Real-Time Personalization of Mobile Interventions

Peng Liao[1], Kristjan Greenewald[3], Predrag Klasnja[1] and Susan Murphy[2]

University of Michigan[1], Harvard University[2], IBM Research[3]

HeartSteps Study

- HeartSteps (HS) is a mobile health study aiming to support maintenance of physical activity after cardiac patients finish cardiac rehabilitation.
- HS 1.0: 37 participants, 42 days study. **HS 2.0 under planning:** 80 participants, 90 days study.
- One of the intervention component in HS 2.0 is the tailored activity message.
- **Goal**: at each decision time, determine whether or to send the message if available based on user's context and history, with the goal to maximize the total proximal outcomes.
- **Reinforcement Learning (RL) Framework**

a) Decision Time: 5 times per day, 90 days

b) Action (treatment): send vs do nothing

c) Reward: number of steps taken in the 30 minutes after decision time

Challenges to RL

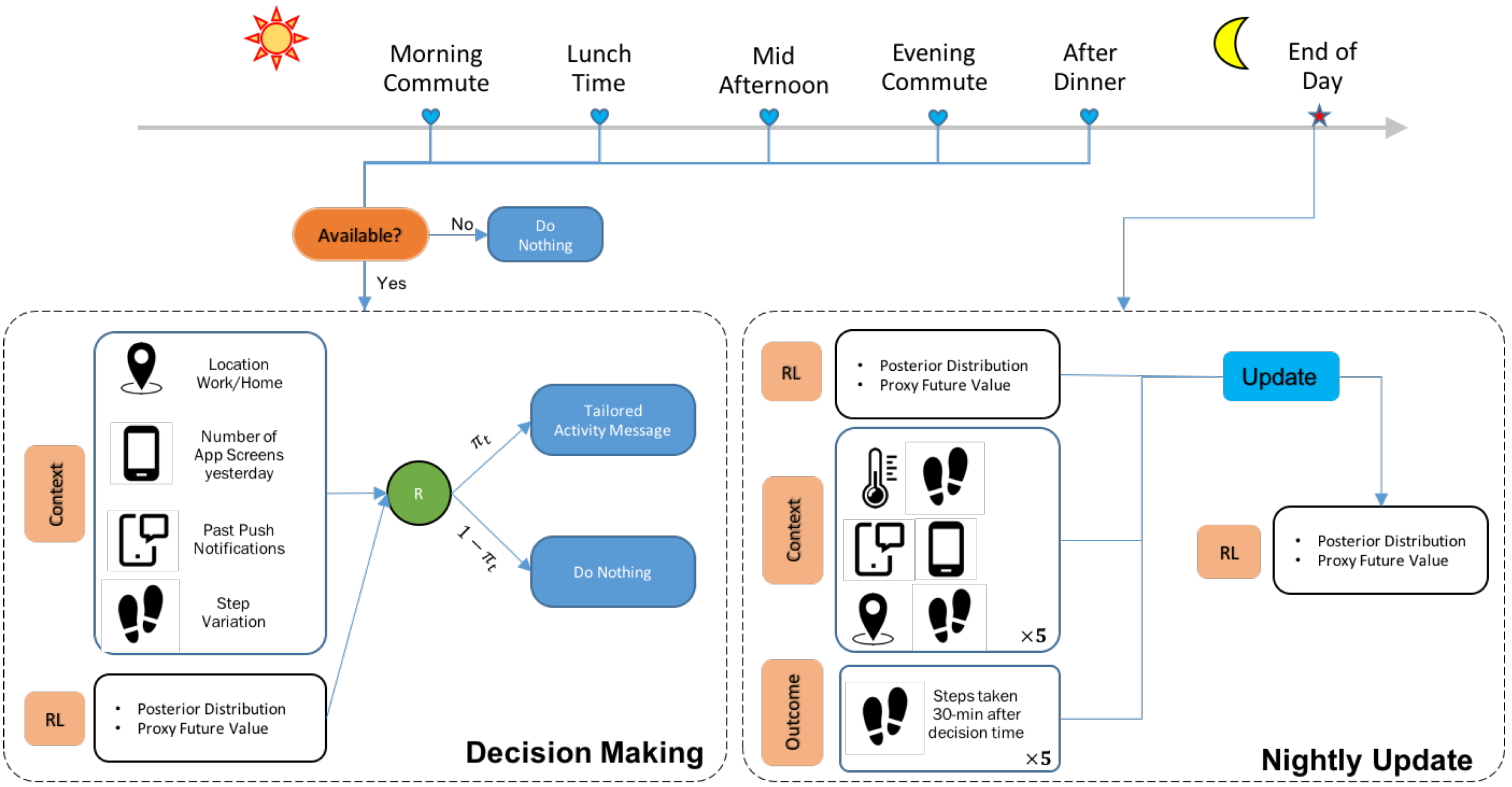
1) Highly noisy (tracker) data
2) Complex and non-stationary reward (due to unobserved state)
3) Treatment tends to have positive effects on immediate rewards, but likely negative impact on future rewards via user habituation/burden.
4) Need to ensure the stability of off-policy learning and the causal analysis after the study

Contextual Bandit with Proxy Value

- *Linear Thompson Sampling Bandit:*
  - Basic Idea: linear parameterize the reward; require prior distribution; choose action with the posterior probability of being optimal
  - Why? Bandit learns faster than full RL. The use of prior (built from HS 1.0) reduces the impact of noise (**Challenge 1**). TS is a stochastic algorithm, allowing for the causal analysis and off-policy learning (**Challenge 4**)
- *Action-centering*
  - Basic Idea: use the hierarchical linear model with the centered action by randomization probability
  - Why? Robust to the model misspecification of baseline reward (**Challenge 2**)
- *Gaussian Process Prior*
  - Basic Idea: construct a Gaussian process prior (over time) for the parameters in treatment effect model
  - Why? Protect against the non-stationarity in the reward and ensure continuing exploration (**Challenge 2 and 3**)
- *Probability Clipping*
  - Restrict the probability of sending treatment within [0.1, 0.8]
  - Why? Ensure the stability of data analysis after the study is over (**Challenge 4**)
- *Winsorization*
  - Basic Idea: replaces the "outliers" in each of the state variable with cutting-point values
  - Why? Reduce the impact of outliers on the estimated policy (**Challenge 1**)
- *Proxy Value*
  - Construct a "dosage" variable based on the past pushes to capture the user's burden/habitation. Increase by 2 if a push was sent since last decision time, otherwise decrease by 1.
  - Use the current dosage to form a proxy of the future discounted sum of the reward under a working model.
  - Select the current action based on the immediate reward and the proxy of future discounted rewards.
  - Why? Pick up the negative impact of sending the treatment (**Challenge 4**)

Future Work

- Implement and evaluate the proposed algorithm in the upcoming HS 2.0 study.
- Provide a theoretical guarantee (regret bound)
- Online model checking/monitoring?
- Design the RL algorithm that pulls the information across different participants in the study.



Morning Commute

Lunch Time

Mid Afternoon

Evening Commute

After Dinner

End of Day

Available?

No

Do Nothing

Yes

Context

Location Work/Home

Number of App Screens yesterday

Past Push Notifications

Step Variation

RL

- Posterior Distribution
- Proxy Future Value

R

π_t

$1 - \pi_t$

Tailored Activity Message

Do Nothing

Decision Making

RL

- Posterior Distribution
- Proxy Future Value

Update

Context

×5

Outcome

Steps taken 30-min after decision time

×5

RL

- Posterior Distribution
- Proxy Future Value

Nightly Update

MD2K Center of Excellence for Mobile Sensor Data-to-Knowledge

Advancing biomedical discovery and improving health through mobile sensor big data

Cornell Tech ♦ Georgia Tech ♦ U. Memphis ♦ Northwestern ♦ Ohio State ♦ Open mHealth
Rice ♦ UCLA ♦ UC San Diego ♦ UC San Francisco ♦ UMass Amherst ♦ U. Michigan ♦ WVU

MD2K is supported by the National Institutes of Health Big Data to Knowledge Initiative Grant #1U54EB020404