Optimizing Adaptive Notifications in Mobile Health Intervention systems

Shihan Wang, Chao Zhang, Ben Kröse and Herke van Hoof

Intelligent Systems Lab, Utrecht University

May 4nd, 2021

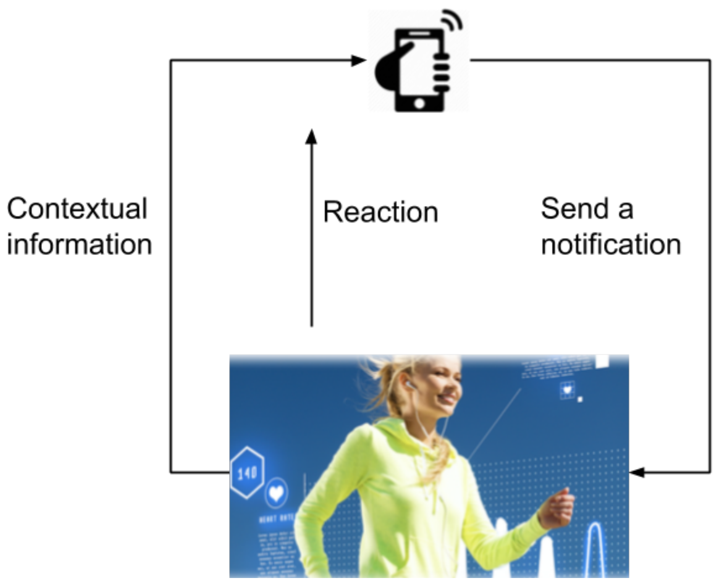
MOBILE HEALTH INTERVENTION SYSTEMS

- Use interventions to interact with users for healthcare purposes
- Just-in-time adaptive intervention [Nahum-Shani, et al. 2017]: provide right type of interventions at the right moment

Deliver interventions by adapting to user's momentary contextual situation

OUR RESEARCH FOCUS

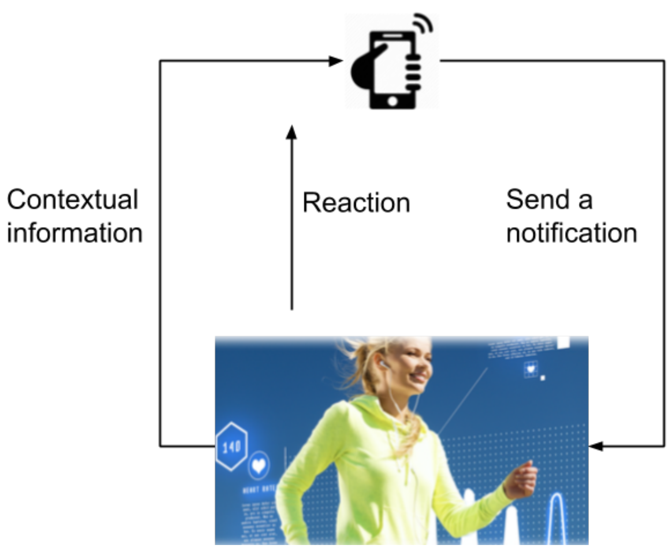
- Optimize the delivery of context-aware notifications in mobile health intervention systems:

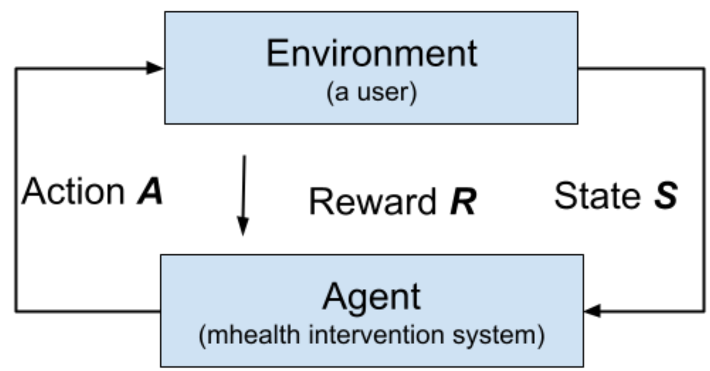


Find the optimal strategy to deliver notifications with respect to the user's momentary contextual information for promoting a target activity

USE REINFORCEMENT LEARNING (RL)

- RL is widely used as it can take users' feedback into consideration for adaptively optimizing the delivery strategy.
- The optimization task can be modeled it as a Markov decision process (MDP):





S: observed contextual features

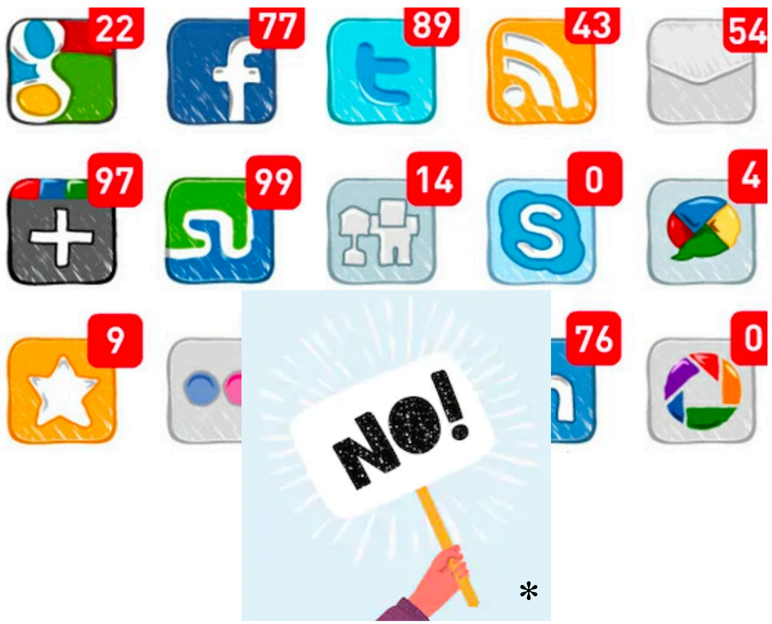
A: 'sending notification' or not

R: performance of target activities

Utrecht University

CHALLENGE

RL approaches often require the agent to interact many times with the environment prior to performing well.



Not desirable in practical applications like ours!

RL TO OPTIMIZE MOBILE HEALTH INTERVENTIONS

Learn from online data

- Learn fast using data from similar users [Tabatabaei et al. 2018 & Tomkins et al. 2019]
- Transfer the common knowledge acquired in other environments to get faster convergence [Gonul et al. 2018]

Still interact too much in the short period of early stage

Learn from historical data

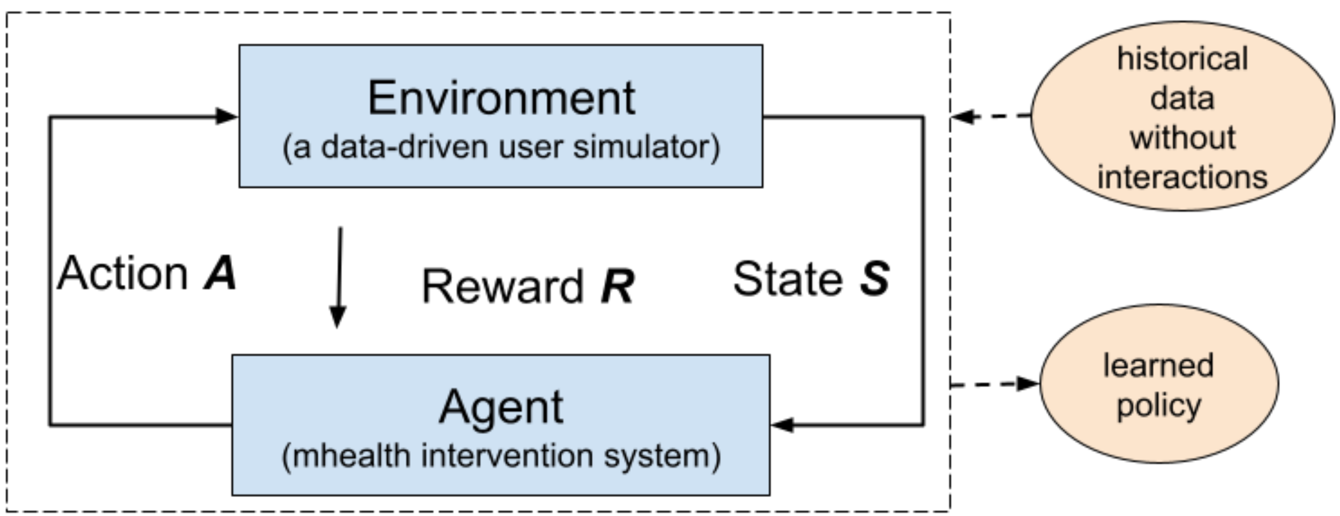
- Integrate collected data in a multi-armed bandit optimization process [Ameko et al. 2020]
- Integrate collected data in a MDP optimization process [Liao et al. 2020]

Data collection is hard:
- small size (with interactions)
- miss outcome information (without interactions)

Utrecht University

OUR APPROACH

1. Learn from historical data that can be collected without interactions
2. Introduce constrains of interaction frequency in the RL algorithm

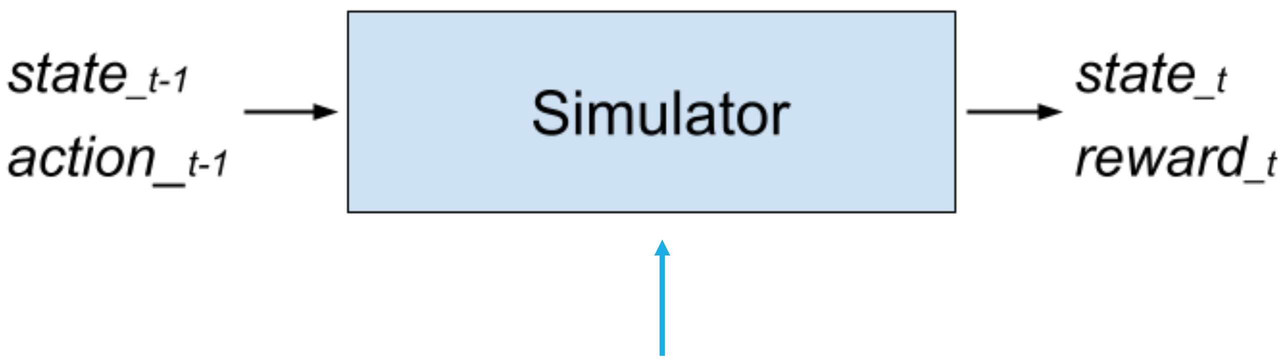


Utrecht University

1. CONTEXT-AWARE RL

- Incorporate the prior knowledge from historical data & psychological insights to optimize the delivery strategy

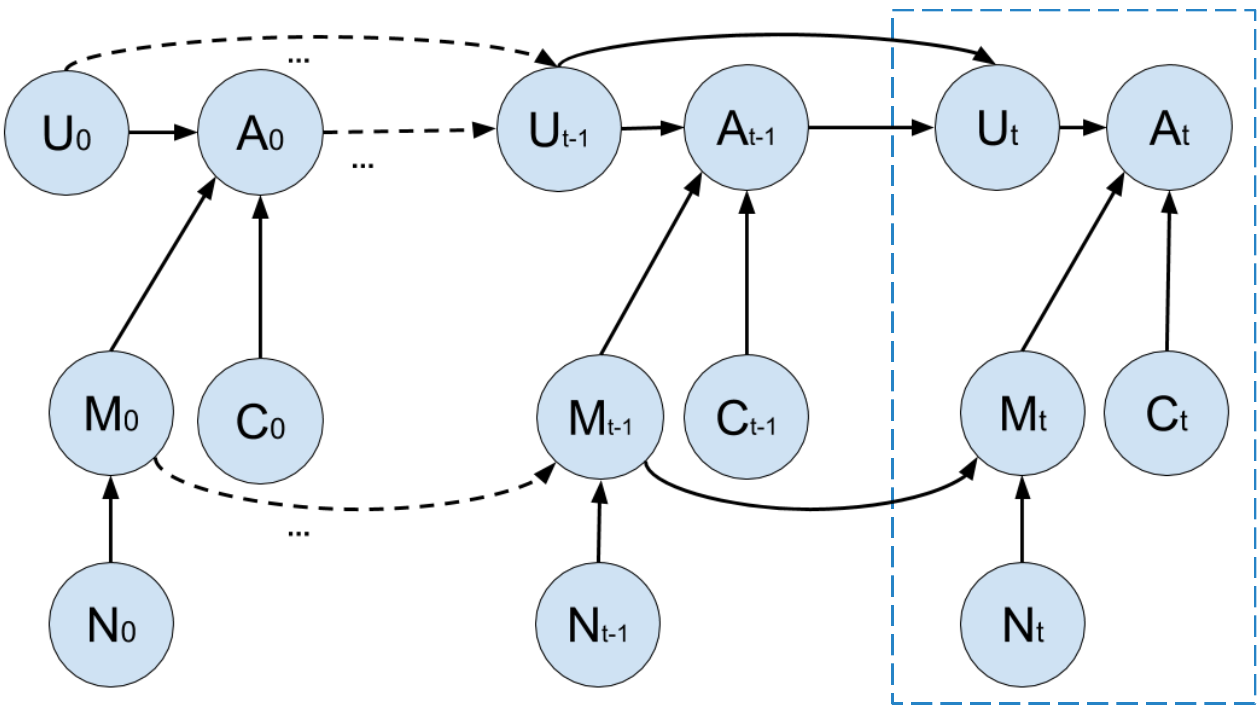
- Build up a simulator that behaves like a human:



Pr (target activity | contextual state, action)

DYNAMIC BAYES NETWORK

Integrate the influence of contextual state and notification:



A: whether the target activity is performed

C: user's momentary context

U: user's urge to perform activity

M: user's memory strength of activity

N: whether the notification is received

INFERENCE FOR DECISION MAKING

Probability of performing the activity after receiving notifications:

$$
\begin{aligned}
&P(A_t|M_{0\dots t-1}, U_{0\dots t-1}, C_{0\dots t}, N_{0\dots t}) \\
&= P(A_t|C_t, N_t, A_{t-1}, M_{t-1}, U_{t-1}) \\
&= \sum_{M_t}\sum_{U_t} \frac{P(C_t|A_t)}{P(C_t)} \cdot \underline{P(A_t|U_t, M_t)} \cdot \underline{P(M_t|N_t, M_{t-1})} \cdot \underline{P(U_t|A_{t-1}, U_{t-1})}.
\end{aligned}
$$

Define by psychological theories

Learn from historical data

A: target activity C: context U: urge M: memory strength N: notification

Utrecht University

2. RESTRICTED RL

Restrict the maximum number of notifications sent in each episode

1. Add a variable of state in the MDP model
2. Adapt the RL algorithm

Each state *s*:

-Notification left

-Time from last run

-Time from last notification

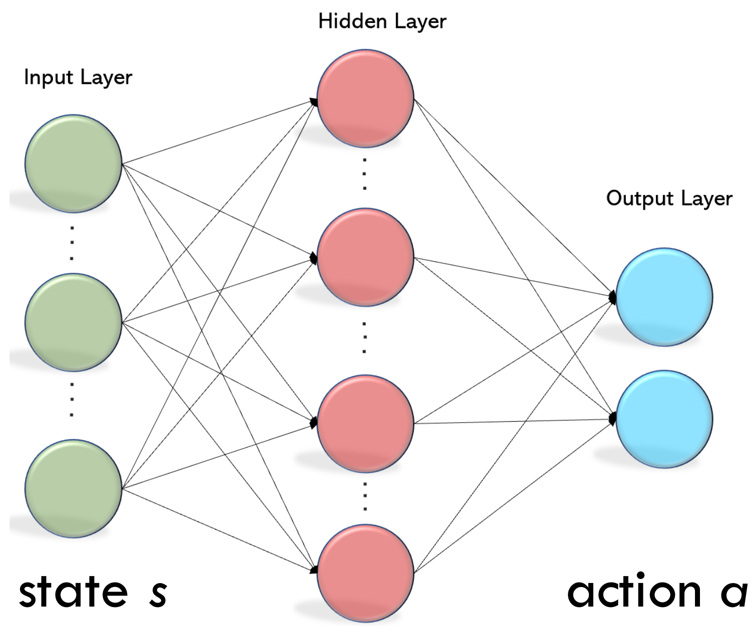
-Contextual variables…

Utrecht University

THE RESTRICTED REINFORCE ALGORITHM

- A policy gradient RL algorithm [Williams, 1992]
- Optimize the mapping between state and action $\pi(A|S,\theta)$
- Learn to 'send restricted amount of notifications' in each episode

Restrict the probability of a certain action being taken in each episode



CASE STUDY

Utrecht University

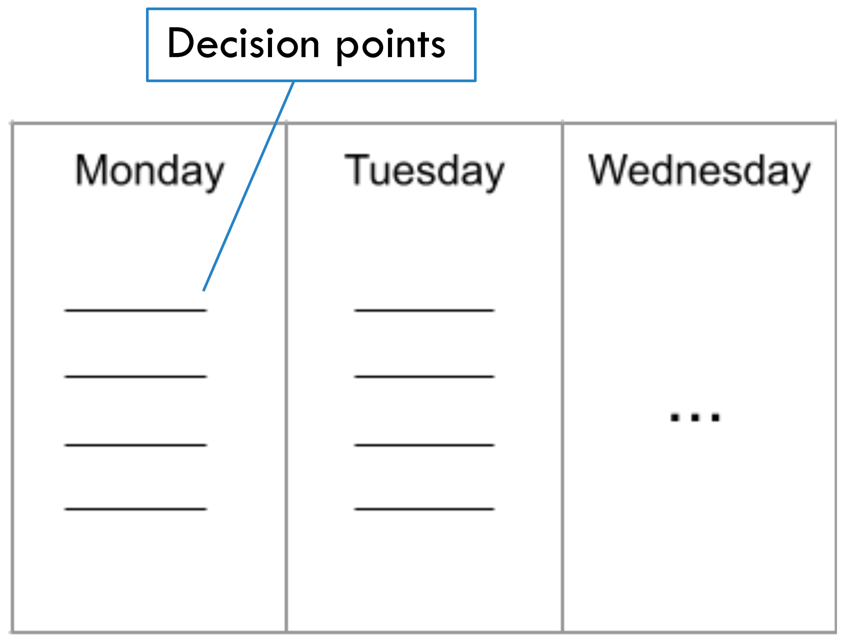
Playful Data-driven Active Urban Living (PAUL): funded by NWO and SIA grant 629.004.013

Motivate people to participate in more running activities using an intelligent mobile system

PRACTICAL TASK

- Find the optimal strategy to deliver notifications with respect to the user's momentary contextual information for promoting running activities

*



Restrict interaction frequency: 14 notifications within a week

THE DATASETS

The historical running data: $P(C_t|A_t)$

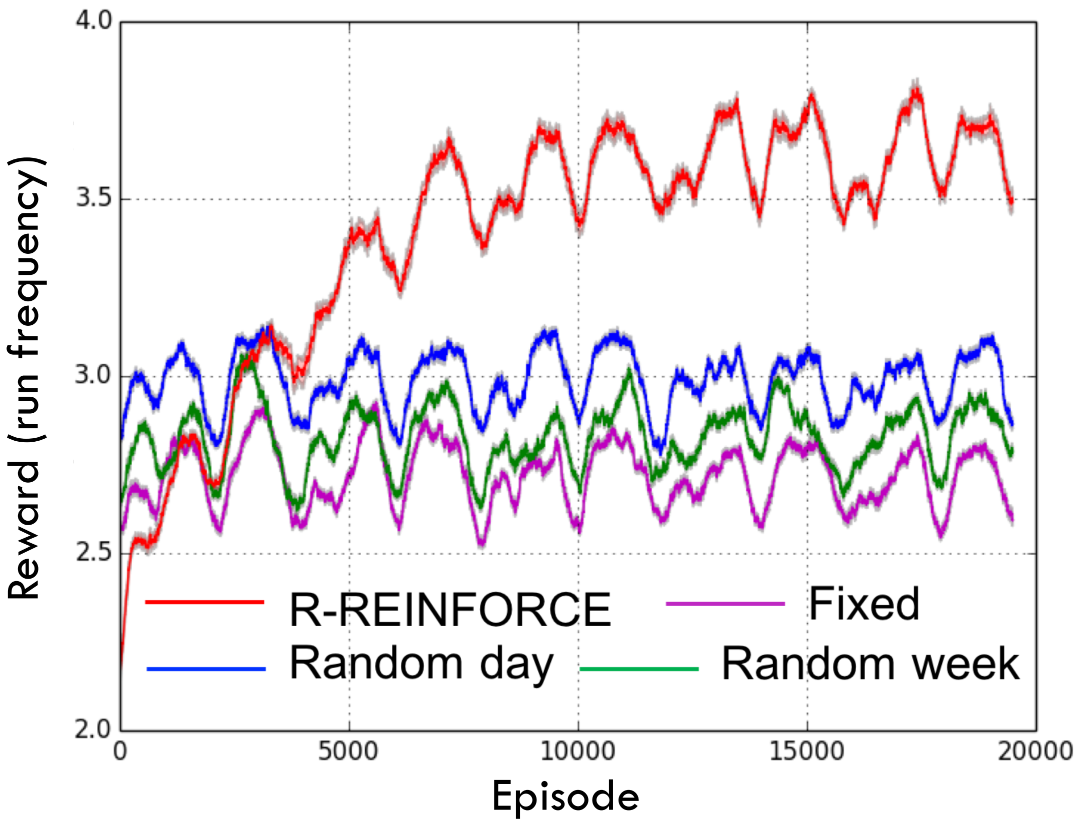
- Collected by fitness mobile application and sensors
- Including about 440K runs of over 10K users from 2013 to 2017
- Covering time and weather data at the start point of a run

The Dutch weather data: $P(C_t)$

- Collected by KNMI from 2013 to 2017
- Covering hourly time and weather data around the Netherlands

RESULT 1: EFFICIENCY OF CONTEXT-AWARE RL

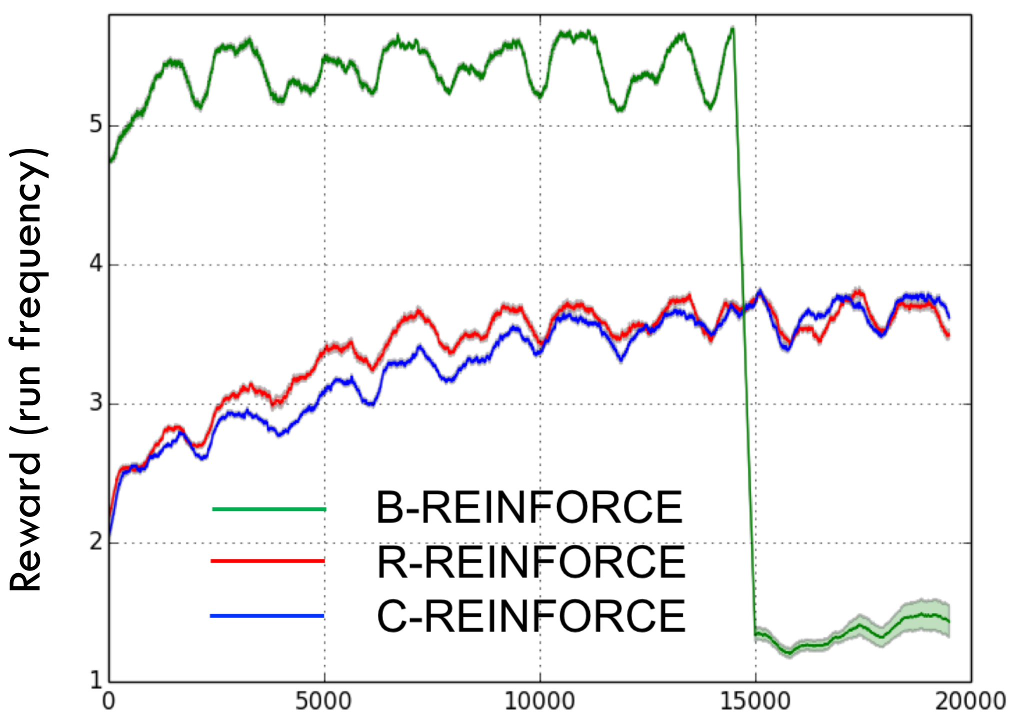
The learned policy from data & psychological insights is more efficient than rule-based policies.



Incorporating the prior knowledge from data & psychological insights can optimize the strategy of delivering context-aware notifications

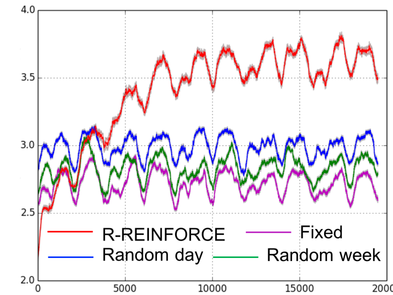
RESULT 2: EFFICIENCY OF RESTRICTED RL

Integrating interaction frequency ‘restriction’ in learning process is more efficient than learning without restriction.

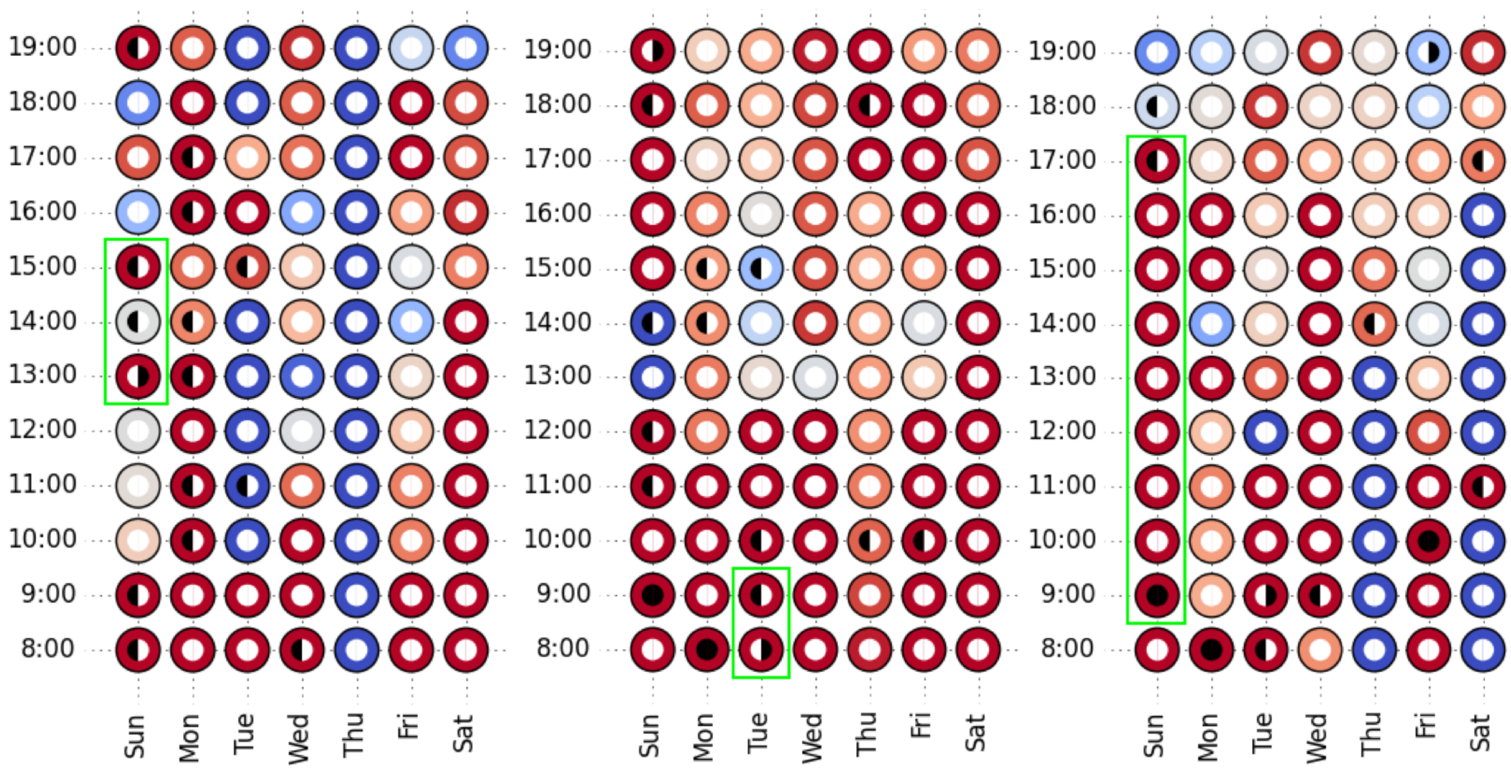


Learning with interaction frequency restriction is essential for practical application like ours.

INTERPRETATION OF LEARNED POLICY



- Information of three episodes in the learning process of the R-REINFORCE agent



(a) episode 100 (b) episode 1500 (c) episode 16000

TAKE-AWAY MESSAGES & FUTURE WORK

✓ We explore the practical usage of RL-based agents in mobile health intervention, and provide an approach to restrict interaction burden in the RL learning procedure.

1. Incorporate the prior knowledge from data & psychological insights can optimize the strategy of delivering context-aware notifications
2. It is essential to take the frequency restriction of certain actions into account in the learning process of RL agents.

➢ What's next ?

- A feasibility study (7 users in one week) has been conducted
- A pilot study with comparable user groups is planned (delayed because of the coronavirus situation)

REFERENCE

[1] Nahum-Shani, I. et al.: Just-in-time adaptive interventions (jitais) in mobile health: key components and design principles for ongoing health behavior support. Annals of Behavioral Medicine (2017)

[2] Gonul, S. et al.: Optimization of just-in-time adaptive interventions using reinforcement learning. In: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (2018)

[3] Tabatabaei, S.A. et al.: Narrowing reinforcement learning: Overcoming the cold start problem for personalized health interventions. In: Inter- national Conference on Principles and Practice of Multi-Agent Systems (2018)

[4] Tomkins,S. et al. : Intelligent pooling in thompson sampling for rapid personalization in mobile health (2019)

[5] Ameko, M.K. et al: Offline contextual multi-armed bandits for mobile health interventions: A case study on emotion regulation. Fourteenth ACM Conference on Recommender Systems (2020) .

[6] Liao, P. et al: Personalized heartsteps: A reinforce- ment learning algorithm for optimizing physical activity. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (2020)

[7] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine learning (1992)

All images with * are from the internet.

Q & A

*

Contact dr. Shihan Wang: s.wang2@uu.nl