Evaluating Restless Multi-Armed Bandit Solutions to Resource Allocation Problems in Public Health

Aditya Mate[1,2], Aparna Taneja[1], Milind Tambe[1]

[1] Google Research, [2] Harvard University

EAAMO2022

Harvard John A. Paulson School of Engineering and Applied Sciences

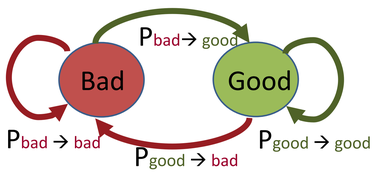
Google Research

❑ MOTIVATION

- Common low-resource public health challenge: Optimization of limited monitoring resources
- **Example**: mMitra program by ARMMAN in India
  - Millions of enrolled mothers
  - Severly limited health workers
  - Service calls boosts engagement
  - Whom to select for service call?

- Popular solution framework:
  - Restless Multi-Armed Bandit

❑ RESTLESS BANDITS MODEL

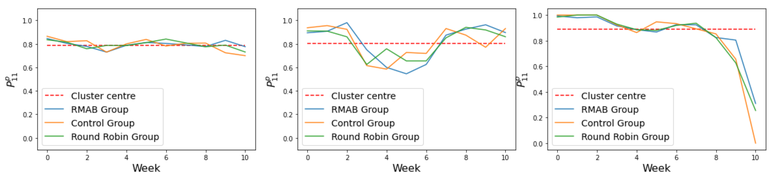


2 states

2 actions

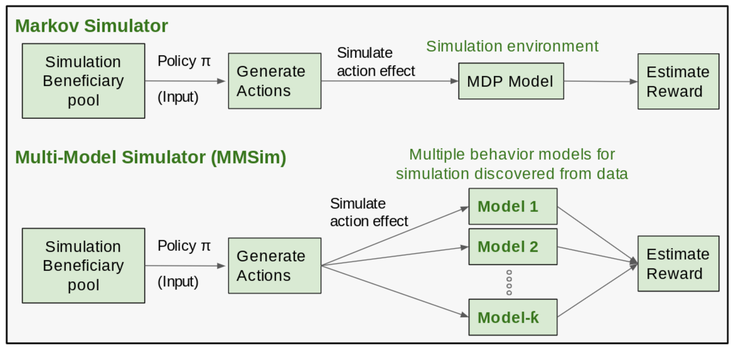
Transition matrix

- Each beneficiary (arm) is a 2-state 2-action MDP
- Transition probabilities of arm are stationary (Markov assumption)
- RMAB consists of N arms; planner can pull k (k<N)
- Real data (~23,000 real beneficiaries) does not conform to Markov model well



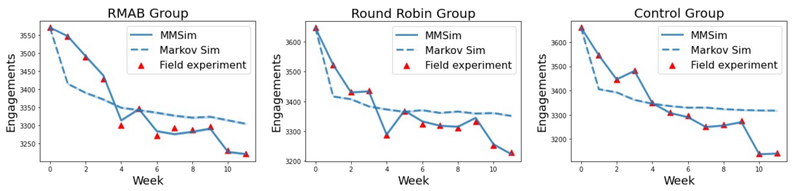
❑ MMSim: PIPELINE OVERVIEW

- Goal is to build a more accurate RMAB simulator
- Identify unique, richer behavior patterns for agents from data and use as model to simulate

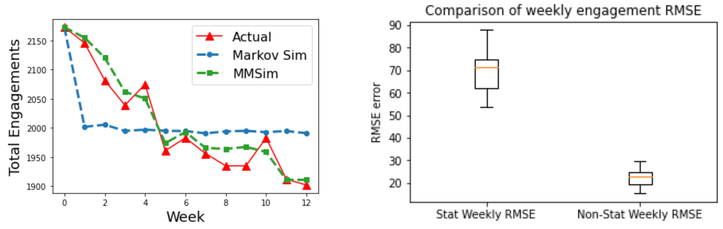


❑ EVALUATION of MMSim

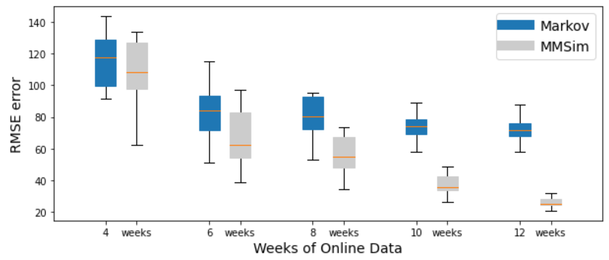
- Even with perfect data, Markov simulator is not expressive enough; only crudely fits actual data
- MMSim build richer models, hence is more expressive and fits data closely



- Apportioning Improved Quality to Simulator

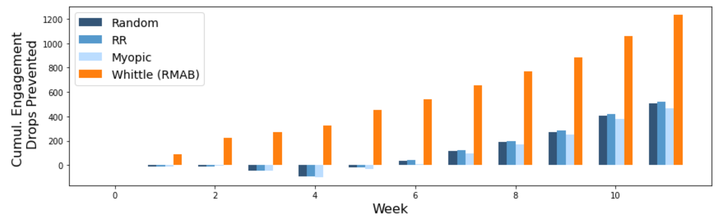


- Online Learning Setup

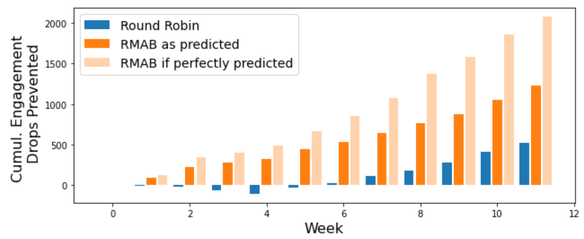


❑ UTILITY OF MMSim

- Evaluating new policies in simulation



- Evaluating performance loss from inaccurate predictive model



.

❑ CONCLUSION

- Markov assumption of RMAB model may not represent real data well
- Transition probabilities could be non-stationary
- Building simulators capturing non-stationary probabilities for evaluating RMABs can be valuable

Contact: Aditya Mate aditya_mate@g.harvard.edu | **Website**: https://projects.iq.harvard.edu/adityamate/home