

Learning Index Policies for Restless Bandits with Application to Maternal Healthcare

Arpita Biswas, Gaurav Aggarwal, Pradeep Varakantham, Milind Tambe

Google Research

Motivation

Maternal Healthcare

Pregnancy
Childbirth
Postnatal



ARMMAN's mMitra program
(non-profit organization)

Problem Statement

Goal

Learning intervention scheme under budget constraints

Post-intervention improvements in engagements are uncertain.

The engagement dynamics of the beneficiaries are unknown.

Uncertainty

Human behavior

Cost of Intervention

Sequential decision

No. of healthcare workers much less than the no. of beneficiaries.

Decisions taken at a timestep impacts the behavior of all the beneficiaries in the next timesteps.

Beneficiary's Engagement (listenership)



Call records
(preventive care information)



Healthcare worker visiting patients
(interventions)

Timely interventions → Increase in engagement

Model: Restless Multi-Armed Bandits

$$\max_{\pi} \liminf_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\sum_{i \in N} \sum_{h=0}^{t-1} R_i^{X_i(h)}(A_i^{\pi}(h)) \right]$$

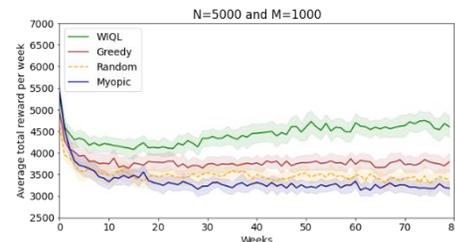
$$\text{s.t.} \quad \sum_{i \in N} A_i^{\pi}(t) \leq M \quad \text{for all } t = \{1, 2, \dots\}$$

$X_i(h) :=$ State of arm i at timestep h .
 $A_i^{\pi}(h) :=$ Intervention using policy π .
 $R_i^{X_i(h)}(A_i^{\pi}(h)) :=$ Reward observed.

Empirical Analysis

Intervention Schemes

- **WIQL**
- **Greedy**
- **Random**
- **Myopic**



Conclusion: WIQL outperforms existing methods.