ADVANCECRT

A World Leading SFI Research Centre

Science Foundation Ireland For what's next

Parallel Offloading over Multi-RATs

Asad Ali, Atri Mukhopadhyay, Marco Ruffini

ABSTRACT: To minimize the service delay of future Multi-Access Edge Computing (MEC) applications, we exploit parallel offloading over multiple radio access technologies (RATs) that a mobile device comes equipped with such as Wi-Fi Direct, Wi-Fi and 5G. However, the RATs differ in response to physical conditions and performance. As a result, inappropriate marshalling of these RATs may prove counterproductive. To address this problem, we measure the performance of every RAT and optimally allocate load shares according to the performance to optimally utilize the capacity. Furthermore, to ensure smooth relay of MEC data, capacity distribution is optimized at relay nodes according to incoming load. Numerical results show significant improvement in terms of service delay.

Introduction

To support remote computation of future applications, we exploit parallel offloading over multiple Radio Access Technologies (RAT) that a mobile device comes equipped with such as Wi-Fi Direct, Wi-Fi and macro-cellular technology such as 5G. However, these technologies differ in performance and response to physical conditions. As a result, inappropriate load sharing and scheduling will lead to problems such as out-of-order packet arrival, under-utilization of capacity, lower throughput, higher end-to-end delay. To address these problems, we formulate a Continuous Non-Linear Program (CNLP) that maximizes system throughput and ensure in-order arrival of data packets. Furthermore, to ensure smooth relay of MEC data, capacity distribution is optimized at relay nodes according to incoming load.

Literature Review

Authors in [1], [2] offload data on the basis of the tasks. Distributing data on the basis of the computational tasks can lead to packet reordering delay as tasks can differ in size and computation. This leads to severe degradation in MEC services. In [3], the RATs are not used simultaneously, rather the choice is made for best radio-edge pair.

System Model and Problem Formulation

Assumed system model is shown in Figure 1 below, where a UE is connected to a peer node, a Wi-Fi Base-station and 5G Base-station. The data in wireless networks is governed mainly by four different types of delays namely queuing delay, slot synchronization delay, transmission delay and propagation delay. Assuming m total links from source to destination and merging the four types of delays, the end-to-end delay of a single RAT r is formulated as follows [4].

$$D_r = \sum_{u=1}^{m}\left(\frac{1}{2\mu\zeta_u} + \frac{1}{\mu\zeta_u - \lambda_u}\right) + \sum_{u=1}^{m}\theta_u \qquad (1)$$

Where μ is packet size, ζ_u and λ_u are the capacity and load on the link u respectively and θ_u is the propagation constant. θ_u does not depend on the load. Therefore, we measure and add its value at the end of the computation. We formulate our continuous non-linear program as follows.

$$Minimize\ D_r$$

Subject to

$$D_r = D_t = D_v \qquad (2)$$

Where D_r, D_t, D_v are end-to-end delay of RAT r, t and v respectively. The constraint in (2) means that all RATs must have equal delay. With equal delay, simultaneous arrival of packets is ensured to avoid any packet reordering delay.

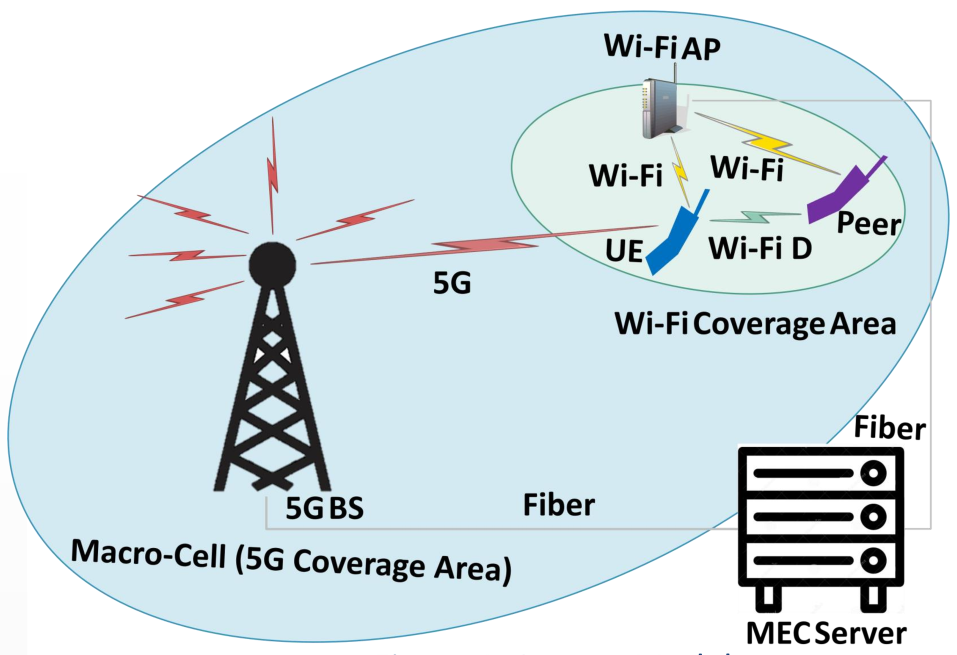

Figure 1: System Model

Results

For different loads, the delay is same. The proposed scheme has lesser delay than existing schemes.

Conclusion

We formulated a CNLP to optimize load shares and traffic scheduling. We also optimized capacity distribution at relay nodes. We used Lagrange Multiplier Theorem to solve our programs.

Proposed Solution

We used Lagrange Multiplier Theorem to get optimal load shares according to the performance measures of the RATs as follows. Taking partial derivative of D_r with respect to λ_u and equaling to 0, we get.

$$\frac{2}{\mu\zeta_r - \lambda_r} = -\frac{1}{\mu\zeta_r}\left(\frac{-\frac{2}{\zeta_r} + \frac{1}{\zeta_t} + \frac{1}{\zeta_v}}{2\mu} + \frac{1}{\mu\zeta_t - \lambda_t} + \frac{1}{\mu\zeta_v - \lambda + \lambda_r + \lambda_t}\right)$$

We obtain load share λ_r for RAT r by solving equation above for λ_r. Next, by the constraint if $D_t = D_v$, putting their respective values from (1) and solving for λ_t, we get;

$$\lambda_t = \frac{\mu\zeta_t\left(\lambda_r(3\zeta_r - \zeta_t) - 3\mu\zeta_r(\zeta_r - \zeta_t)\right)}{\lambda_r(\zeta_r - \zeta_t) - \mu\zeta_r(\zeta_r - 3\zeta_t)}$$

Finally, we get λ_v by subtracting λ_t and λ_r from total load λ, that is $\lambda_v = \lambda - \lambda_r - \lambda_t$.

For relay nodes, to ensure smooth relay of MEC data, we optimally distribute the capacity according to the incoming data. At relay node, if k nodes are connected to it, total capacity will be distributed among these k nodes. Mathematically, we can write.

$$\zeta_t = \sum_{u=1}^{k}\zeta_u \qquad (3)$$

Also, system delay is given by [4];

$$\Gamma = \frac{1}{T}\sum_{u=1}^{k}\frac{\lambda_u}{\mu\zeta_u - \lambda_u}$$

We again use Lagrange multiplier theorem and re-write our capacity optimization problem as follows

$$W = \frac{1}{T}\sum_{u=1}^{k}\frac{\lambda_u}{\mu\zeta_u - \lambda_u} - K\left(\sum_{u=1}^{k}\zeta_u - \zeta_t\right)$$

Here K is the Lagrange multiplier and $\left(\sum_{u=1}^{k}\zeta_u - \zeta_t\right)$ is the capacity conservation constraint as shown in (3) . Taking $\frac{\partial W}{\partial \zeta_u}$ and equalling to 0, we get;

$$\zeta_u = \frac{\lambda_u}{\mu} + \frac{\left(\zeta_t - \sum_{u=1}^{k}\frac{\lambda_u}{\mu}\right)\cdot\sqrt{\lambda_u}}{\sum_{k=1}^{n}\sqrt{\lambda_u}}$$

ζ_u above is the optimal capacity that the relay node will assign link u according to the incoming load λ_u.

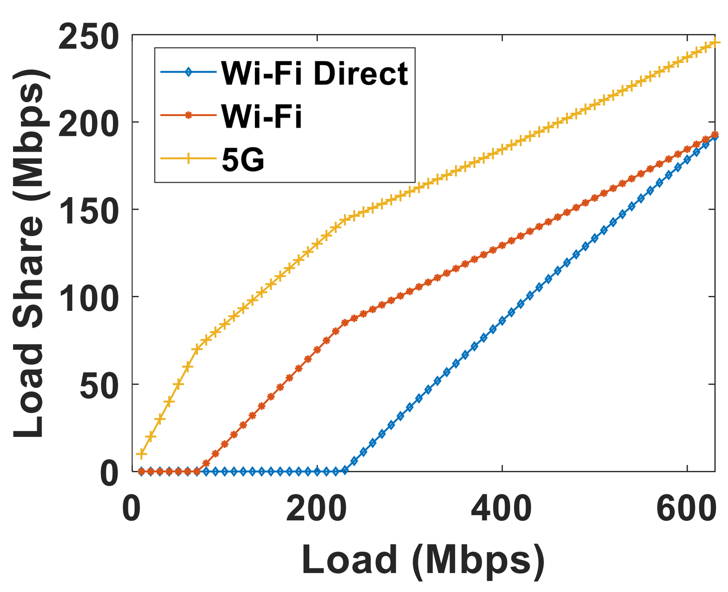

Figure 2: Incoming Load Vs. Load Shares

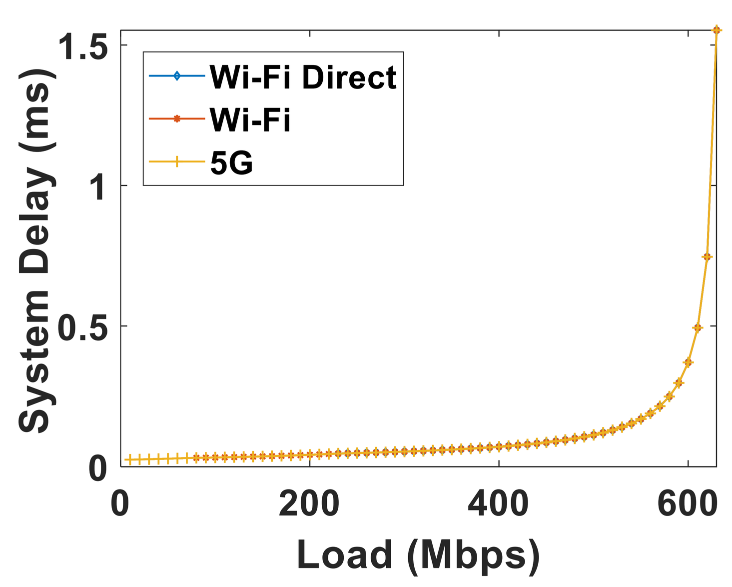

Figure 3: Same Delay for Different Loads

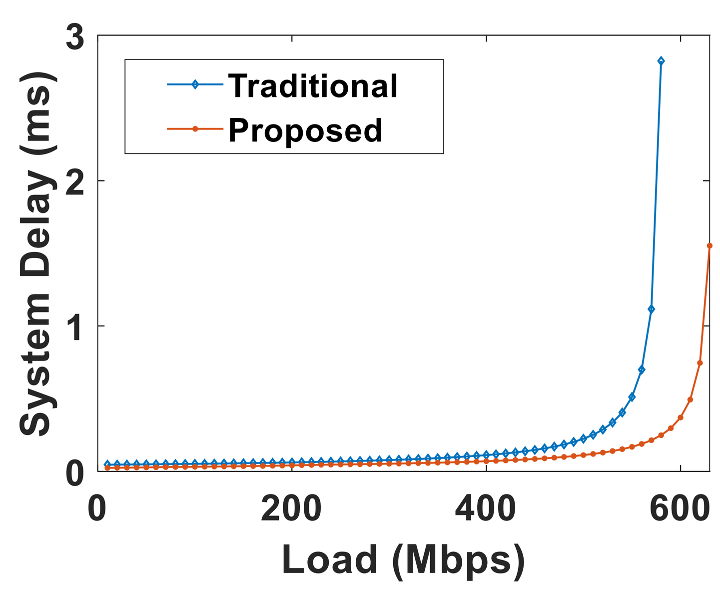

Figure 4: Lesser Delay Than Task-Based Distribution Schemes

HOST INSTITUTION

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

PARTNER INSTITUTIONS

References

[1] H. Lin, et. al, “A survey on computation offloading modeling for edge computing,” Journal of Network and Computer Applications, vol. 169, pp. 1–25, 2020.
[1] T. Braud, et. al., “Multipath computation offloading for mobile augmented reality,” in 2020 IEEE PerCom, 2020, pp. 1–10.
[2] . Jing, Q. et. al“Long-term maxmin fairness guarantee mechanism for integrated multi-rat and MEC networks,” IEEE TVT, pp. 1–15, 2021
[3] . C. Lin, et. al, “Communication and computation offloading for multi-rat mobile edge computing,” IEEE Wireless Comm. vol. 26, no. 6, pp. 180–186, 2019.
[4] A. Ali, et. al. “Multi-RAT Parallel Offloading in MEC: Optimizing Load Shares, Scheduling and Capacity”, Submitted to IEEE IoT Journal, 2021.