Optimizing Rebalance Scheme for Dock-less Bike Sharing Systems with Adaptive User Incentive

Yubin Duan and Jie Wu

Email: yubin.duan@temple.edu

Dept. of Computer and Info. Sciences, Temple University

Temple University, USA

Road Map

- Introduction
- Problem Formulation
- Algorithm Design
- Simulations
- Summary

1. Introduction

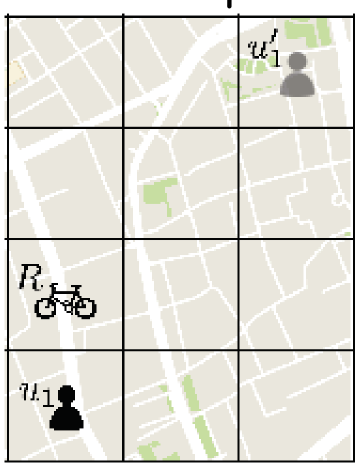
- Rebalancing dock-less bike sharing systems
  - underflow: an area lack of bikes, users cannot rent bikes
  - overflow: an area full of bikes, users cannot return bikes

- Existing incentive scheme: source detour
  - Incentivize users to rent bikes at alternative locations with detour

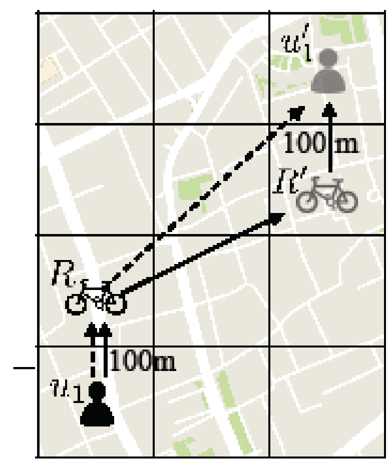
[1] A Deep Reinforcement Learning Framework for Rebalancing Dockless Bike Sharing Systems (AAAI '19)

Motivation

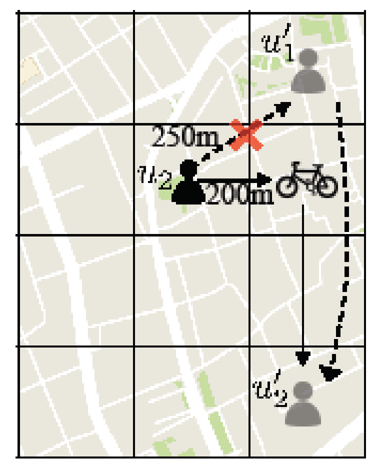
- Destination detour
  - A complement of source detour
  - Example of destination incentive
    - temporal and spatial domains are discretized



user u_1 arrives with destination u_1'



offer source incentive only (dashed lines)
vs.
offer source and dest. Incentive (solid lines)

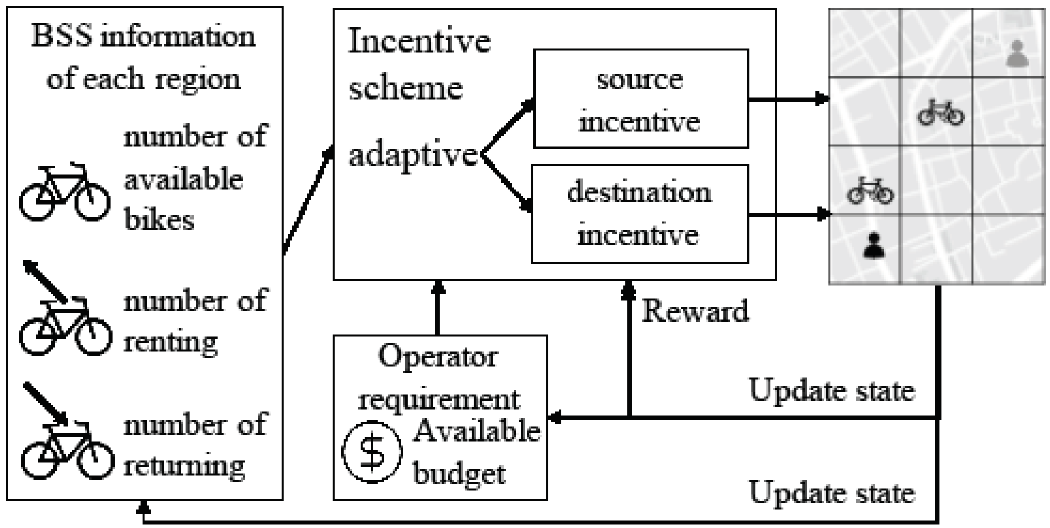


user u_2 arrives with destination u_2'

- Objective
  - Maximize the total number of satisfied users over a day under a given budget

2. Model & Problem Formulation

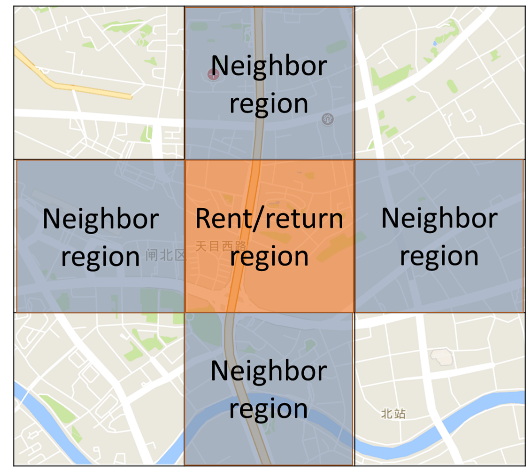
- Rebalance scheme overview



- System dynamic are modeled by Markov Decision Process (MDP)
  - state: supply *S(t)*, rent demand *D(t)*, arrival *A(t)*, and remaining budget *RB(t)*
  - action: source incentive price $P^{+}(t)$, and dest. incentive price $P^{-}(t)$
  - reward: number of satisfied requests (could successfully rent and return)

2. Model & Problem Formulation

- User model (Environment model)
  - Cost of detour δ [2]
    - 0 if the alternative renting or returning locations are in the same region
    - $C + \eta\delta^2$ if the alternative renting or returning locations are in neighbor regions
    - $+\infty$ otherwise



  - A user accepts the incentive only if the incentive price is higher than the cost

- Problem formulation

$$\max \quad \sum_{t=1}^{m} \sum_{i,j=1}^{n} \tau_{ij}(t) \qquad (1)$$ objective

$$\text{s.t.} \quad \sum_{t=1}^{m} \sum_{i=1}^{n} p_i^{+}(t) < B - B^{-} \qquad (2)$$ source incentive budget

$$\sum_{t=1}^{m} \sum_{i=1}^{n} p_i^{-}(t) \le B^{-} \qquad (3)$$ dest. incentive budget

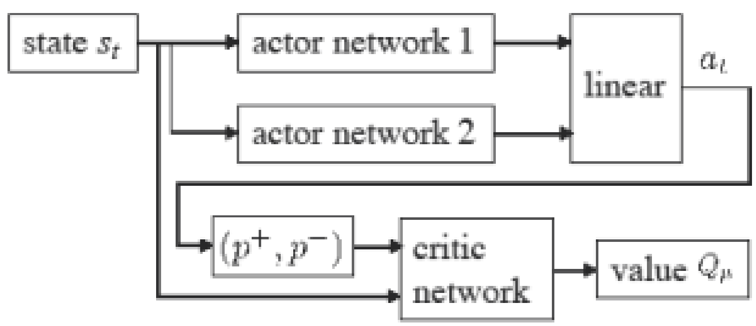
$$\sum_{j=1}^{n} \tau_{ji}(t) - \sum_{j=1}^{n} \tau_{ij}(t) \le \varphi_i(t), \forall i, t \qquad (4)$$ inventory constraint

$$\varphi_i(t+1) = \varphi_i(t) + \sum_{j=1}^{n} (\tau_{ji}(t) - \tau_{ij}(t)) \forall i, t. \qquad (5)$$ inventory evolution

[2] Incentivizing users for balancing bike sharing systems (AAAI '15)

3. Algorithm Design

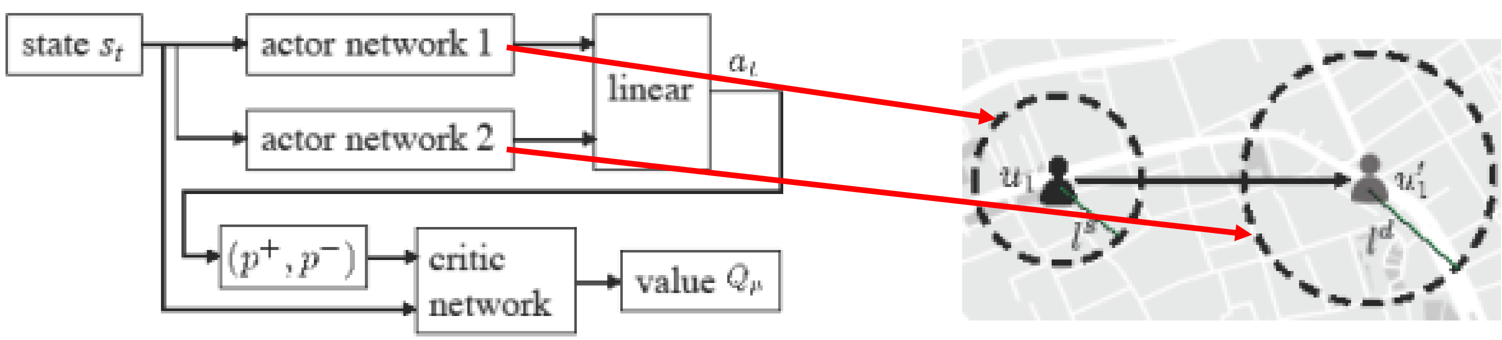
- An existing solution for source incentive[1]
  - Based on the deep deterministic policy gradients algorithm[3]
  - They decompose the Q-value of the entire area of interest into multiple sub-Q-values of smaller regions.
- A hybrid incentive scheme
  - Considering both source and destination incentives
  - Extend the actor-critic framework used in [1]



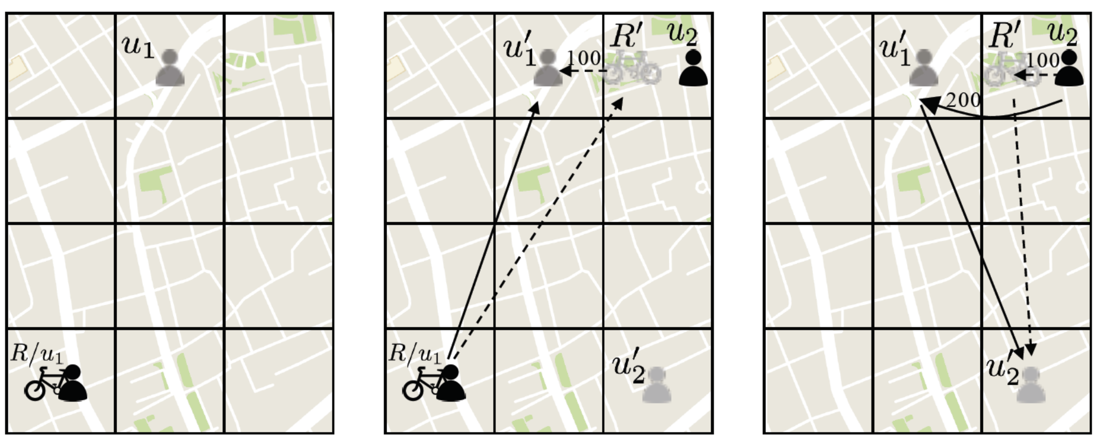
[1] A Deep Reinforcement Learning Framework for Rebalancing Dockless Bike Sharing Systems (AAAI '19)
[3] Continuous control with deep reinforcement learning (ICLR '16)

3. Algorithm Design

- Intuition of parameters in actor networks
  - Adaptively adjust the strength (the maximum detour distance) for source and destination incentives



  - Example: avoiding long-distance detour in our cost model ($\propto \delta^2$)



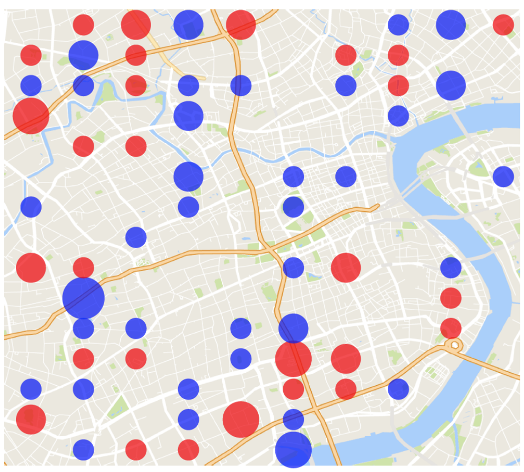
(a) t=0 (b) t=1 (c) t=2

solid lines: source incentive only

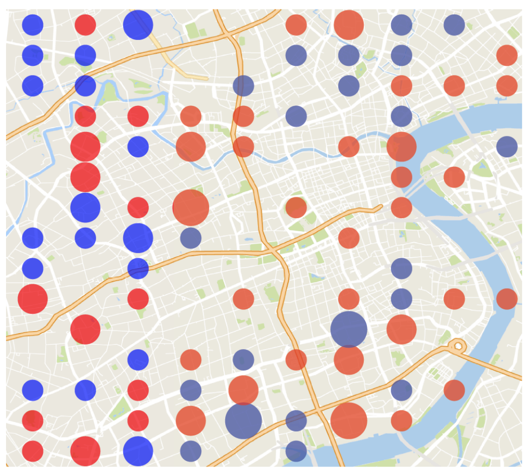
dashed lines: combining source and destination incentives (two shorter detours)

4. Simulations

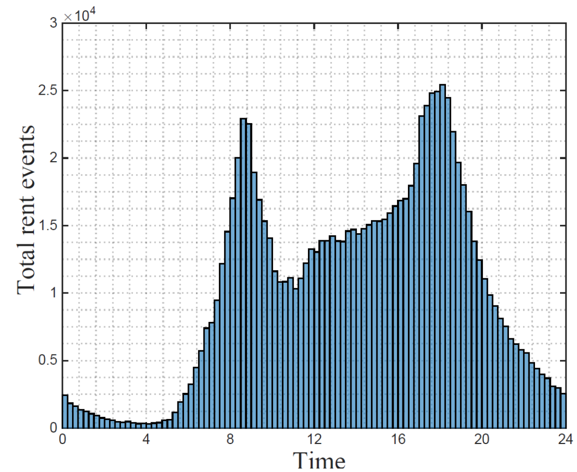
- Mobike dataset
  - One-month history trip data from 8/1/2016 to 9/1/2016 in Shanghai City
- Temporal and spatial imbalance distribution



Pick-up events in 8 a.m. – 9 a.m.



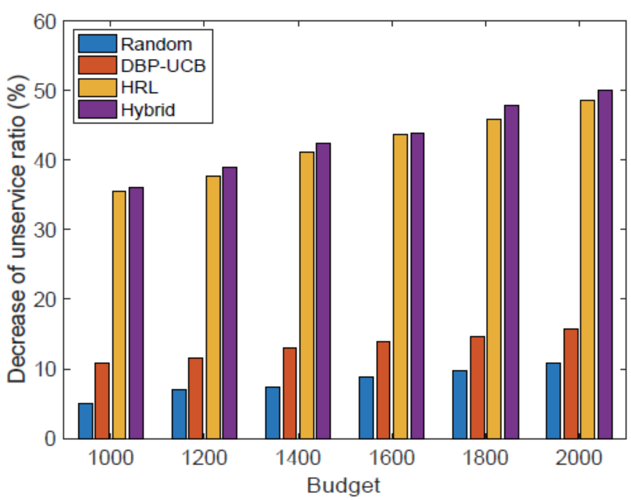
Pick-up events in 6 p.m. – 7 p.m.

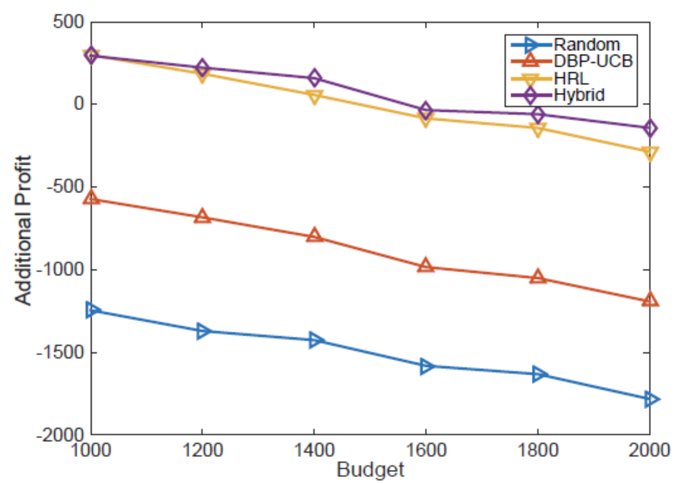


Rent events distribution

4. Simulations

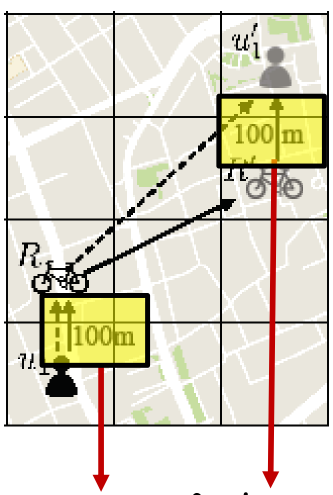
- Comparison in decreased unserviced ratio (DUR)
  - DUR = (N1-N2)/N1
  - N1: number of unsatisfied users **without** incentive
  - N2:number of unsatisfied users **with** incentive
  - The hybrid incentive scheme achieves the highest DUR
- Comparison in profit
  - Assume each successful ride has 1 reward
  - The Hybrid incentive scheme earns positive profit



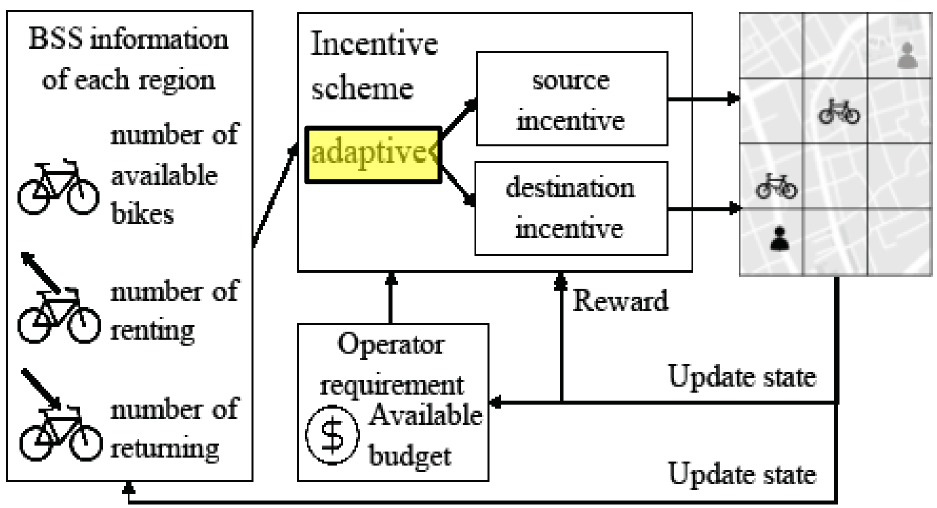


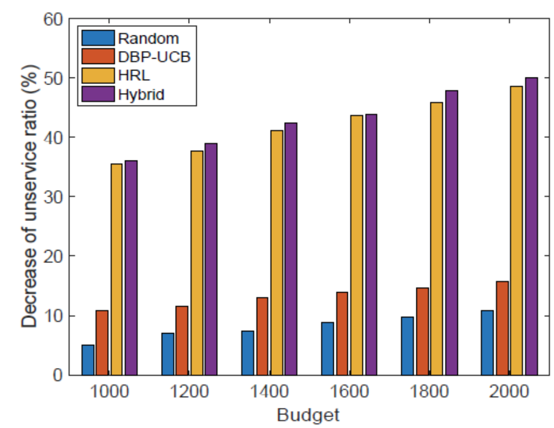
5. Summary

- We show the advantages of destination incentives
- We propose the hybrid incentive scheme
- Experiments on the real-world dataset show the efficiency of our hybrid scheme



source & dest. incentive





Thank you

Q & A