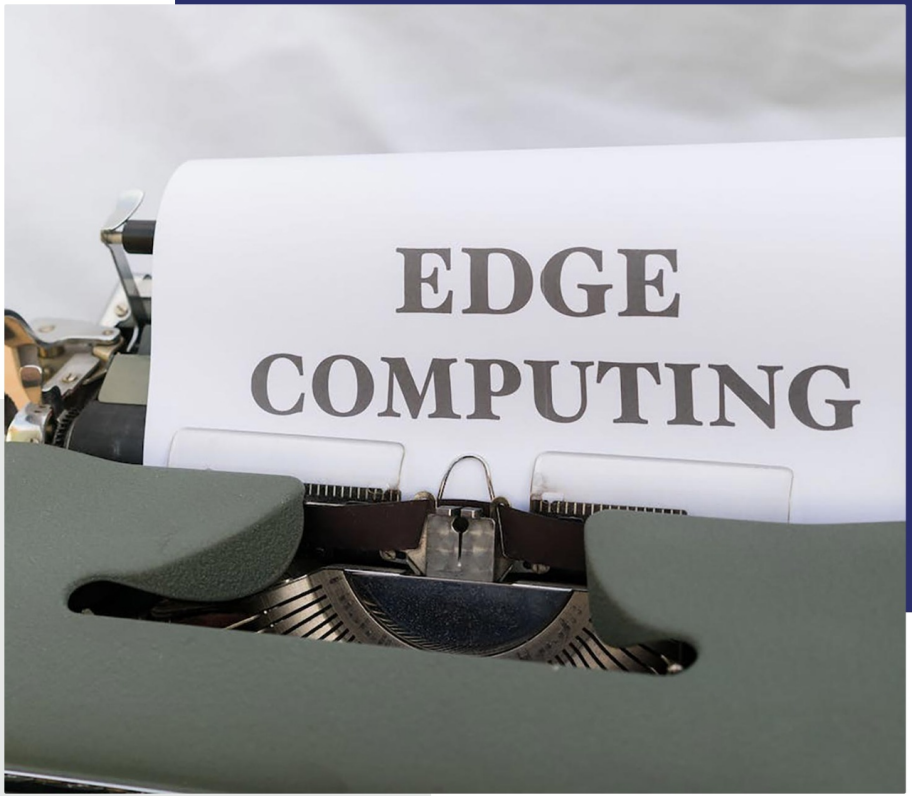


# Joint Service Placement and Request Routing in Multi-cell Mobile Edge Computing Networks

Konstantinos Poularakis, Jaime Llorca

Team:009

# The Role of Mobile Edge Computing in Reducing Latency

Exploring MEC's impact on 5G and low-latency services

# Optimizing Service Placement and Request Routing in MEC Networks

Addressing storage constraints and asymmetric communication in multi-cell environments

**Current studies focus on computing resources efficiency.**

Most research emphasizes efficient resource allocation, neglecting critical aspects of storage and communication.

**Significant storage constraints exist in MEC.**

Modern services like AR require extensive local data storage, which is often limited at Base Stations (BSs).

**Asymmetric communication needs remain unaddressed.**

Different bandwidth requirements for uplink and downlink are not thoroughly explored in existing literature.

**This project aims to fill the literature gap.**

It addresses joint optimization of service placement and request routing under multi-dimensional constraints.

# Joint Service Placement and Request Routing in Mobile Edge Computing

Exploring solutions to optimize service placement at Base Stations and request routing

- **Optimizing Service Placement**

  Determining which services to place at each Base Station (BS) to maximize limited storage efficiency.

- **Efficient Request Routing**

  Routing user requests to BSs while respecting computation and bandwidth limitations to enhance performance.

- **Joint Decision Optimization**

  Simultaneously optimizing service placement and routing to increase the number of requests served at the edge.

- **Minimizing Cloud Reliance**

  Reducing dependence on high-latency centralized cloud services by maximizing edge computing utilization.

- **Holistic Approach in Multi-cell**

  Considering the interplay of storage, computation, and communication in multi-cell environments with overlapping coverage.

# Optimizing Edge Computing with JSPRR and Randomized Rounding

A comprehensive overview of the JSPRR problem and the innovative Randomized Rounding solution for request routing.

### JSPRR Problem

The Joint Service Placement and Request Routing (JSPRR) aims to minimize centralized cloud load by maximizing edge BS service requests.

### User Request Routing

Each user request must be directed to a nearby BS or the cloud, adhering to service availability constraints.

### Resource Constraints

Constraints include BS capacity limits for storage, computation, and bandwidth for both uplink and downlink operations.

### NP-Hard Complexity

The JSPRR problem is NP-Hard, making traditional optimization methods inadequate to find optimal solutions efficiently.

### Linear Relaxation

Relaxing the integer constraints allows for a fractional solution, solvable in polynomial time using linear programming.

### Randomized Rounding

A novel approach where fractional solutions are rounded to integers, ensuring service placement and request routing are probabilistic.

# Architecture for Simulation in Multi-cell MEC Networks

Illustrating the integration of cloud resources and edge computing nodes for efficient service placement and request routing.

# Phased Project Plan for Cloud Resource Management

A detailed overview of Simulation, Evaluation, Analysis, and Reporting phases in the project

## 3

### Simulation Execution

**Person 3** will run extensive simulations based on the scenarios outlined in the paper's evaluation, testing parameters like storage, computation, and bandwidth capacities.

### Debugging Process

**Persons 1 & 2** are tasked with debugging the integrated model and troubleshooting any issues that arise during the large-scale simulations to ensure accuracy.

### Results Collection

**Person 4** will gather results from all implemented algorithms, including Randomized Rounding, Greedy, and Linear Relaxation, to facilitate comparative analysis.

### Performance Visualization

**Person 4** will create plots and charts that compare the performance of different algorithms to visualize the cloud load against capacity, reflecting the original paper's figures.

### Resource Utilization Analysis

**Person 3** will analyze resource utilization data to evaluate how the proposed algorithm manages multi-dimensional constraints effectively in cloud environments.

### Final Reporting

**Persons 1 & 2** will co-author the final project report and presentation, detailing the methodology, results, and conclusions derived from replicating the paper's findings.

# Optimizing Resource Management in MEC Networks

Exploring outcomes and conclusions of the JSPRR problem in 5G environments.

- **Simulation of the JSPRR problem**

  A successful simulation of the **Joint Service Placement and Request Routing** (JSPRR) problem within a **multi-cell Mobile Edge Computing** (MEC) environment has been achieved.

- **Performance of Randomized Rounding**

  The **Randomized Rounding algorithm** has been shown to significantly outperform the **Greedy baseline**, particularly when resource capacities are variable, highlighting its effectiveness in dynamic environments.

- **Proximity to optimal lower bound**

  The performance of the Randomized Rounding algorithm closely aligns with the **optimal lower bound** established by the **Linear Relaxation solution**, indicating high efficiency and effectiveness in resource allocation.

- **Joint optimization benefits**

  Insights reveal how **joint optimization** of storage, computation, and bandwidth resources effectively minimizes latency and reduces the load on the central cloud in MEC networks.

- **Relevance to 5G and mobile computing**

  This project addresses a crucial and complex problem at the forefront of **5G** and **mobile computing**, providing a foundational method for future enhancements in application responsiveness.

# Join us in advancing resource management strategies through Randomized Rounding.

To enhance resource management and service delivery, we propose a comprehensive implementation strategy for the Randomized Rounding algorithm, emphasizing real-world testing within Multi-cell Mobile Edge Computing networks to optimize operational effectiveness.