Balancing Accuracy and Fairness for Interactive Recommendation with Reinforcement Learning

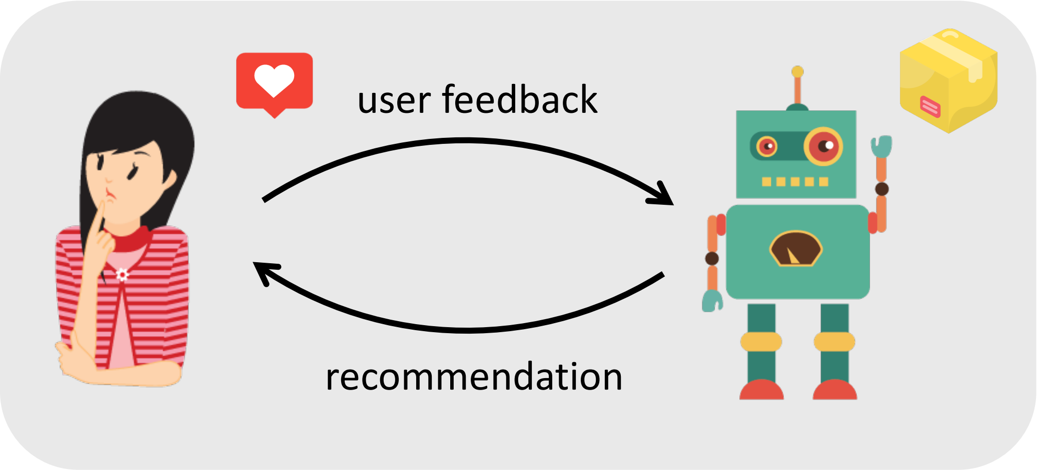
Weiwen Liu, Feng Liu, Ruiming Tang, Ben Liao, Guangyong Chen, Pheng Ann Heng

Huawei Noah’ s Ark Lab

liuweiwen8@Huawei.com

Interactive Recommender Systems (IRS)

- IRS consecutively recommend items to individual users and receive their feedback in **interactive** processes.
  - Users are involved in the recommendation procedure.
  - Gradually refine the recommendation policy according to the obtained user feedback.
  - To maximize the total utility over the whole interaction period.



Conversion Rate (CVR)

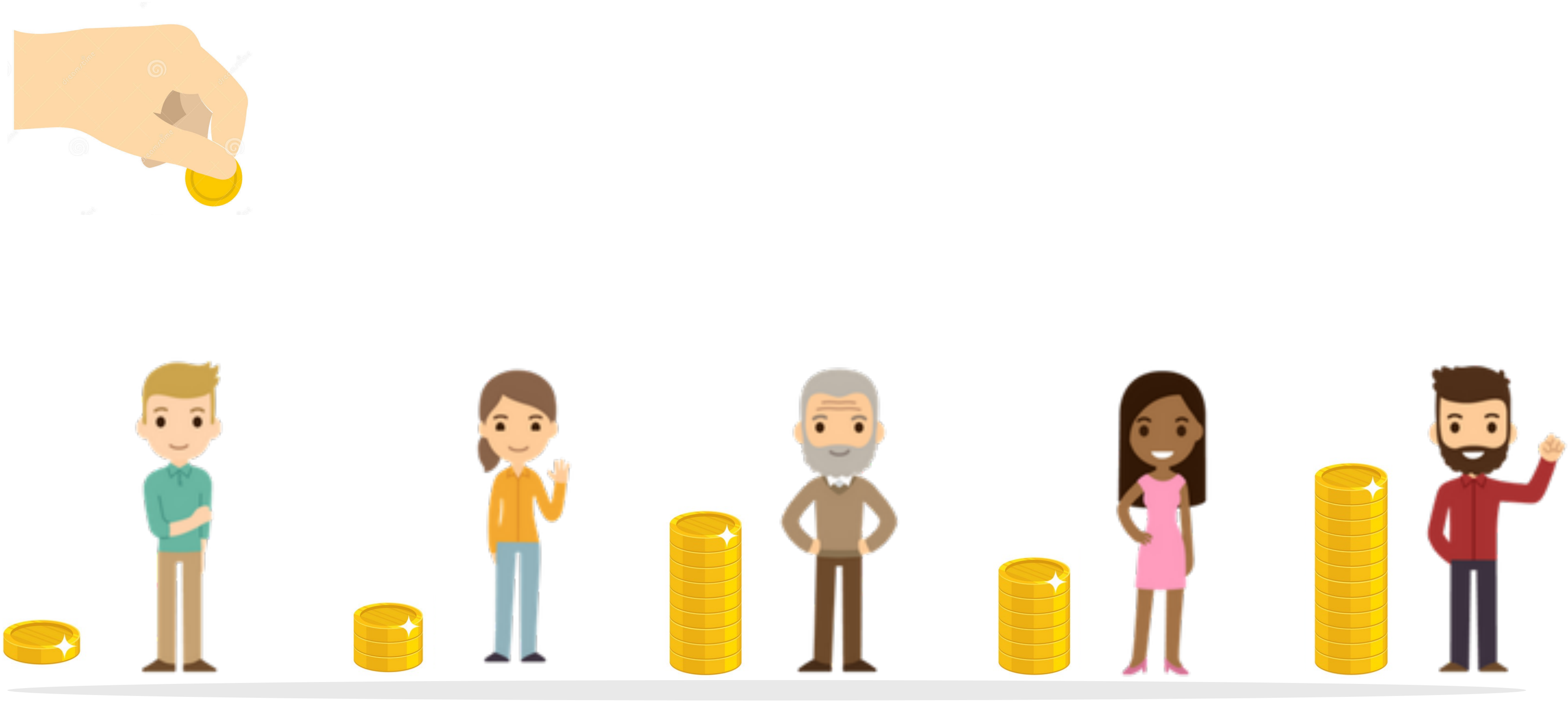
- Measuring recommendation acceptance:

$$\frac{\#system's\ desired\ activity}{\#impressions}.$$

- A system's desired activity could be downloading from App stores, purchases, or making loans for microlending.
- CVR is one of the most commonly used **objectives** for interactive recommender systems.

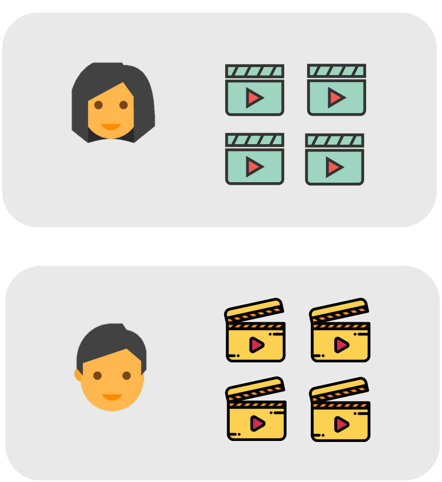
Fairness Issues

- Optimizing CVR solely may result in fairness issues, one of which is the unfair allocation of **desired activities** over different demographic groups.

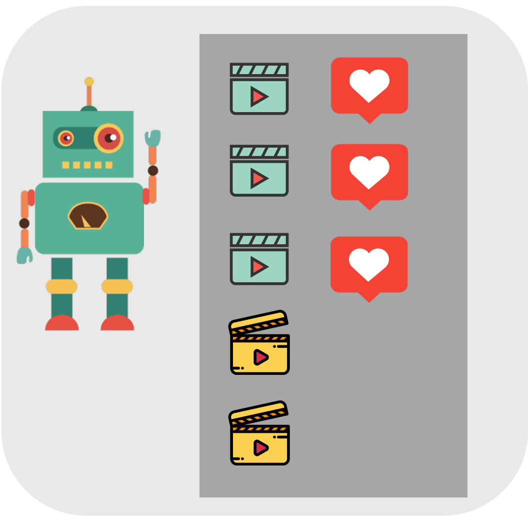


Importance of a Fair Allocation

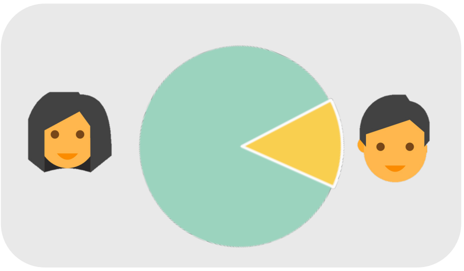
- **Legal**
  - In the setting of employment, education, housing, or public accommodation, a fair treatment with respect to race, color, religion, etc., is required by the anti-discrimination laws.
- **Financial**
  - Under-representing for some groups leads to the abandonment of the system.



creators (item groups)



recommended lists



Fairness vs. Recommendation Accuracy

- **Fairness**
  - Ideal Fairness ⇒ **equally** divide the recommendation opportunities to each item group.
  - Users' **satisfaction** will be affected.

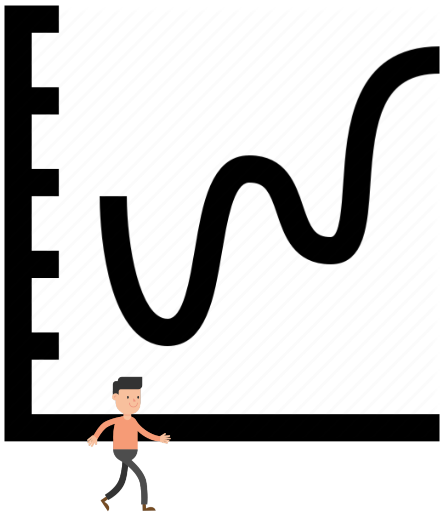
- **Recommendation Accuracy**
  - Recommendation Accuracy / Personalization ⇒ to estimate users' preferences accurately, maximizing **CVR**.
  - Has been proved to favor **popular items** [Òscar et al., 2008].
  - Usually leads to extremely **unbalanced** recommendation results.

Can we achieve a fairer recommendation while preserving recommendation accuracy?

[Òscar et al., 2008] Celma, Òscar, and Pedro Cano. "From hits to niches? or how popular artists can bias music recommendation and discovery." *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*. 2008.

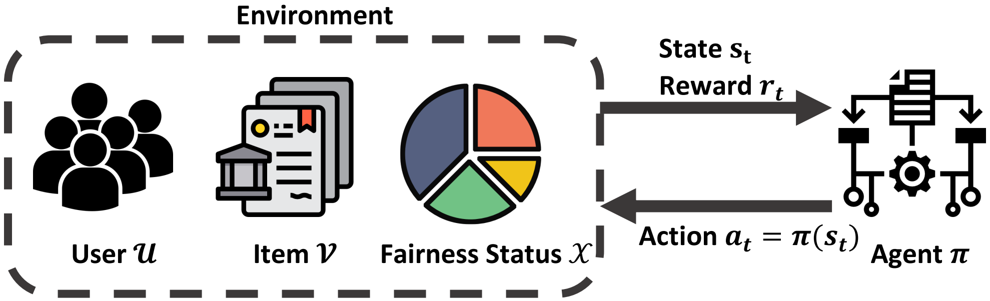
Fairness vs. Recommendation Accuracy

- **Long-term Cumulative Utility**
  - **Users with particular favor:** focusing on improving accuracy (CVR)
  - **Users with diversified interests:** focusing on improving fairness
  - the lack of fairness at one point time can be compensated for a later time.
- **Fair allocation of desired activities:**
  - Existing work only considers the allocation of the number of recommendations (exposure).
  - The distribution of desired activities has much larger commercial value.



Fairness-aware Recommendation (FairRec)

- When a user u arrives at time step $t = 1, \dots, T$,
  - The current state s_t is observed:
    - **User Preference State:** the user's N most recent positively interacted items.
    - **Fairness State:** the current allocation distribution of the desired activities x_t at time t.
  - The system takes an action a_t and recommends an item to the user.
  - The user views the recommended item and provides feedback y_{a_t}.
  - The system then receives a reward r_t (a function of y_{a_t}) and updates the model. The objective function is the long-term discounted reward: $R_t = \sum_{k=t}^{T} \gamma^{k-t} r_k$.



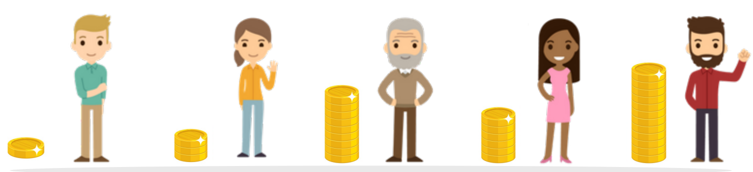
Weighted Proportional Fairness for IRS

Definition 1 (Weighted Proportional Fairness) [Frank et al., 1998].

An allocation of desired activities x_t is weighted proportionally fair if it is the solution of the following optimization problem,

$$\max_{x_t} \sum_{i=1}^{l} w_i\left(\log x_t^i\right), \quad s.t. \sum_{i=1}^{l} x_t^l = 1,\ x_t^l \geq 0, i = 1,..,l\,.$$

- The coefficient $w_i \geq 0$ is a pre-defined parameter weighing the importance of each group.
- The optimal solution can be easily solved by standard Lagrangian multiplier methods, namely $x_*^i = \frac{w_i}{\sum_{j=1}^{l} w_j}$.

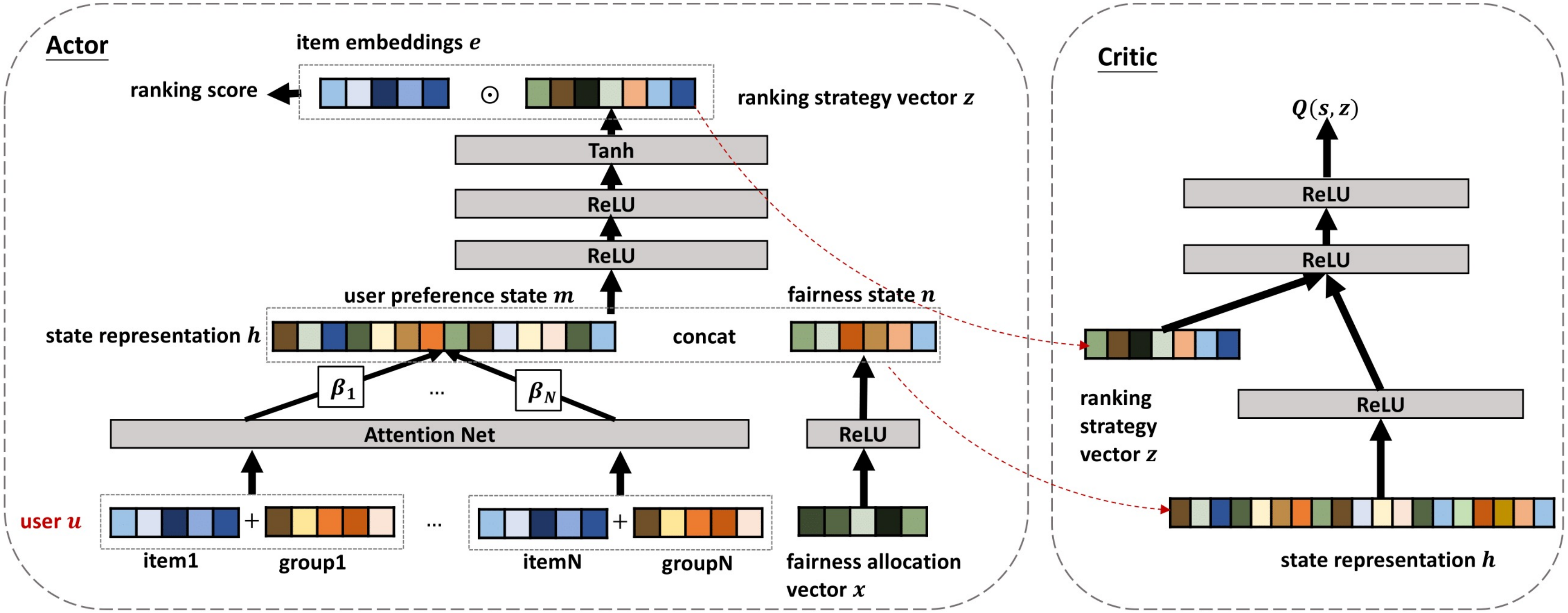


$x_t = [0.03,\quad 0.1,\quad 0.29,\quad 0.19,\quad 0.39]$

[Frank et al., 1998] Kelly, Frank P., Aman K. Maulloo, and David KH Tan. "Rate control for communication networks: shadow prices, proportional fairness and stability." *Journal of the Operational Research society* 49.3 (1998): 237-252.

Fairness-aware Recommendation: FairRec

- We adopted an **actor-critic architecture** in reinforcement learning.
  - **Personalized Fairness-aware State Representation:** User preferences and the system's fairness status are jointly compressed into the state representation.
  - **Reward Function Design:** measuring the system's gain regarding accuracy and fairness.



Reward Function Design

- We incorporate the deviation from the optimal solution $x_*^i - x_t^i$ into the reward as the fairness indicator:

$$r_t = \begin{cases} \sum_{i=1}^{l} \mathbb{I}_{A_i}(a_t)\left(x_*^i - x_t\right) & \text{if } y_{a_t} = 1 \\ -\lambda. & \text{if } y_{a_t} = 0' \end{cases}$$

where $\mathbb{I}_A(x)$ is the indicator function and is 1 when $x \in A$, 0 otherwise, x_t^i is the allocation proportion of group i at time t. The constant $\lambda > 1$ is the penalty value for inaccurate recommendations and manages the accuracy-fairness tradeoff.

Experiments

Table 1: Experimental results on MovieLens and Kiva.

	MovieLens			Kiva		
	CVR	PropFair	UFG	CVR	PropFair	UFG
NMF	0.7972	0.8592	4.2362	0.4211	0.8473	1.4635
SVD	0.8478	0.8337	5.4795	0.4870	0.8686	1.6931
DeepFM	<u>0.8612</u>	0.8098	5.8323	0.6349	0.8671	2.3752
LinUCB	0.8577	0.8464	5.9476	0.6517	0.8697	2.4970
DRR	0.8592	0.8470	<u>6.0177</u>	<u>0.6567</u>	0.8645	<u>2.5183</u>
MRPC	0.8361	<u>0.8608</u>	5.2508	0.4286	<u>0.8761</u>	1.5332
FairRec	**0.8702***	**0.8666***	**6.6776***	**0.6905***	**0.8838***	**2.8555***

Table 2: Ablation study on MovieLens and Kiva.

	MovieLens			Kiva		
	CVR	PropFair	UFG	CVR	PropFair	UFG
FairRec(reward-)	0.8561	0.8053	5.5957	**0.6935**	0.8670	2.8290
FairRec(state-)	0.8194	**0.8758**	4.8494	0.6723	0.8746	2.6688
FairRec	**0.8702**	0.8666	**6.6776**	0.6905	**0.8838**	**2.8555**

Conclusions

- We formulate **a fairness objective** for IRS.
- We propose a reinforcement learning based framework, FairRec, to dynamically maintain a **balance** between accuracy and fairness in IRS.
- Experiments show that FairRec can achieve a better balance between accuracy and fairness, compared to the state-of-the-art methods.

THANKS☺

REFERENCES

- [Òscar et al., 2008] Celma, Òscar, and Pedro Cano. "From hits to niches? or how popular artists can bias music recommendation and discovery." *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*. 2008.
- [Frank et al., 1998] Kelly, Frank P., Aman K. Maulloo, and David KH Tan. "Rate control for communication networks: shadow prices, proportional fairness and stability." *Journal of the Operational Research society* 49.3 (1998): 237-252.