Joint Task Offloading and Routing in Wireless Multi-hop Networks Using Biased Backpressure Algorithm

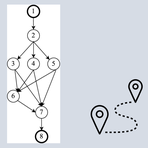
Zhongyuan Zhao, Jake Perazzone†, Gunjan Verma†, Kevin Chan†, Ananthram Swami †, Santiago Segarra**

*Rice University, USA † US Army's DEVCOM Army Research Laboratory, USA

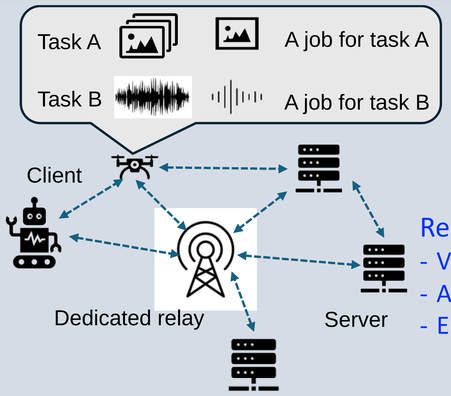
2025 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING
April 06 - 11, 2025 Hyderabad, India

Recurrent task offloading for ad-hoc edge computing

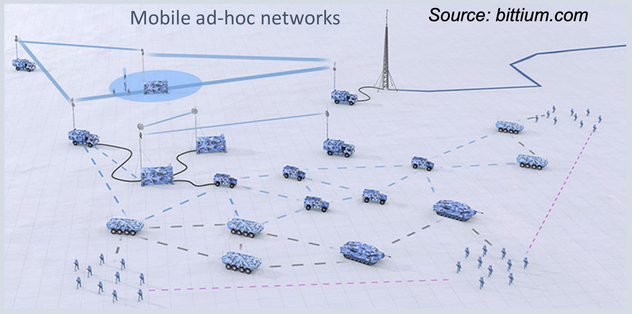
Paper ID: 2496



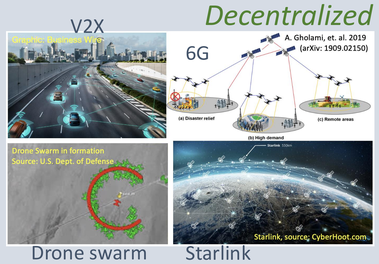
One-time on-demand computation offloading
- Path finding
- Scientific computing
- Data analysis



Recurrent task (our focus)
- Video surveillance & analytics
- Autonomous robots
- Environmental monitoring



Infrastructureless
Self-organizing
Decentralized



Existing solutions: joint v.s. separated decisions

[Joint_LP] Joint decisions with linear programming

- Requires a centralized scheduler
- Offline computing, joint decisions in batches
- Linear relaxation of mixed-integer programming
- Model task as liquid flows
- Require linear cost function
- Ignore interference in wireless networks

[SPBP_SP-BP] Ideal distributed offloading & routing

- Mobile devices know immediate Queueing States of all servers
- Online, asynchronized decision making
- State-of-the-art decision mechanism (SP-BP, [Zhao 2024])
- Two-step decision: first offloading destination, then the route to it.

Existing approach: destination of a job cannot be changed once it is sent out! (But ours do)

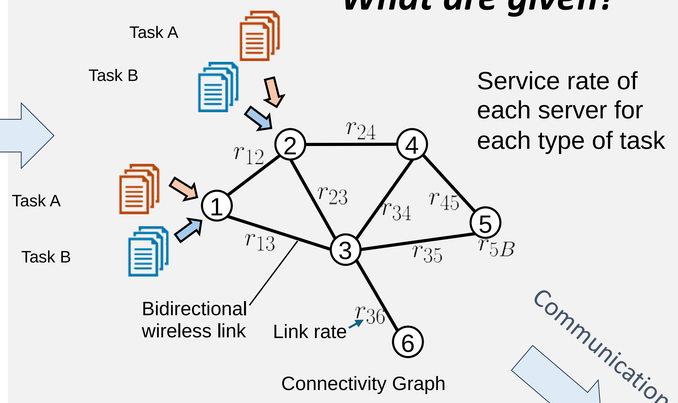
References

- Z. Zhao, B. Radojičić, G. Verma, A. Swami and S. Segarra, "Biased Backpressure Routing Using Link Features and Graph Neural Networks," in IEEE TMLCN, vol. 2, pp. 1424-1439, 2024
- Z. Zhao, J. Perazzone, G. Verma, and S. Segarra, "Congestion-aware distributed task offloading in wireless multi-hop networks using graph neural networks," in IEEE ICASSP, pp. 8951–8955, 2024.
- G. Feng, X. Li, Z. Gao, C. Wang, H. Lv, and Q. Zhao, "Multi-path and multi-hop task offloading in mobile ad hoc networks," IEEE Trans. Vehicular Tech., vol. 70, no. 6, pp. 5347–5361, 2021.
- R. Lin, Z. Zhou, S. Luo, Y. Xiao, X. Wang, S. Wang, and M. Zukerman, "Distributed optimization for computation offloading in edge computing," IEEE Trans. Wireless Commun., vol. 19, no. 12, pp. 8179–8194, 2020.

Offloading without destination, routing without routes

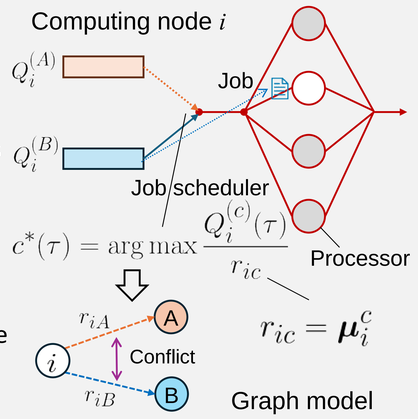
Job arrival rate of each task initialized by each client

What are given?



Service rate of each server for each type of task

Job scheduling operation
Whenever there is free processor on a server, send it a job from the queue requires the longest time to clear



Virtual link modeling

Conflict: virtual links on the same server conflicting with each other

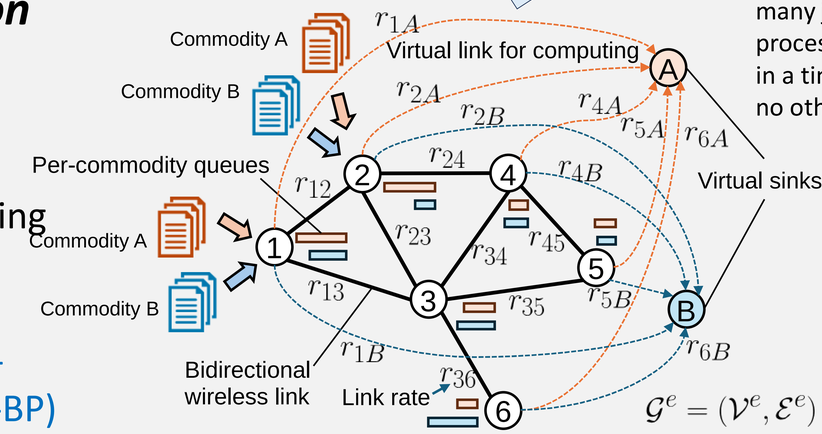
Graph model

Communications

Computation

A simple 3-step solution

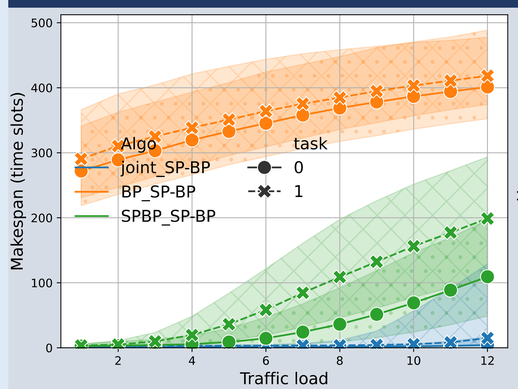
1. Model computing as sending a job to a virtual sink over a virtual ink
2. Joint offloading and routing → packet routing
3. Pick the best routing algorithm, shortest path-biased Backpressure (SP-BP)

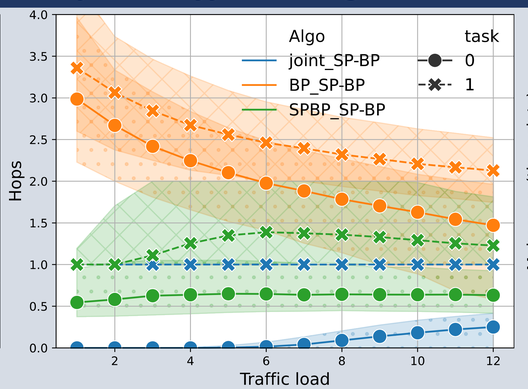


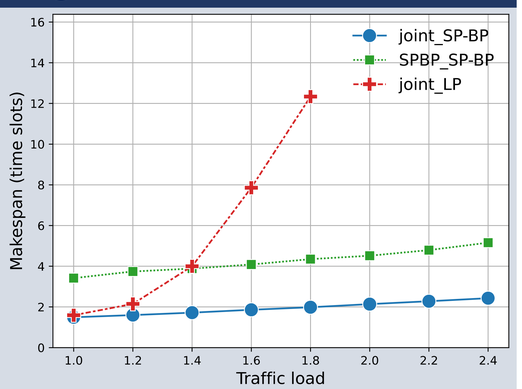
Virtual link rate: how many jobs can be processed by the server in a time slot, assuming no other types of jobs

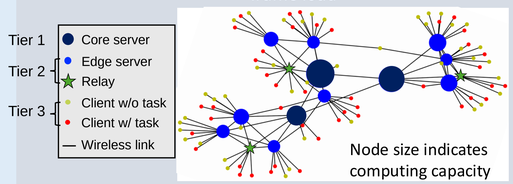
Paper, preprint, code, slides, poster

Numerical Results: *distributed joint offloading and routing beats all*









[BP_SP-BP]: only consider queueing states, offload to remote servers (worst)
[SPBP_SP-BP]: queueing states, distance, and capacity of servers (a lot better)
Ours [Joint_SP-BP]: decide the next step of a job on the fly (best, leading everyone by a large margin) **we don't predict the future, we adapt on-the-fly**
[Joint_LP]: make joint decisions offline, unable to adapt and scale to heavier loads