Task Recommendation in Crowdsourcing

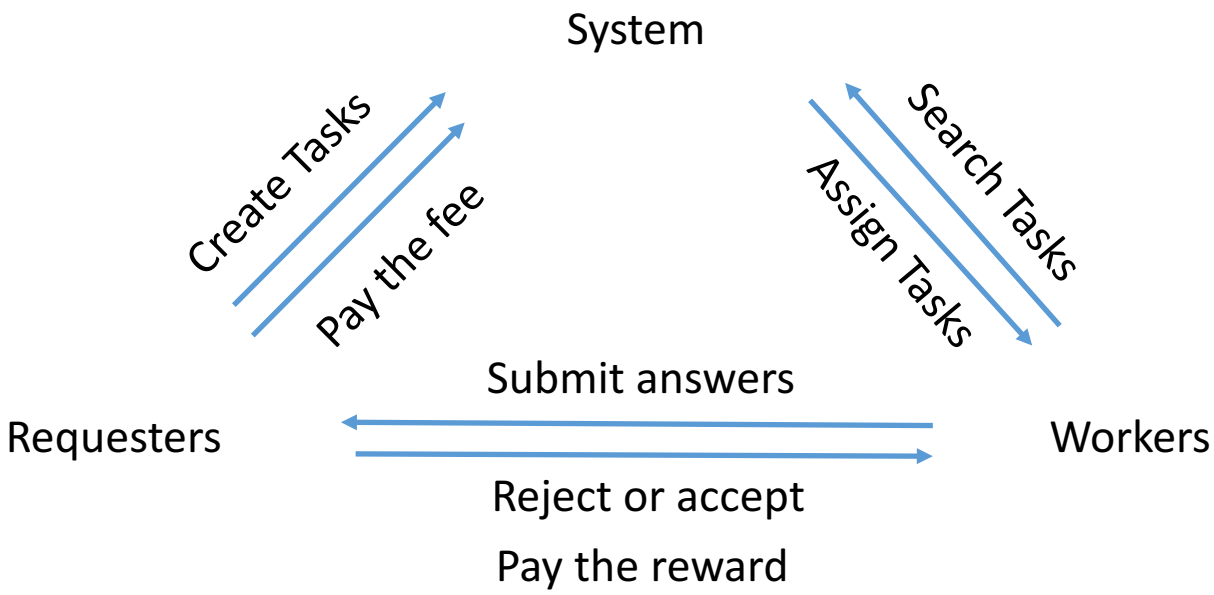
You Ming Hsu

2018.10.16

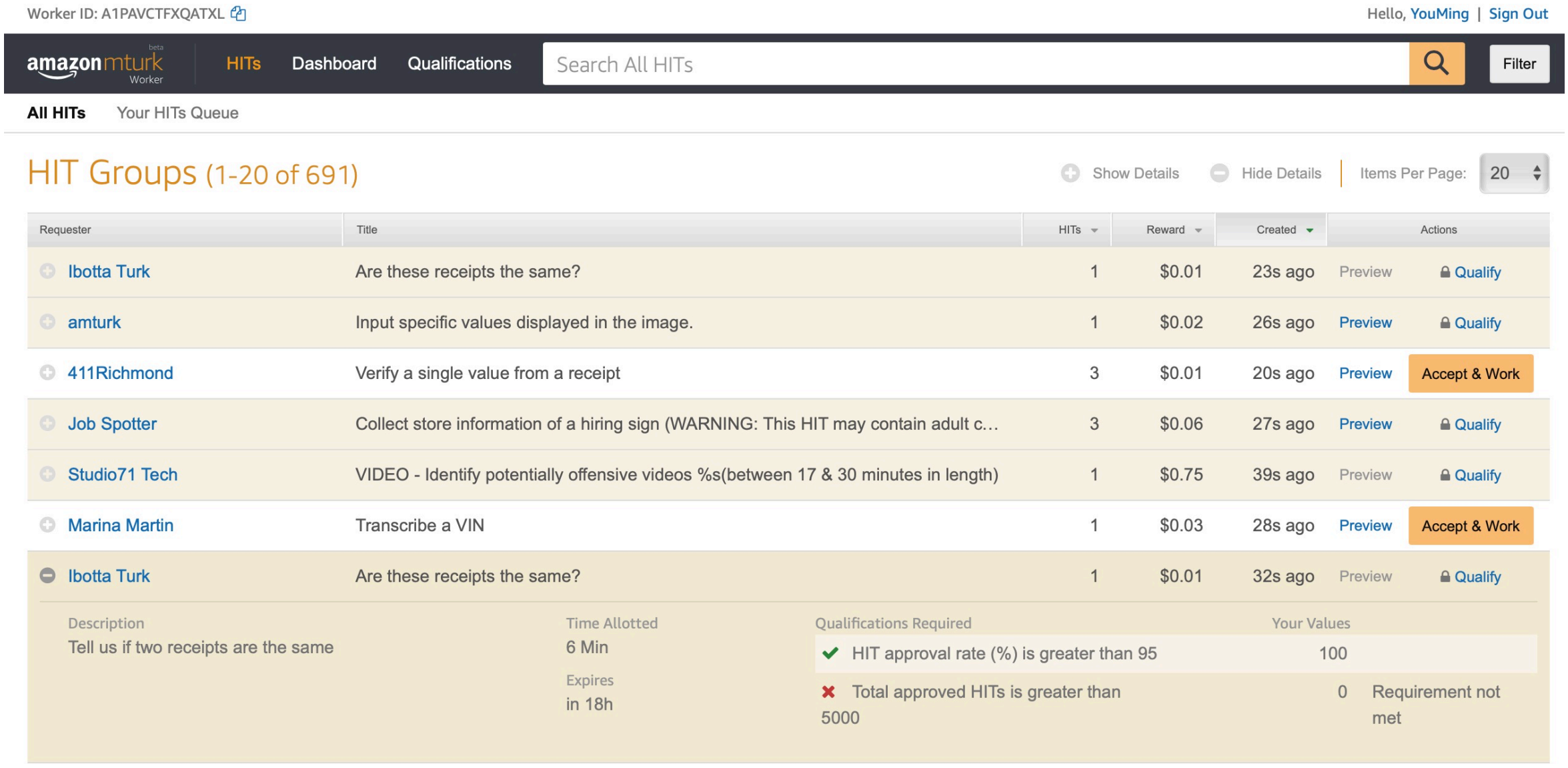
Outline

- Purpose
- Crowdsourcing Recommendation Papers
- Conclusion and Future work
- References

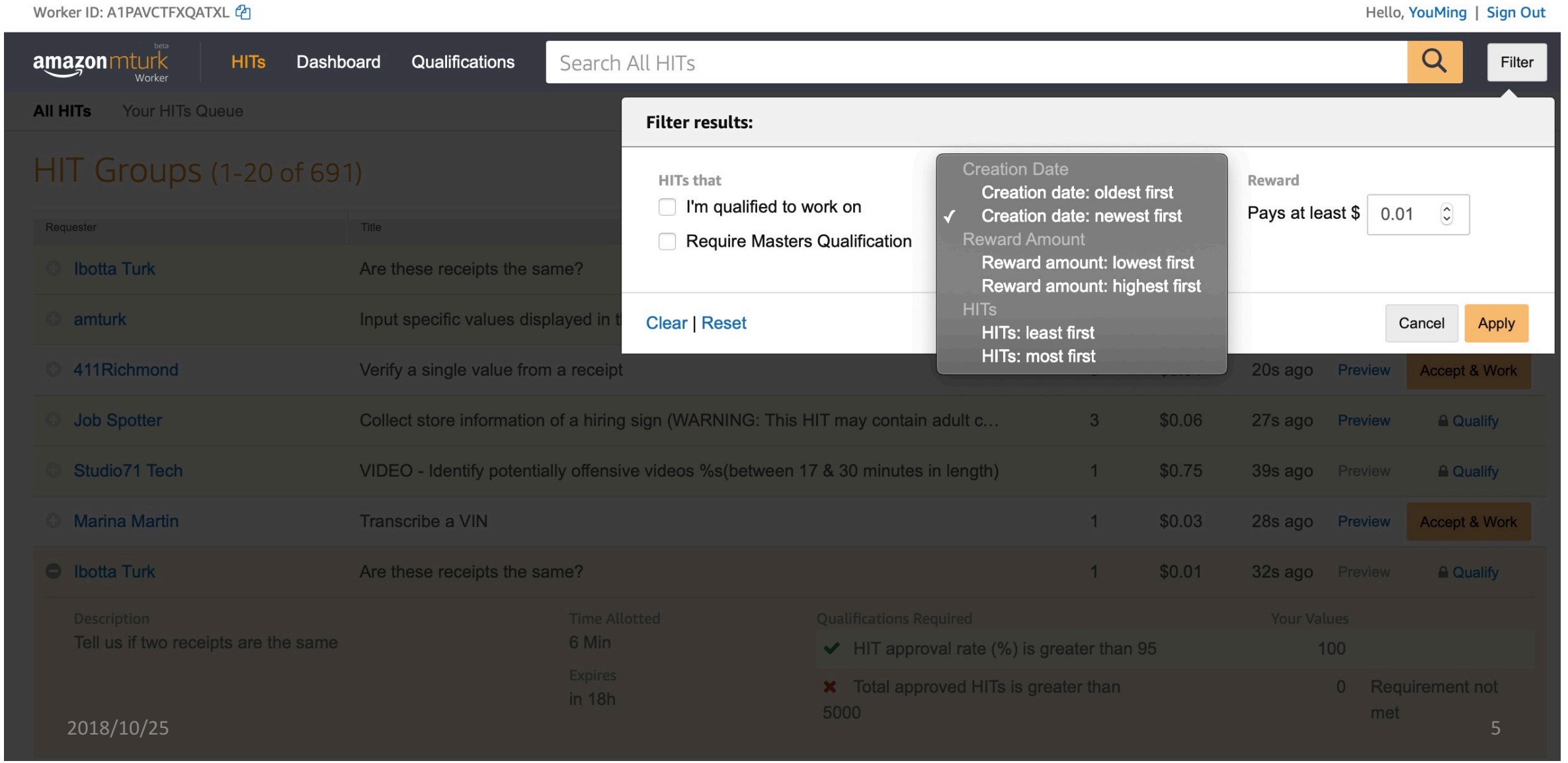
Crowdsourcing Workflow



Amazon Mechanical Turk



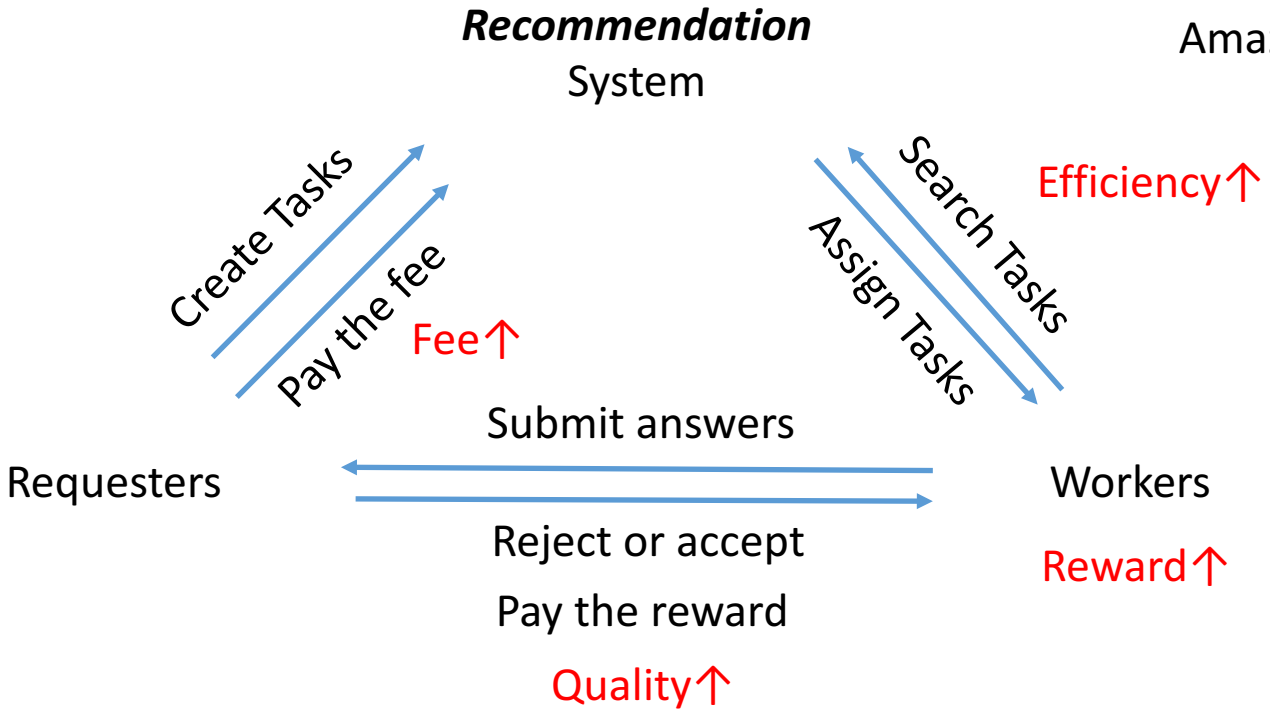
Amazon Mechanical Turk



Purpose

- Crowdsourcing Workflow

In 2011, a writer in San Francisco, claimed that he earned a measly USD$ 4.38 for spending eight hours in a day crowdsourcing for Amazon Mechanical Turk.



Bag-of-Words Approach

- The bag-of-words approach uses the vocabulary of the task description and computes the similarity as the overlap in their vocabularies.

$$bow(t) = \frac{1}{|H|} * \sum_{i=1}^{|H|} c_i * | Voc(t_i) \cap Voc(t) |$$

- H = {(t, c)}, the history of a specific user.
- t is the task and its associated features.
- c indicate the scale of preference of the user for the task, here we only consider a binary preference, and therefore c $\in$ {−1, 1}.

Ambati,V.,Vogel,S.,& Carbonell,J.G.(2011,August). Towards Task Recommendation in Micro-Task Markets. In *Human Computation*

Classification Based Approach

- The bag-of-words approach can not incorporate the other features of a task like timestamp, reward etc.
- Using maximum entropy classifier, the classifier probability can be defined as:

$$Pr(c_i|t) = \frac{1}{Z(t)} exp\left(\sum_{j=1}^{n} \lambda_j f_{ij}(c_i, t)\right)$$

- t is a task and associated features.
- c_i is the class, where c = 1 indicates a user will be interested in the task and c = −1 otherwise.
- f_{ij} are feature functions and Z(t) is a normalizing factor.
- λ_j are the weights for the feature functions.

Ambati,V.,Vogel,S.,& Carbonell,J.G.(2011,August). Towards Task Recommendation in Micro-Task Markets. In *Human Computation*

TaskRank

- Define a score of a task t_{ij} in category c_i for a worker v_y is given as follows:

$$\mathcal{TR}_{i,j}(v_y) = \mathcal{AR}_i(v_y) * \mathcal{CPS}_i(v_y) * \left(1 - \left|\frac{m_{i,j} - \mathcal{RPS}_i(v_y)}{\mathcal{RPS}_i(v_y)}\right|\right) * \left(1 - \left|\frac{e_{i,j} - \mathcal{TPS}_i(v_y)}{\mathcal{TPS}_i(v_y)}\right|\right)$$

- $AR_i(v_y) = \frac{|T_i'(v_y)|}{|T_i(v_y)|}$, the acceptance rate of tasks in category c_i for worker v_y.
- $CPS_i(v_y) = \frac{|T_i(v_y)|}{\sum_{p=1}^{C_N} |T_p(v_y)|}$, task category preference score on category c_i for worker v_y.
- $RPS_i(v_y) = \frac{\sum_{j=1}^{T_N} m_{ij}}{|T_i(v_y)|}$, reward preference score on category c_i for worker v_y.
- $TPS_i(v_y) = \frac{\sum_{j=1}^{T_N} e_{ij}}{|T_i(v_y)|}$, time alloted preference score on category c_i for worker v_y.

 Yuen, M. C., King, I., & Leung, K. S. (2011, October). Task matching in crowdsourcing. iThings/CPSCom. IEEE

Recover the worker-task preferring matrix

- To transform workers' behaviors into values as follows:

Worker Behavior		Value
Worker's work done is accepted by requester.	$\longrightarrow$	5
Worker's work done is rejected by requester.	$\longrightarrow$	4
Worker completes a task and submits the work done.	$\longrightarrow$	3
Worker selects a task to work on but not complete it.	$\longrightarrow$	2
Worker browses the detailed information of a task.	$\longrightarrow$	1
Worker does not browse the detailed information of a task.	$\longrightarrow$	0

Yuen, M. C., King, I., & Leung, K. S. (2012, August). Task recommendation in crowdsourcing systems. CrowdKDD
Yuen, M. C., King, I., & Leung, K. S. (2014). TaskRec : A Task Recommendation Framework in Crowdsourcing Systems. Neural Processing Letters

Recover the worker-task preferring matrix

- Example for worker-task matrix:

	v_1	v_2	v_3	v_4	v_5	v_6
w_1		5	2		3	
w_2	4			3		4
w_3			2			2
w_4	5			3		
w_5		5	5			3

- Using Probabilistic Matrix Factorization (PMF) to recover this matrix.
- Select the top few of the tasks for recommendation.

Yuen, M. C., King, I., & Leung, K. S. (2012, August). Task recommendation in crowdsourcing systems. CrowdKDD
Yuen, M. C., King, I., & Leung, K. S. (2014). TaskRec : A Task Recommendation Framework in Crowdsourcing Systems. Neural Processing Letters

Conclusion and Future work

- Ranking of all tasks to recommend, based on the several factors such as worker history, interest, performance and requester's feedback.
- If a worker gets a list of related recommended tasks:
  - Improving the quality of the task.
  - Increasing the solution efficiency.
  - Worker and system may get more reward.
- These kind of worker history data is difficult to obtain.

  —> Synthetic recommendation system
    - Which can evaluate the performance between different methods.

References

- Yuen, M. C., King, I., & Leung, K. S. (2012, August). Task recommendation in crowdsourcing systems. In *Proceedings of the First International Workshop on Crowdsourcing and Data Mining* (pp. 22-26). ACM.
- Ambati,V.,Vogel,S.,&Carbonell,J.G.(2011,August).Towards Task Recommendation in Micro-Task Markets. In *Human Computation* (pp. 1- 4).
- Yuen, M. C., King, I., & Leung, K. S. (2011, October). Task matching in crowdsourcing. In Internet of Things (iThings/CPSCom), 2011 International Conference on and 4th International Conference on Cyber, Physical and Social Computing (pp. 409-412). IEEE.
- Eman Aldhahri, Vivek Shandilya, and Sajjan Shiva. 2015. Towards an Eective Crowdsourcing Recommendation System: A Survey of the State-of-the-Art. InService-Oriented System Engineering (SOSE), 2015 IEEE Symposium on. IEEE, 372–377.
- Mnih, Andriy, and Ruslan Salakhutdinov. “Probabilistic matrix factorization.” NIPS 2007