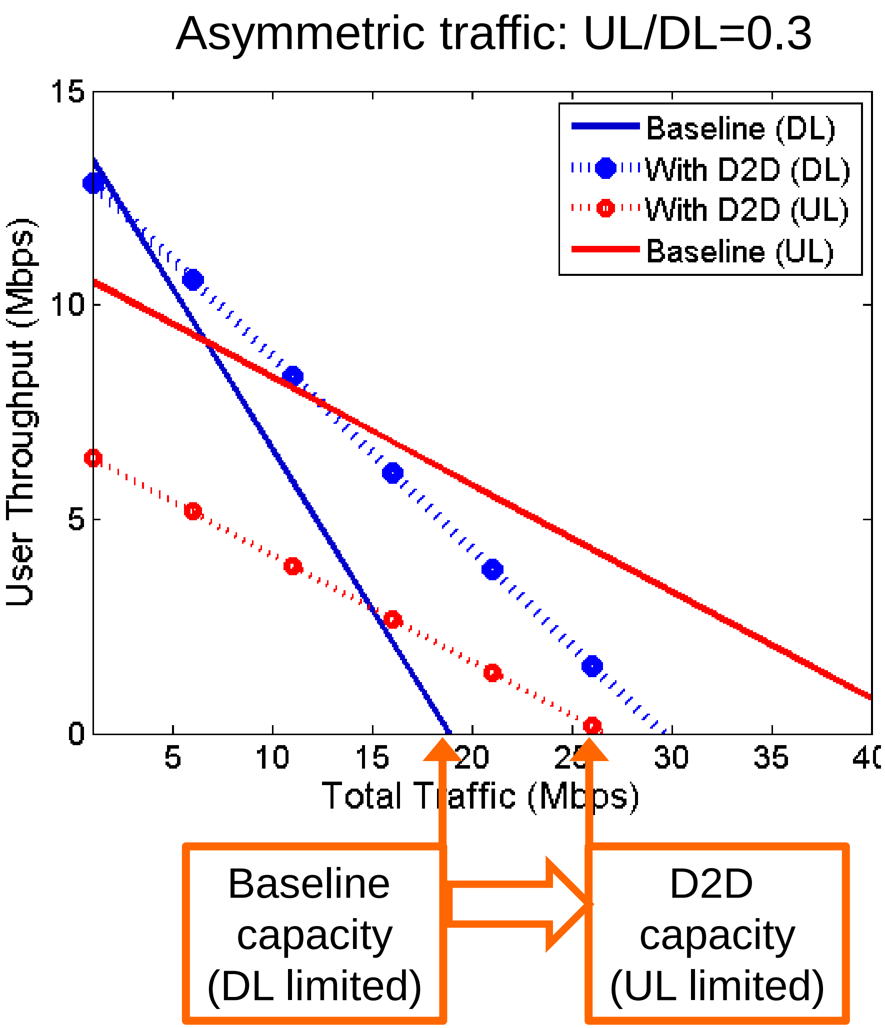
50 devices in proximity; hit rate P=0,35

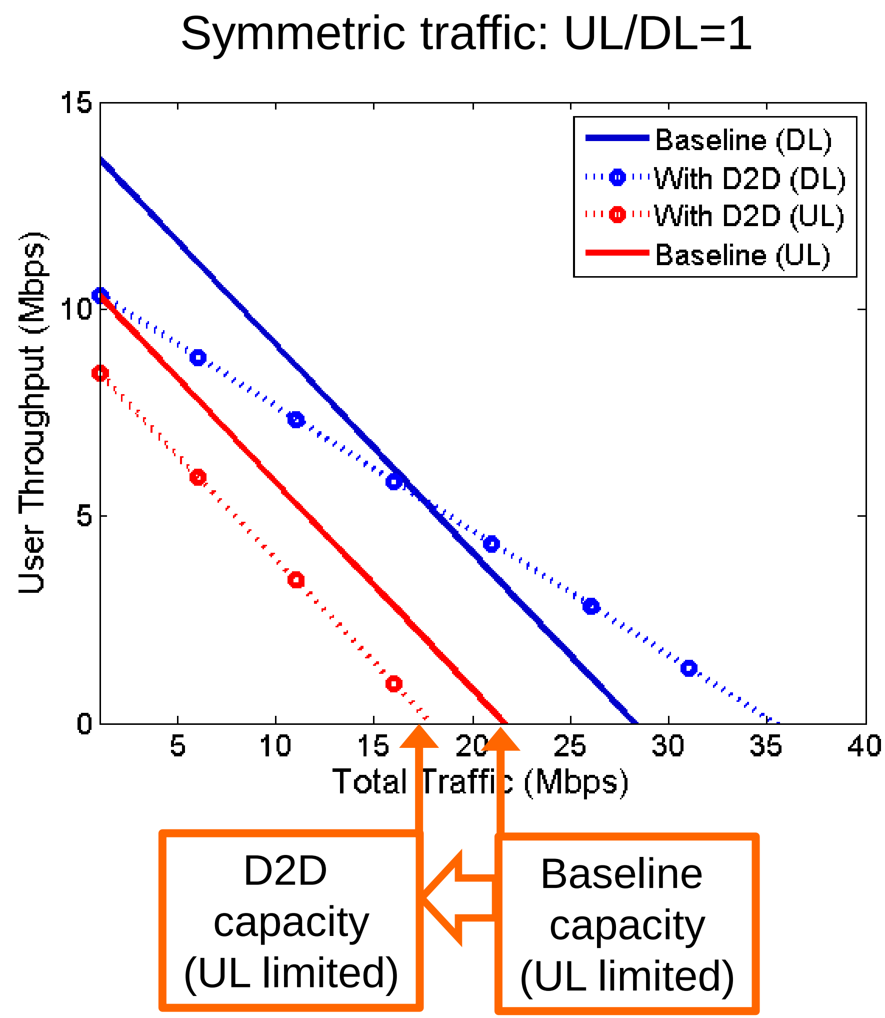
Optimal resource allocation: α*=0,6



- D2D increases the stability condition
  - but decreases the throughput for low traffic
- The design of the system (α*) depends largely on the hit rate (density of devices, efficiency of the caching solution)

D2D on the uplink spectrum



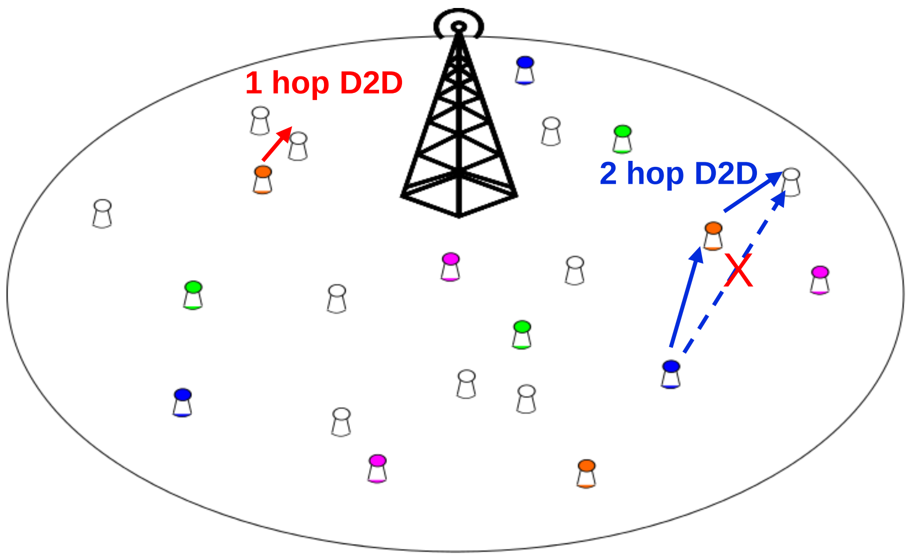


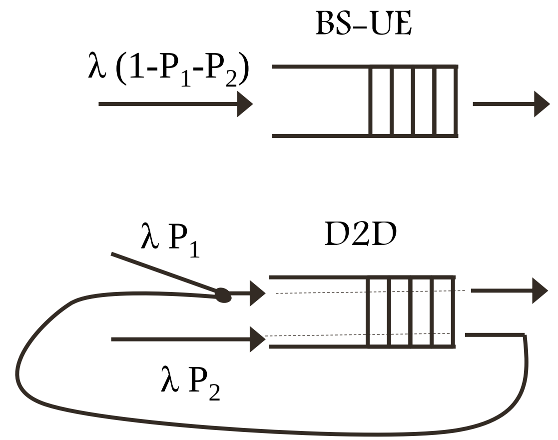
- In traffic asymmetry case, D2D can increase capacity by a natural load balancing between UL and DL
- When uplink traffic is high, D2D has to be performed on DL spectrum

Outline

- System design principles
- Model and results
  - Baseline model
  - Queuing model for D2D
  - Extension to multiple-hop D2D
- Conclusions

Multihop D2D

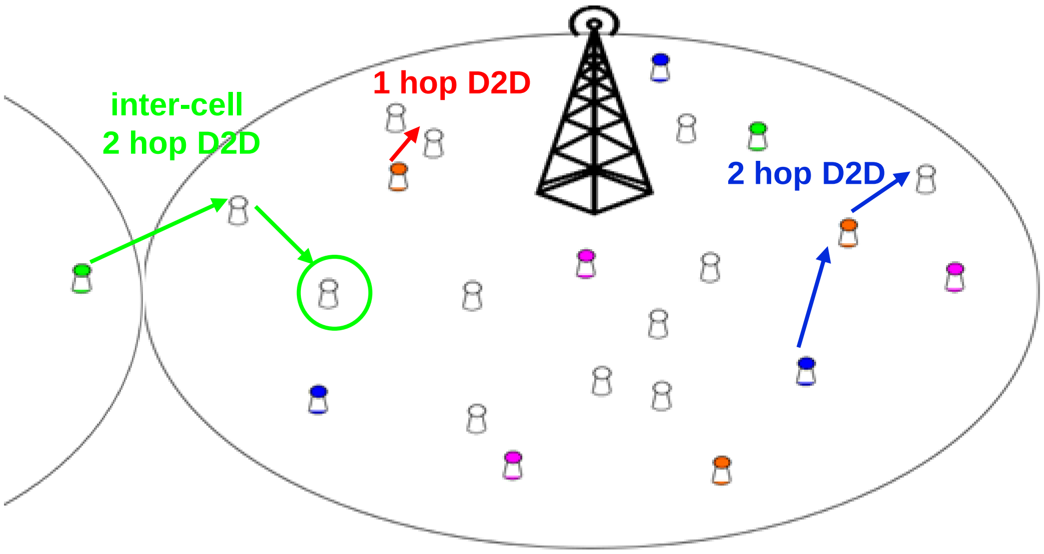


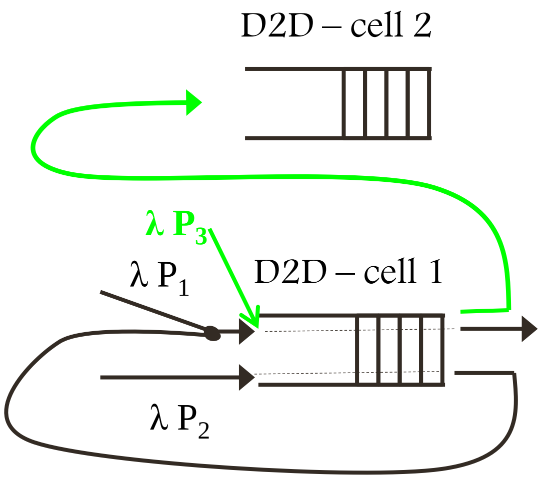


P_1: proba of finding content within 1 hop
P_2: proba of needing 2 hops

- D2D communications are limited in range
- multiple hops can increase offload capacity: larger number of devices
- The system is modeled as a queue with multiple classes
  - a 2-hop flow re-enters the system as a 1-hop flow
- Stability condition: $\lambda V < \min\left(C_{BS-UE} \frac{\alpha}{1-P_1-P_2}, C_{D2D} \frac{1-\alpha}{P_1+2P_2} \right)$
- Optimal D2D resource allocation: $\alpha^* = \frac{C_{D2D} \big/ (P_1+2P_2)}{C_{D2D} \big/ (P_1+2P_2) + C_{BS-UE} \big/ (1-P_1-P_2)}$

Inter-cell Multihop D2D

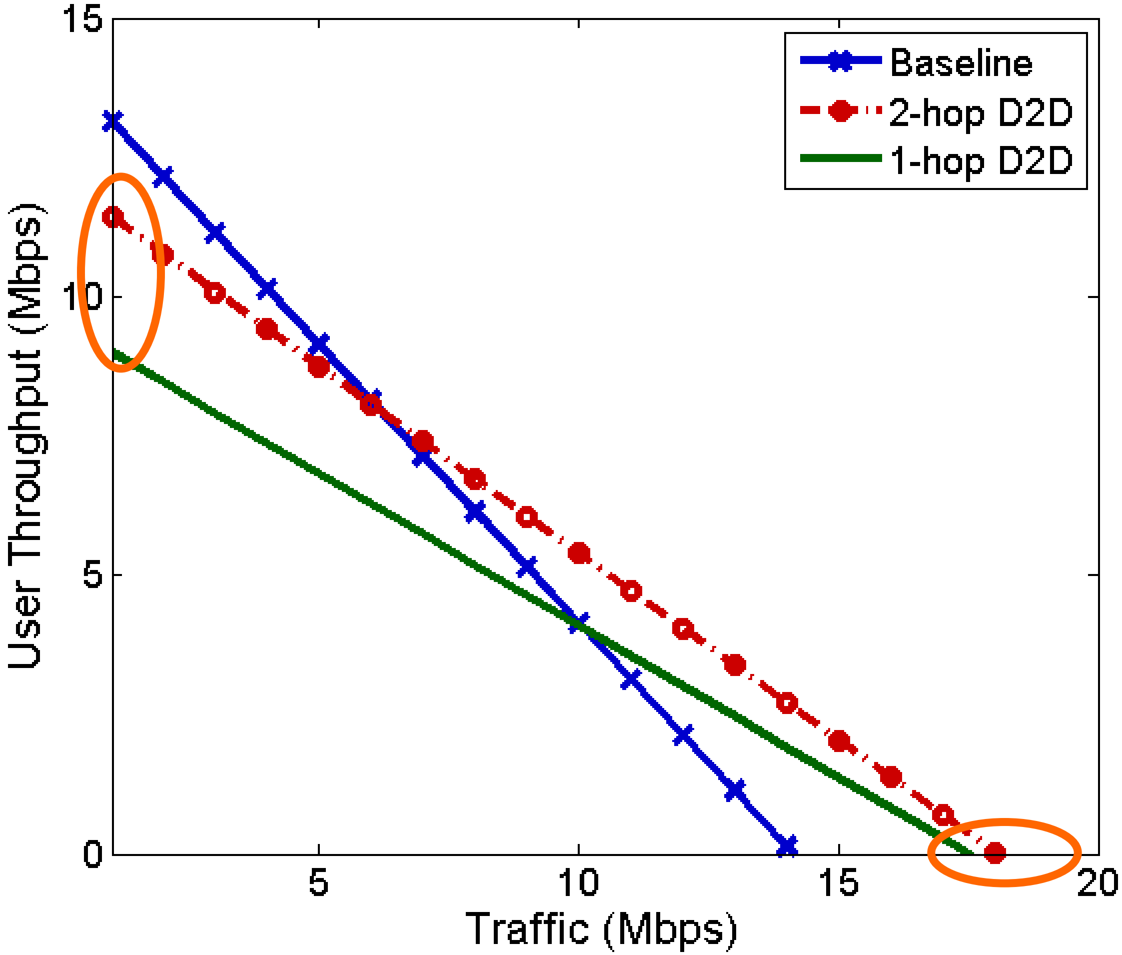


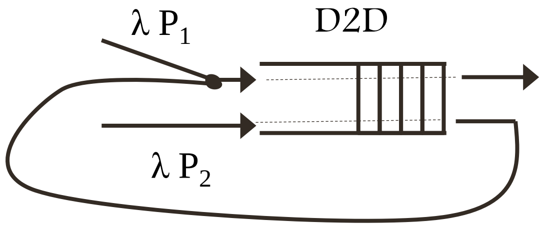


P_1: proba of finding content within 1 hop
P_2: proba of needing 2 hops (intra-cell)
P_3: proba of needing 2 hops (inter-cell)

- Inter-cell D2D may increase the hit rate
- However, it introduces coupling between adjacent cells
- Gain is not obvious, as cells have to serve additional traffic

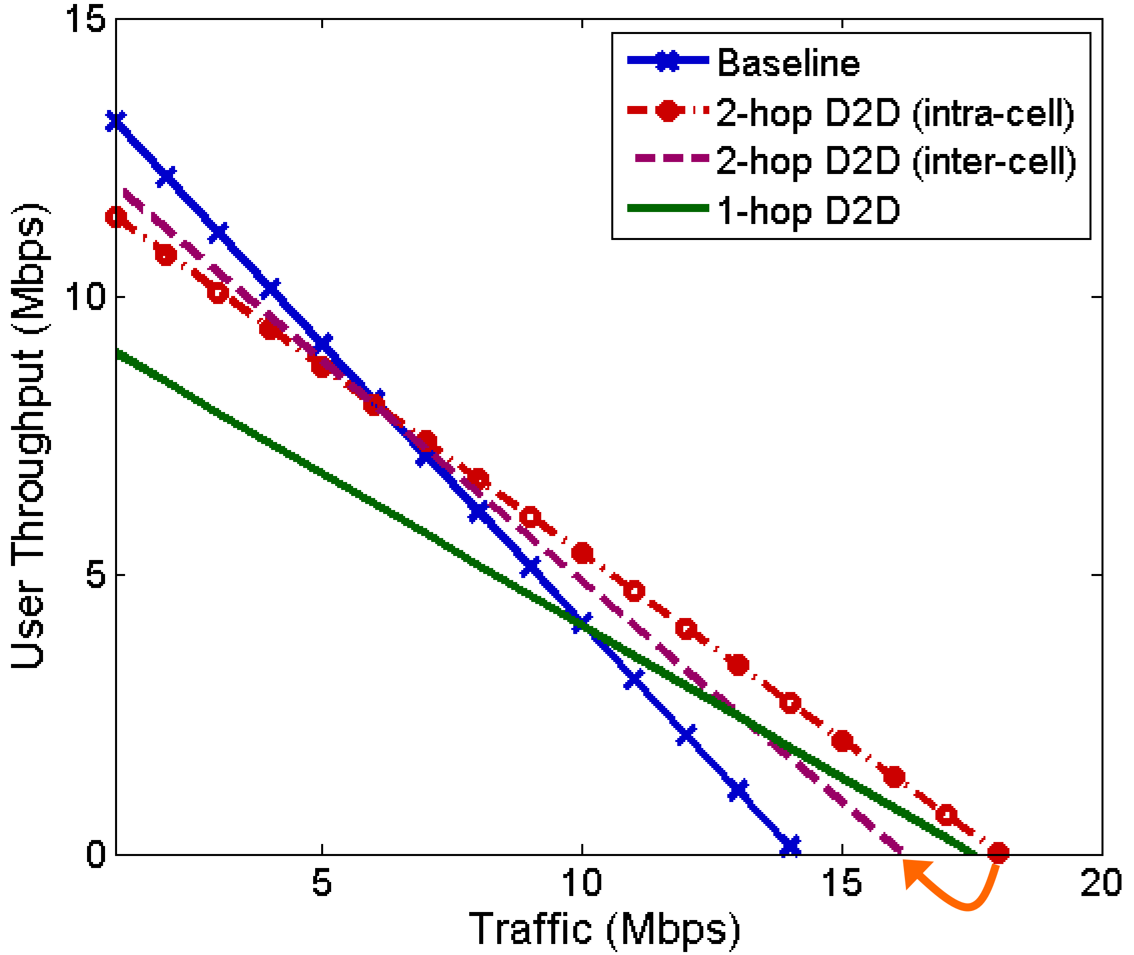
Is intra-cell multihop D2D effective?

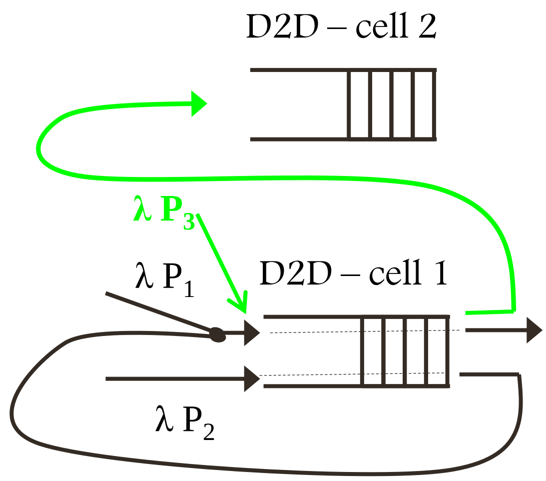




- Intra-cell Multihop brings a better capacity and a better throughput at low traffic

Is inter-cell multihop D2D effective?





- Intra-cell multihop brings a better capacity and a better throughput at low traffic
- Inter-cell multihop decreases capacity, as cells are already too loaded to serve additional traffic

Outline

- System design principles
- Model and results
- Conclusions

Conclusion and future works

- A queuing theory analysis for D2D communications
  - Closed-form solutions for the capacity region

- D2D combined with caching provide capacity gains
  - depend on the efficiency of the caching strategy
  - resource allocation is to carefully optimized

- Future works:
  - Analysis of efficient caching strategies
    - optimal caching (as in Błaszczyszyn and Giovanidis, 2015)
    - coded caching (as in Altman et al., 2014)
    - in the framework of METIS-II 5GPPP project

  - D2D communications as a way for greening the mobile network
    - complement the coverage of cellular networks
    - allow a switch off some of the infrastructure nodes
    - in the framework of SooGreen European project

METIS II

SooGREEN

Thank you!

salaheddine.elayoubi@orange.com
mariaantonia.masucci@orange.com
berna.sayrac@orange.com