ENERGY CONSUMPTION OPTIMIZATION THROUGH PRE-SCHEDULED OPPORTUNISTIC OFFLOADING IN WIRELESS DEVICES

Constandinos X. Mavromoustakis
Department of Computer Science
University of Nicosia, Cyprus

Andreas Andreou
Faculty of Computer Science and Technology
University of Cambridge, UK

George Mastorakis, Stelios Papadakis
Department of Business Administration
Technological Educational Institute of Crete, Greece

Athina Bourdena, Dimitris Stratakis
Department of Informatics Engineering
Technological Educational Institute of Crete, Greece

August 24, 2014 to August 28, 2014

OUTLINE

- MOTIVATION
- PROPOSED FRAMEWORK
- WHEN TO OFFLOAD?
- SIMULATION RESULTS AND OBSERVATIONS

Ίδρυμα Προώθησης Έρευνας

MOTIVATION

Motivation

- In human life Mobile devices e.g., smartphone, tablet pcs, etc) become an essential part of
- Dream of "**Information at your fingertips anywhere anytime",**
- When compared to conventional information processing devices these Mobile devices are lack in resources.

Soultion

- Mobile Cloud Computing (MCC)

WHAT IS MOBILE CLOUD COMPUTING?

- *MOBILE CLOUD COMPUTING = MOBILE COMPUTING + CLOUD COMPUTING*
- MCC refers to an infrastructure where either (or both) data storage and data processing happen outside of the mobile device.
  - To another mobile device in most cases (communicational issues in accessing the other devices)
- Mobile cloud applications move the computing power and data storage away from the mobile devices and into other computing platforms located in clouds (or To another mobile device in most cases) which are then accessed over the wireless connection based on a thin native client

eurostars Ίδρυμα Προώθησης Έρευνας

Aims and objectives for MCC

- Mobile devices face many resource challenges (battery life, storage, bandwidth etc.)
- Cloud computing offers advantages to users by allowing them to use infrastructure, platforms and software by cloud providers at low cost and elastically in an on-demand fashion.
- Mobile cloud computing provides mobile users with data storage and processing services in clouds, obviating the need to have a powerful device configuration (e.g. CPU speed, memory capacity etc), as all resource-intensive computing can be performed in the cloud.

•**APPLICATIONS: Mobile Commerce; Mobile HealthCare; Mobile Learning; Mobile Gaming etc.**

ADVANTAGES

- Extending battery lifetime
- Improving data storage capacity and processing power
- Improving reliability and availability
- Dynamic provisioning
- Scalability
- Multi-tenancy
- Ease of Integration

Issues to deal with..

- **Mobile communication issues**:
  - **Low bandwidth**: One of the biggest issues, because the radio resource for wireless networks is much more scarce than wired networks
  - **Service availability**: Mobile users may not be able to connect to the cloud to obtain a service due to traffic congestion, network failures, mobile signal strength problems
  - **Heterogeneity**: Handling wireless connectivity with highly heterogeneous networks to satisfy MCC requirements (always-on connectivity, on-demand scalability, energy efficiency) is a difficult problem

eurostars
Ίδρυμα Προώθησης Έρευνας

This paper proposes..

- An offloading resource mechanism, which is used in collaboration with an energy-efficient model. The scheme exploits an offloading methodology in order to guarantee that no intermittent execution will occur on mobile devices, whereas the application explicit runtime will meet the required deadlines to fulfil the QoS requirements; and
- Elaborates on the development of an offloading scenario, in which the scheduling policy for guaranteeing the efficiency in the execution of mobile users' tasks/applications can be achieved in an energy-efficient manner.

Ίδρυμα Προώθησης Έρευνας

Pre-scheduled offloading mechanism

- Due to the heterogeneity in the hardware of both mobile devices and the servers on the cloud that the resources will be potentially (based on the proposed scheme) offloaded, the proposed framework encompasses the execution environment volatility and considers the cloud servers' response time, in order to a-priori compare them and select the appropriate server, according to the best fit-case.

Model used..

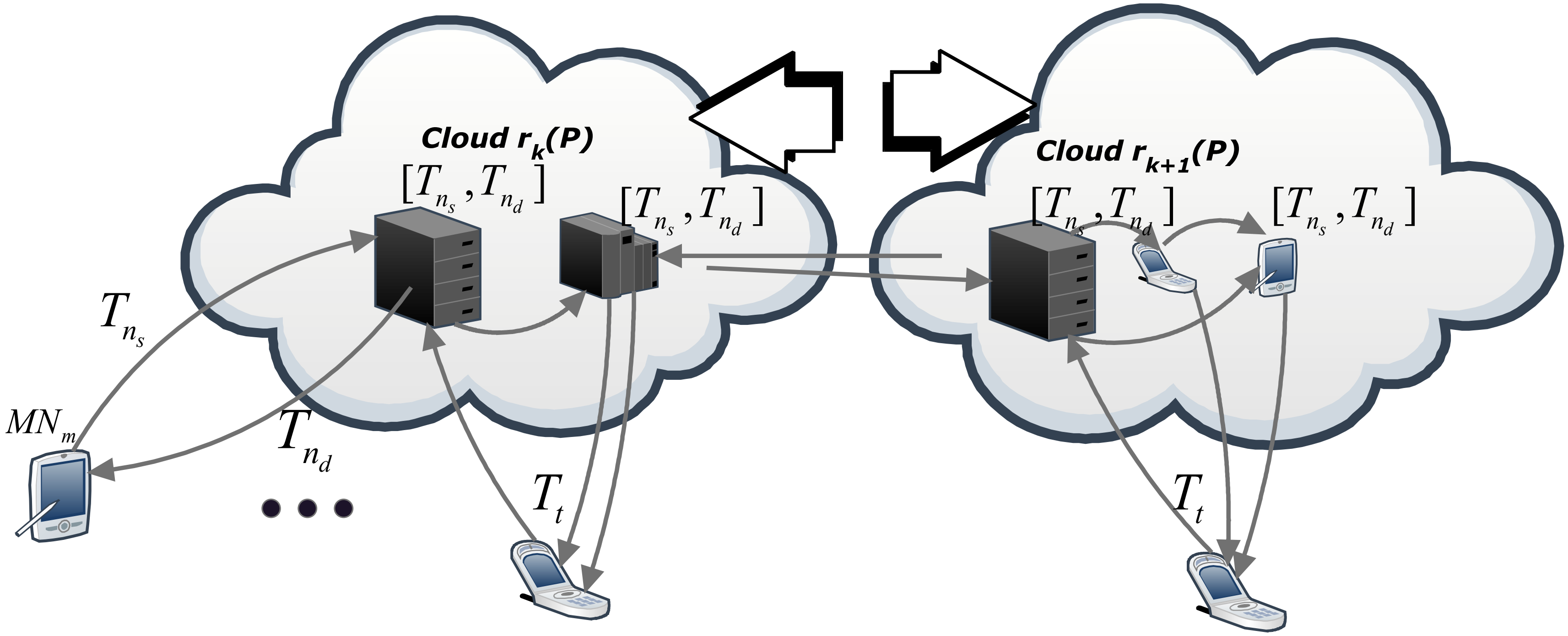
- the modelled parameters can be expressed, for an offloading process *O for an* executable resource task , as a 5-tuple given by:

$$O_{a_j}(MN) = <n_s,\ T_{n_s},\ T_{n_d},\ BW,\ T_t>$$

for the a_j executable task, where n_s is the devices or cloud terminals that the a_j from MN device will be offloaded, T_{n_d} is the source location best effort access time, T_{n_s} is the destination device or cloud location best effort access time from the source, ***BW*** is the required connection bandwidth and T_t is the connection holding duration for the $\boldsymbol{a_j}$ executable resource task.

The BW is set to be evaluated for : $1MBps \le BW \le 20MBps$

Cloud configuration and offloading process in order to achieve the best effort processing on-device power.



eurostars Ίδρυμα Προώθησης Έρευνας

Scheduled tasks

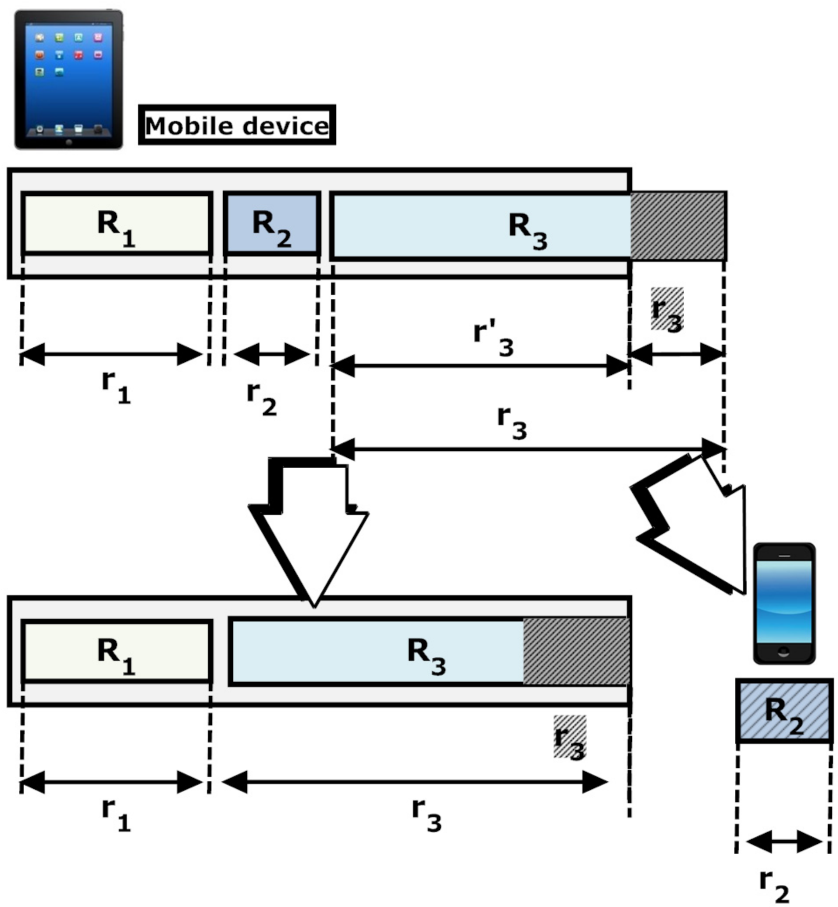
- All tasks are "ranked" according to the resources that are requiring for smooth run/execution onto mobile device.

- This work uses a dynamic offloading scheme, by considering executable resource tasks (t_x partitions) that will be offloaded according to the scheme proposed in the next section to conserve energy.
- Partitionable tasks are offloaded onto other devices or cloud terminals based on the evaluation mechanisms
  - Cost of running the task on a nearby peer
  - Energy consumption of each device should satisfy some metrics/evaluations for each of the resources (executable processes) running onto the device hosting the r_i (resources).

Our proposed scheme aims at conserving energy on each mobile device, while running energy draining applications.

eurostars
Ίδρυμα Προώθησης Έρευνας

Resource partitioning onto mobile device.



r_i resources

Energy-consumption model using temporal capacity measurements

- The measurable energy consumption can be evaluated according to the:

relative energy consumption

processing time of the server

$$E_{r(a_j)} = E_c(a_j) \cdot \frac{C}{S_{a_j}}$$

energy consumption, the parameter C is the number of instructions that can be processed within T_t

Relative energy consumption

resources processing instruction cost for the computation resources

$$E_c(r_i) = \frac{Cost_{c(r_i)}}{S_{c(r_i)}} \cdot W_c$$

energy consumption of the device or server in mW

server processing instruction speed for the computation resources

Relative energy consumption/Cost for offloading

- Hence, for the neighboring devices *N* within *2-hops vicinity coverage* (based on the maximum signal strength and data rate model [1]) should stand:

r_i resources

$$\frac{Cost_{c(r_i)}}{S_{c(r_i)}} \cdot W_c \Big|^{r_i} > \frac{Cost_{c(r_i)}}{S_{c(r_i)}} \cdot W_c \Big|^{1,2..N}$$

$$W_{r_i} > W_c \forall 1,2,3,\ldots N$$

Dynamic Resource-based offloading scheme

TABLE I. DYNAMIC RESOURCE-BASED OFFLOADING SCHEME

Inputs: MN_m, Location($[T_{n_s}, T_{n_d}]$), resources $r_{1,} r_2, r_3, .. r_i \forall MN_m$

for all Cloud devices that stands $r_{1,} r_2, r_3, .. r_i$ find the r_i that can be offloaded to run onto another device

for all MN_{m-1} do{

Estimate the N_i //(as derived in (4))

if (N_i is above a threshold){

while (T_t == TRUE) {

while ($1 \leq t_x \leq$ z*P)

search for MN_{m-1} device that satisfies

$$\frac{Cost_{c(r_i)}}{S_{c(r_i)}} \cdot W_c \Big|^{r_i} > \frac{Cost_{c(r_i)}}{S_{c(r_i)}} \cdot W_c \Big|^{1,2..N}, \; W_{r_i} > W_c \forall 1,2,3,\ldots N$$

offload ($r_i, MN_{k(i)}$) //to $MN_{k(i)}$ to execute resource *(i)*

end **while**

end **while** ($C_{a_j} = \dfrac{T_k^j}{\sum_k T_{a_j}(r)} \forall \min(E_c(r_i)) \in 1,2,..N$)

end **for**

end **if**

end **for**

Minimal loss in the capacity utilization

$$C_{a_j} = \frac{T_k^j}{\sum_k T_{a_j}(r)} \forall \min(E_c(r_i)) \in 1, 2, ..N$$

$$C_{a_j} \cong 1$$

Performance Evaluation Analysis and Experimental Results

- The mobility model used in this work is based on the probabilistic Fraction Brownian Motion (FBM) adopted in [13], where nodes are moving, according to certain probabilities, location and time. Towards implementing such scenario, a common look-up application service for resource execution offloading is set onto each one of the mobile nodes.
- For the simulation of the proposed scenario, the varying parameters described in the previous section were exploited, by using a two-dimensional network, consisting of nodes that vary between 10-150 (i.e. terminal mobile nodes) located in measured area, as well as five cloud terminals statically located on a rack.
- All measurements were performed using WLAN (Wi-Fi based technology specifications) varying with different 802.11X specifications. During simulation the transfer durations are pre-estimated or estimated, according to the relay path between the source (node to offload resources) and

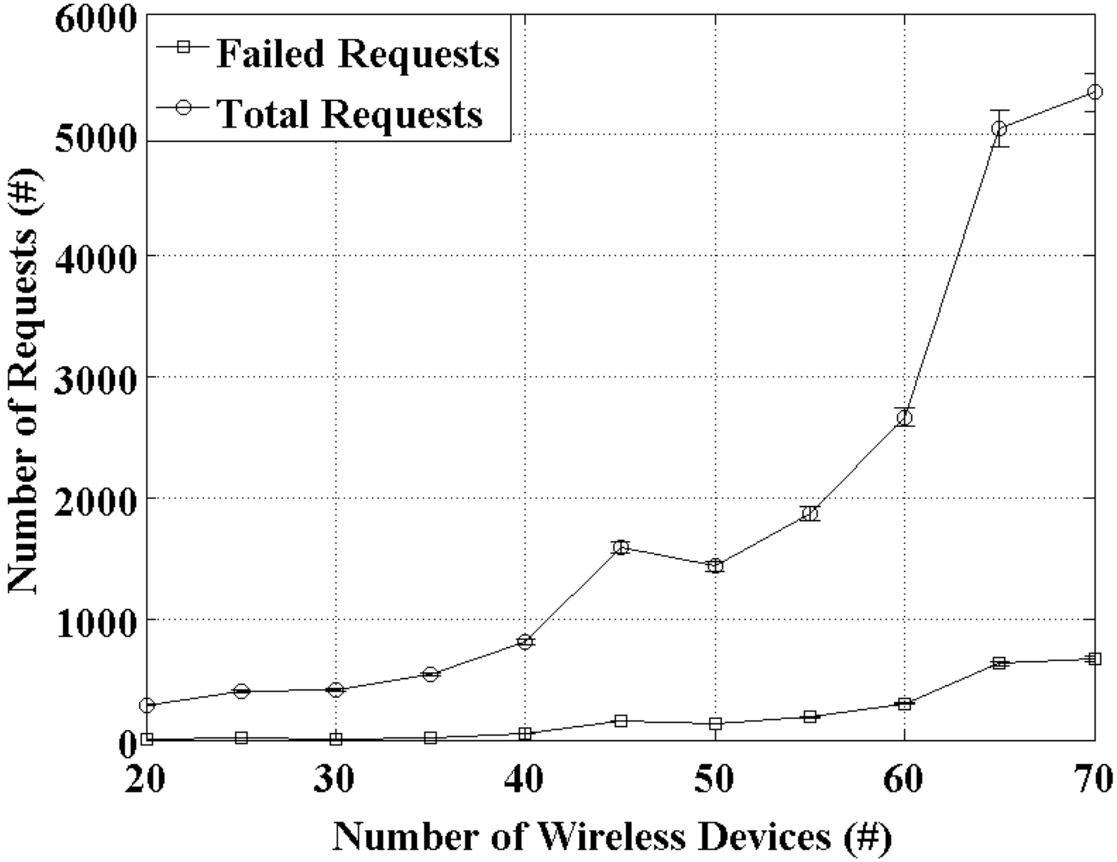
eurostars Ίδρυμα Προώθησης Έρευνας

Cloud rack terminals characteristics

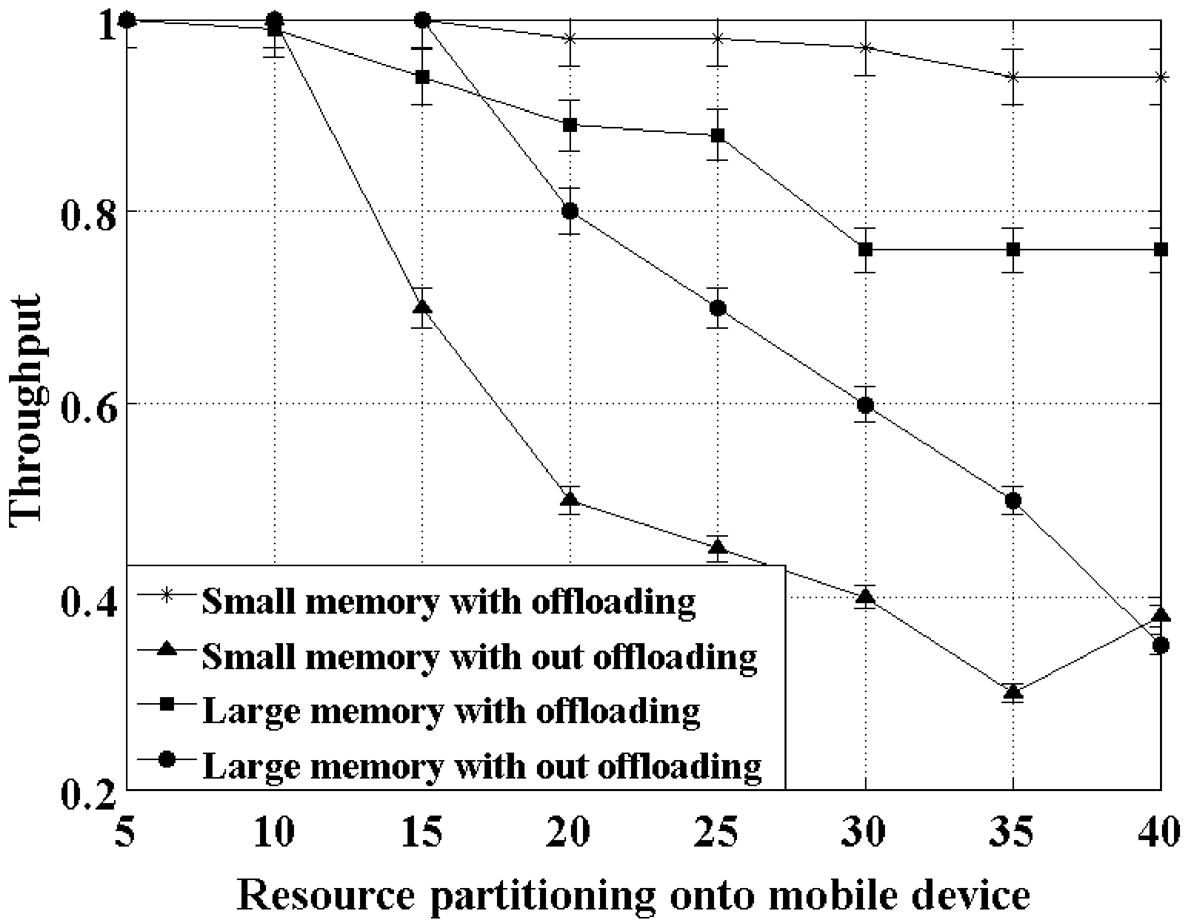
- Cloud rack terminals characteristics.

Device #	CPU (GHz)	RAM (GB)	Core No.	Hard Disk (GB)	Cache (MB)	Core Speed (GHz)	Upload Speed (Mbits/sec)
1	2.1	8	Intel Duo	600	2	5	0.6
2	2.3	16	Quad 6600	500	2	5	2.6
3	2.1	4	i5	400	2	3	2.6
4	4.0	16	i5	1000	2	5	2.6
5	2.1	32	i7	600	2	3	4.6
6	2.3	16	i5	500	2	5	2.6
7	2.1	4	Quad 6600	400	2	3	1.5
8	4.0	16	i5	1000	2	5	2.6

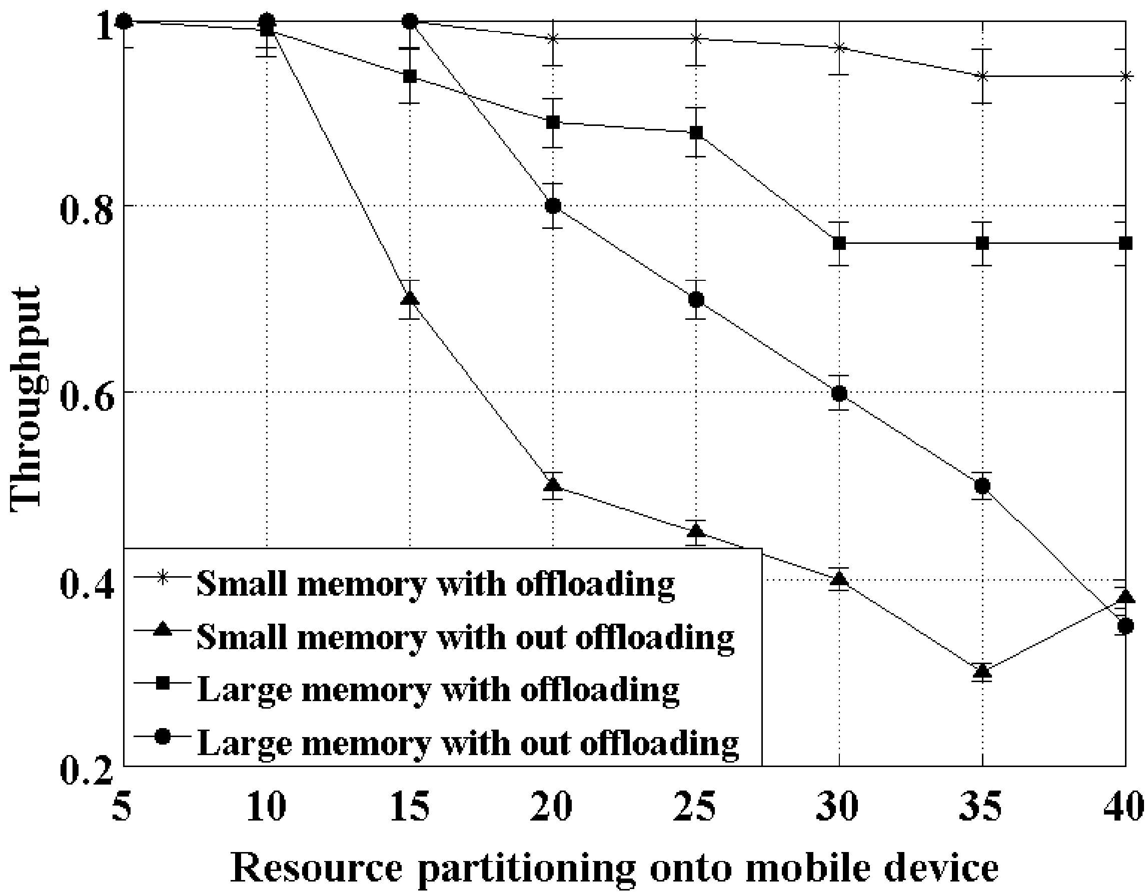
Number of requests with the number of mobile devices participating in the evaluated area.



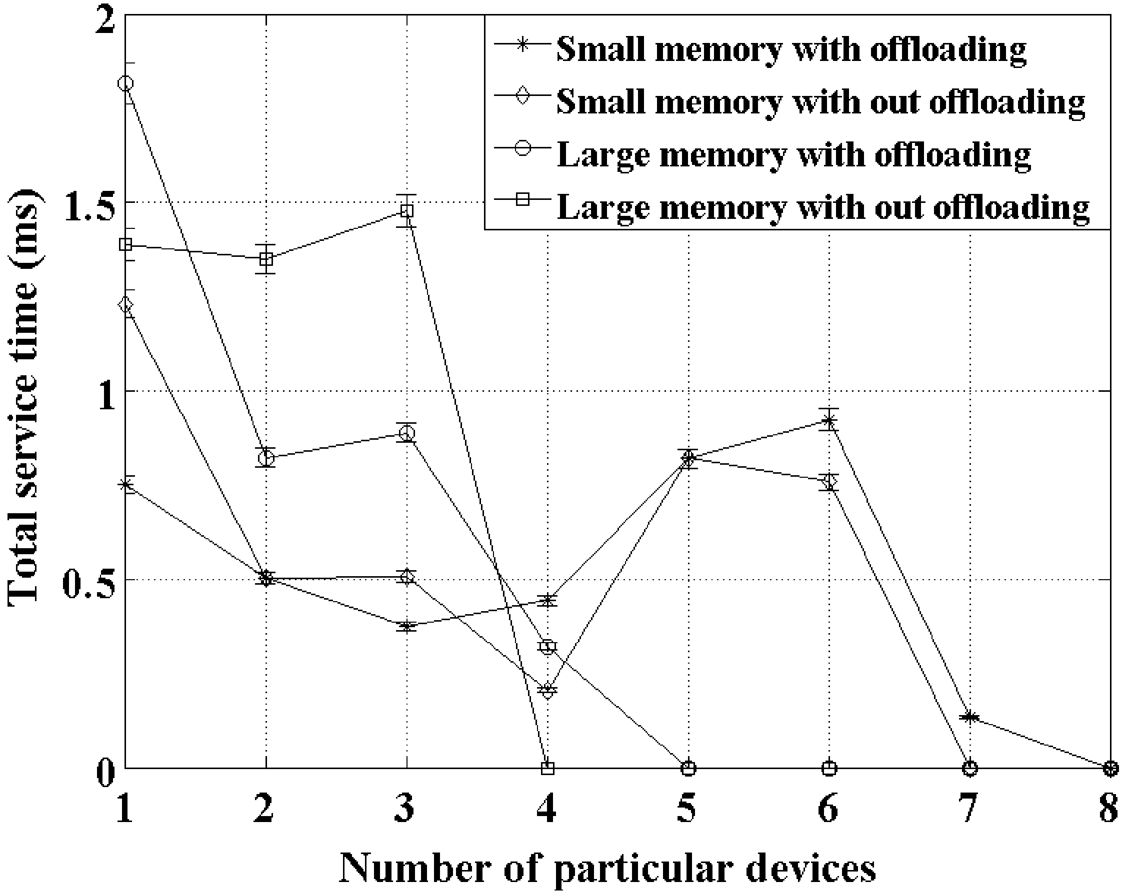
Throughput response with the mean number of executable resources that are partitioned per mobile device.



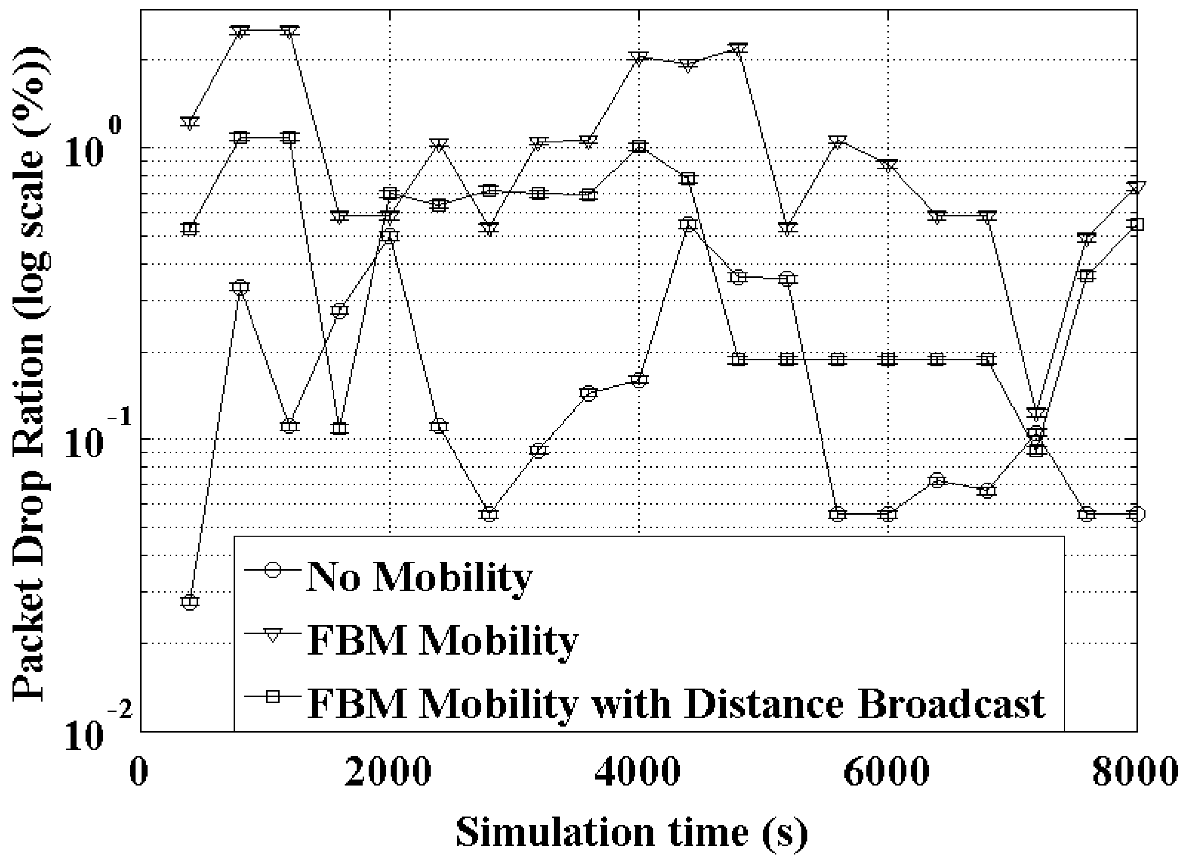
Throughput response with the mean number of executable resources that are partitioned per mobile device



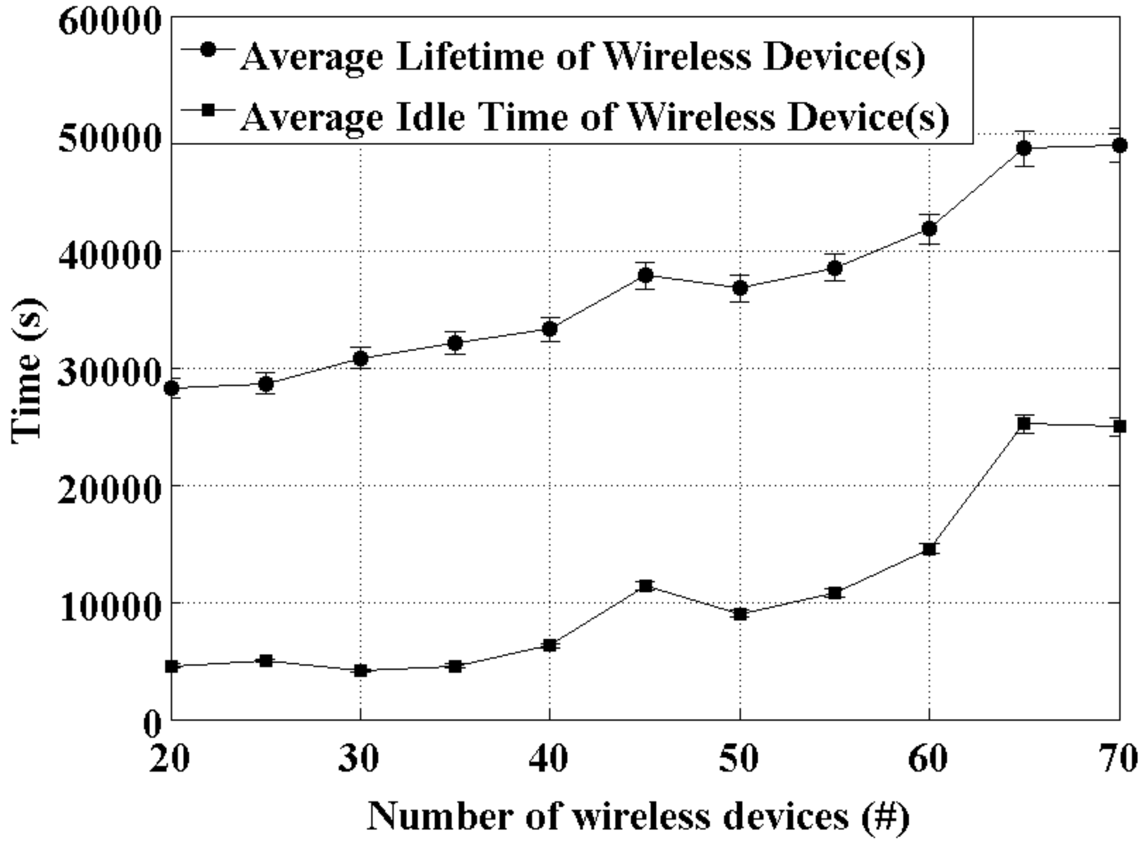
Number of requests with the number of mobile devices participating in the evaluated area



Packet drop ratio of the proposed scheme for different mobility variations and no-mobility model over time



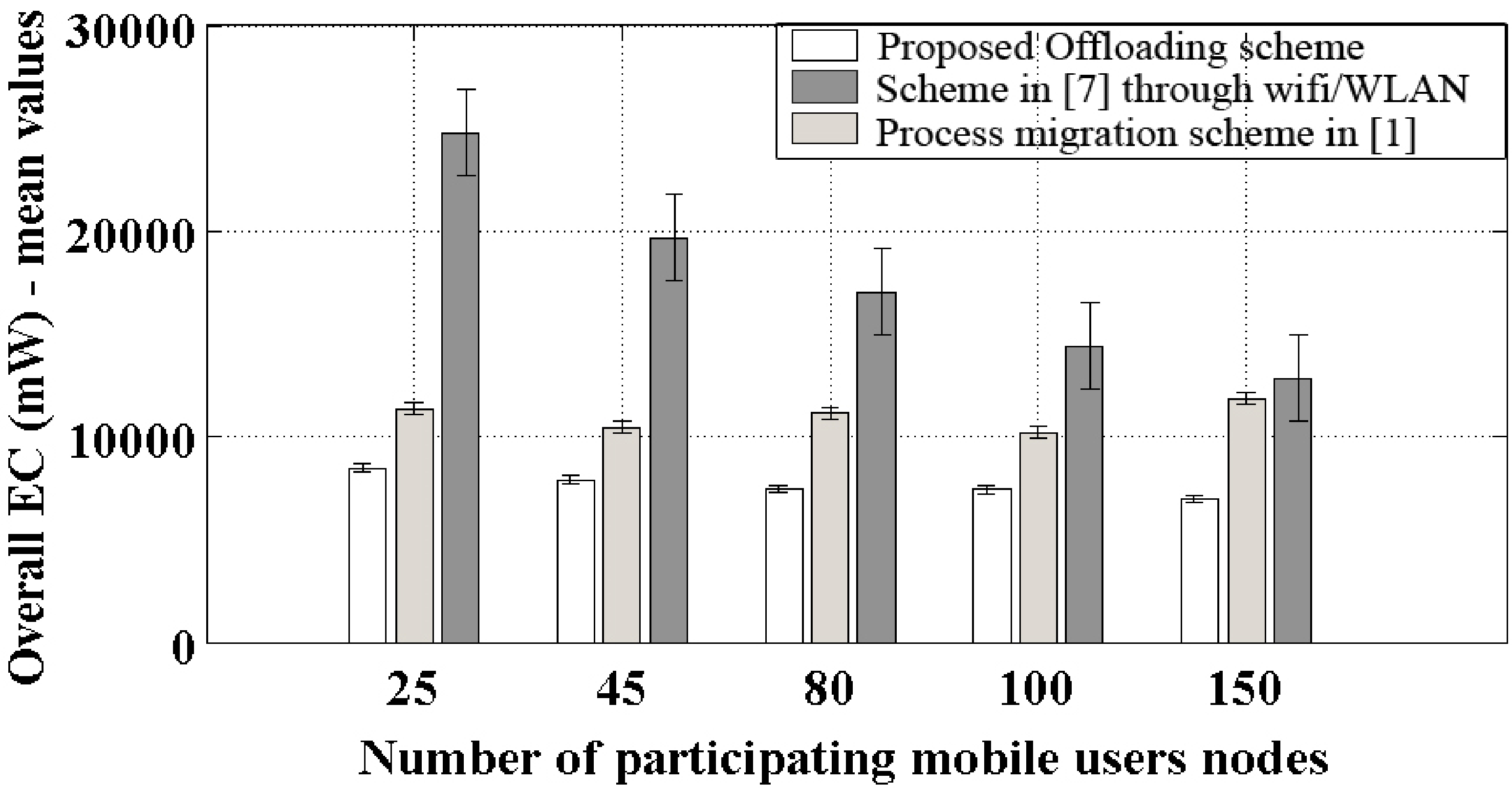
Average lifetime for both active and idle time with the number of mobile devices participating in the evaluated area.



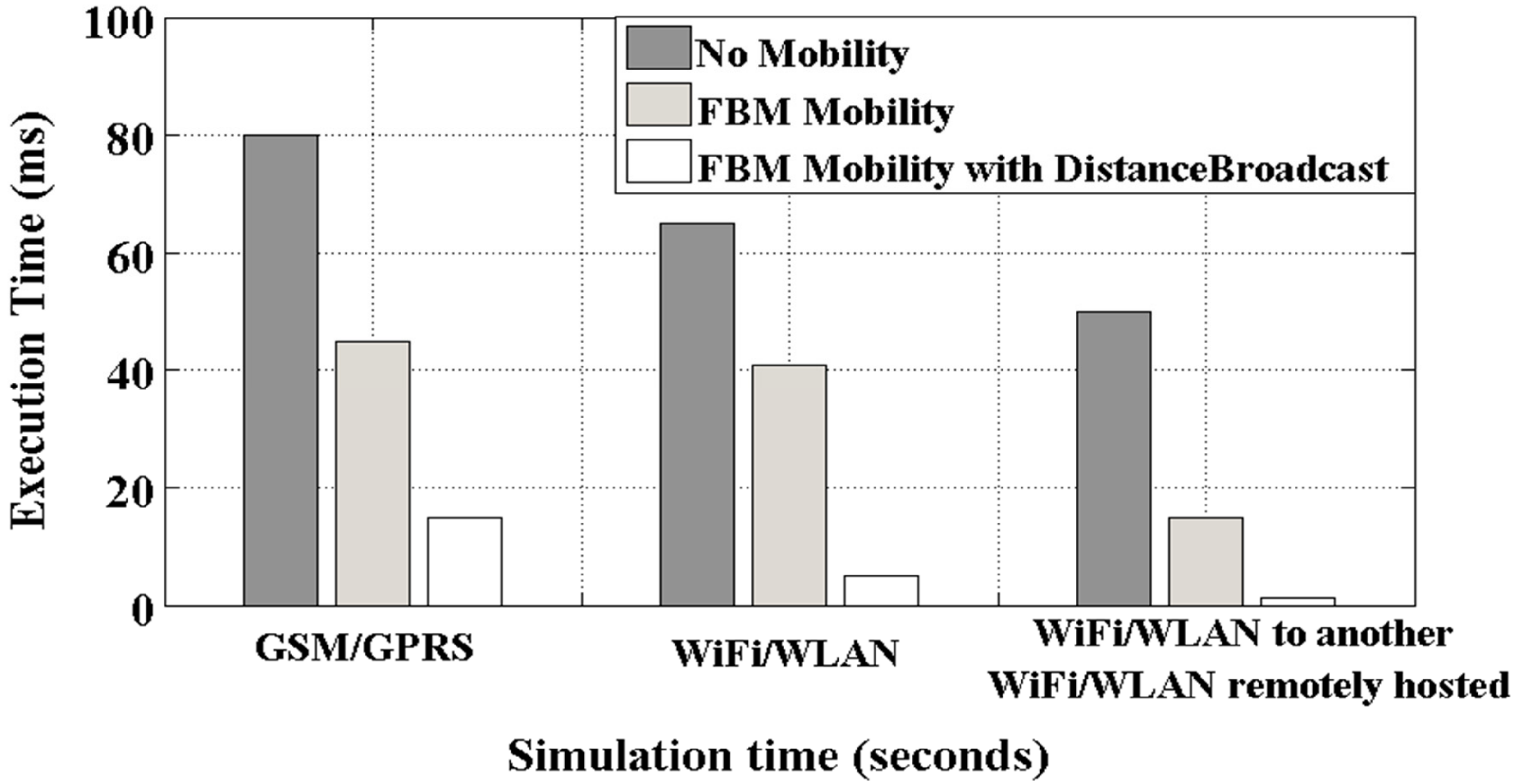
eurostars

Ίδρυμα Προώθησης Έρευνας

Overall energy consumption for each mobile device for three different schemes in the evaluated area (evaluated for the most energy draining resources)



Execution time during simulation for nodes with different mobility patterns for three different schemes of communication



Conclusions

- Computational Offload in terms of energy costs is optimized
- Enhancing the efficiency of data access/process and delay
- Using the MCC mobile users are provided with data storage and processing services (e.g. CPU speed, memory capacity etc.)

=> ***minimization of the energy consumption and the maximization of the lifetime of each mobile device based on the available resources***

- Coming streams in our work
  - Context-aware mobile cloud (Mobile HealthCare; Mobile Gaming etc.)

Eurostars is a joint programme between more than 30 EUREKA member countries and the European Union

Thank you for your patience!

Questions || Comments…

- E-mail me:

mavromoustakis.c@unic.ac.cy

EUREKA

innovation across borders

Ίδρυμα Προώθησης Έρευνας