Bandit Learning-based User Clustering and User Selection for Cellular Networks

Isfar Tariq*, Kartik Patel*, Gustavo de Veciana*, Sanjay Shakkottai*
Collaborators: Salam Akoum, Thomas Novlan, Milap Majmundar, *AT&T Labs*, *Austin*
*6G@UT · Wireless Networking and Communications Group · University of Texas at Austin

Motivation: Optimal MU rate scheduling for a given CQI

Question: Given an observed Channel Quality Index (CQI), what is the optimal subset of users to schedule and the associated transmission rates, i.e. Modulation and Coding Scheme (MCS) indices?
In practice, a semi-static map from CQI to MCS is used but non-adaptivity to specific scenario results in performance loss.
Prior work: Adaptively learn the optimal mapping from CQI to MCS [2]! However, the optimal MCS depends on the choice of scheduled users. Hence, learning is too complex for large number of users.

Learning the mapping with user clustering [1]

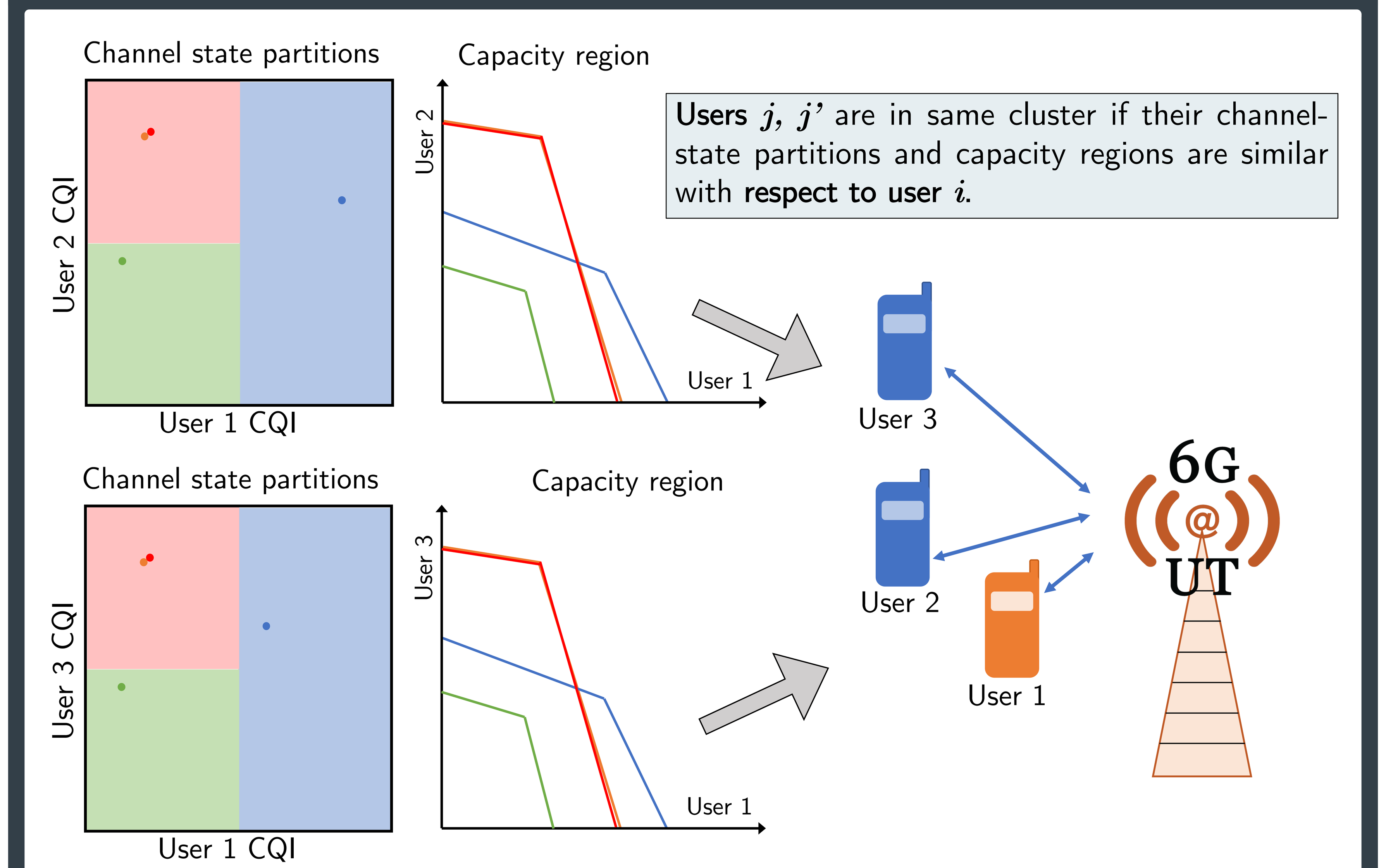


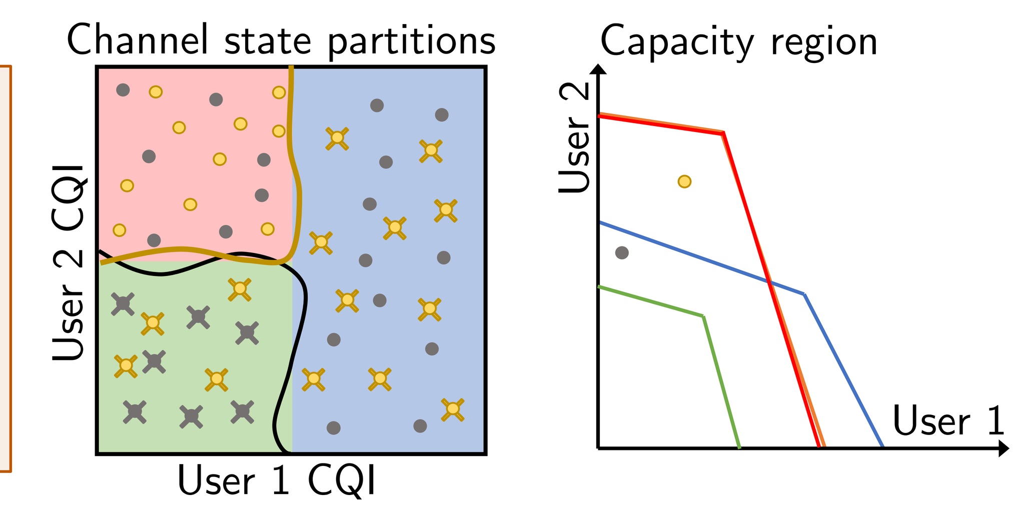
Figure: User clustering based on similarity in channel-state partitions and capacity regions

Benefits of user clustering

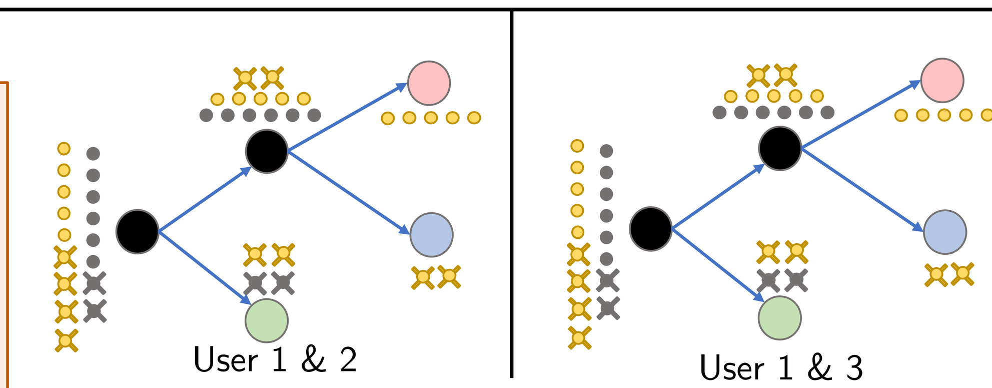
- Reduction in the exploration complexity, thus efficient learning
  - Get preliminary estimate of channel partitions and associated rate regions of user-pairs to create user-clusters.
  - Then estimate and fine-tune the channel partitions and rate regions for each user-clusters.
- **Robust in dynamic user setting** - For new user, the system only needs to assign it into a cluster. Then learned knowledge about cluster can be applied to new user.

Joint user-channel clustering and rate scheduling algorithm

Step 1: For each user-pairs Initializing channel-state classifiers
- Schedule same MCS for different CQI.
- Train binary classifiers based on transmission feedback

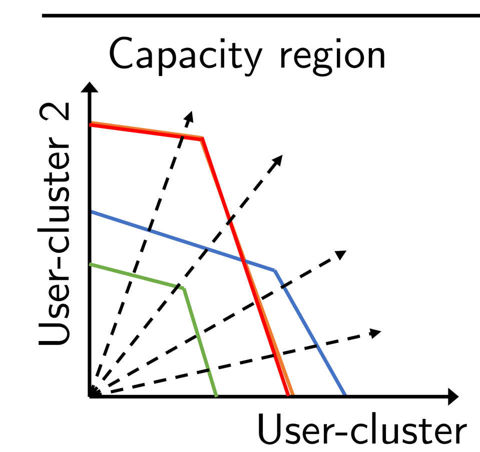


Step 2: User-clustering
- Compare results of binary classification tree of user-pairs (i,j) and (i,j')
- Define channel-classifiers for user-cluster pairs.



For each epoch

Step 3: Class explore (Same as Step 1 but for user-clusters)
- Test one MCS for different channel-states
- Update channel-state classifier tree for each cluster-pairs
- Update the user clusters if required.



Step 4: Capacity explore phase
- For a user cluster-pair and a given direction, estimate optimal rate using binary search.

Step 5: Exploit
- Schedule best rate for a given user-pairs and direction

Theoretical guarantee

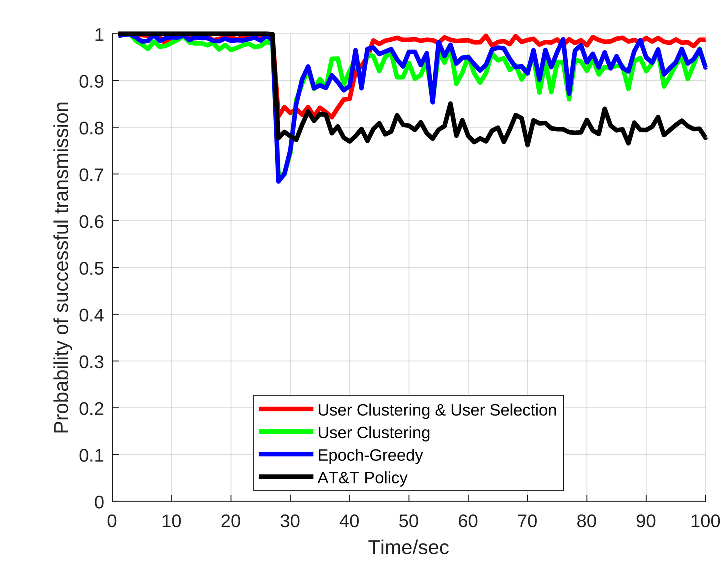
Regret: Difference in the throughput achieved by omniscient genie policy and by the proposed algorithm.
Theorem 1: *Given certain assumptions, with probability at least* $1 - O(KDL^2\delta)$, *the proposed algorithm achieves a regret bound of,*

$$R(T) = \mathcal{O}\left(L^{2/3}T^{2/3}\log\left(\frac{1}{\delta}\right)\left(D\log T + K + \sqrt{V}\right)\right).$$

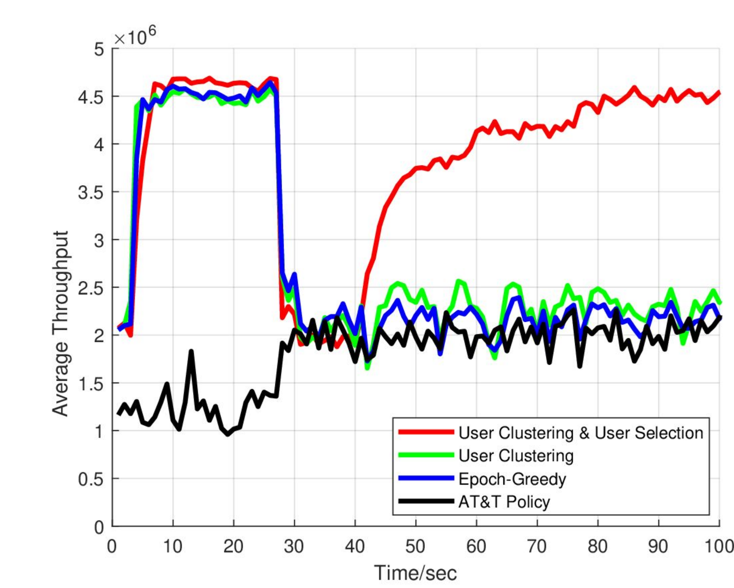
User selection algorithm

- Prioritize the user-cluster pairs which are under-explored or have high throughput.
- Penalize the user-cluster pairs which are well-explored and have lower throughput estimates.

Results



Probability of success



Overall throughput

Figure: Performance of our algorithm (in red) in a setting with dynamic users: (1) compared to benchmark static policy (in black), learning by the proposed algorithm achieves higher throughput. (2) The benefit of user clustering and user selection algorithm can be observed by comparison with epoch-greedy strategy [2] (in blue) and only user-clustering algorithm (in green).

References

[1] Tariq et al., "*Bandit learning-based User Clustering and User Selection for Cellular Network,*" *Under preparation.*

[2] Tariq et al., "*Auto-tuning for Cellular Scheduling through Bandit-Learning and Low-Dimensional Clustering,*" IEEE/ACM Trans. Netw. Vol. 29, 2021.