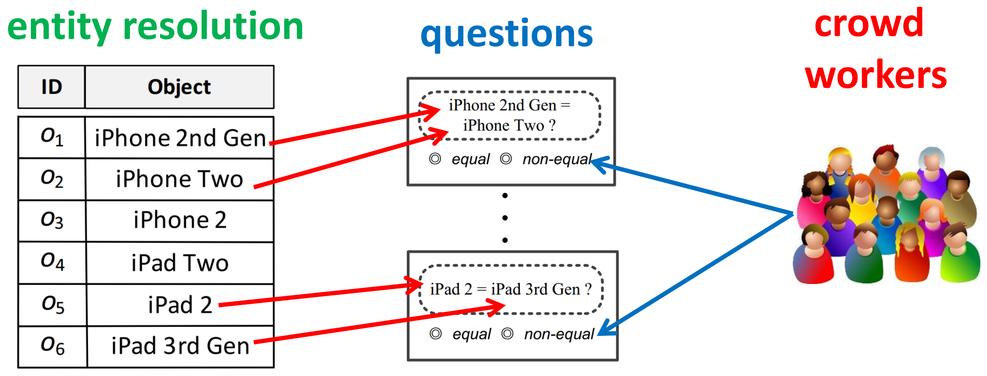
# QASCA: A Quality-Aware Task Assignment System for Crowdsourcing Applications

Yudian Zheng*, Jiannan Wang$, Guoliang Li#, Reynold Cheng*, Jianhua Feng#

#Tsinghua University, *University of Hong Kong, $UC Berkeley

## Crowdsourcing

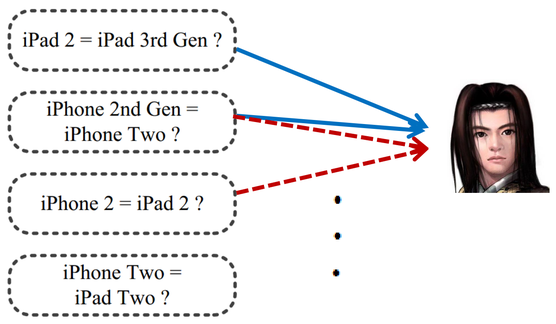
**Coordinate a crowd to answer questions that solve computer-hard applications.**




## Task Assignment

**Given *n* questions, which *k* questions should be batched in a HIT and assigned to a worker?**

**Here we have *n*=4 questions, and a HIT contains *k*=2 questions.**




## Evaluation Metric

**An application is often associated with an *Evaluation Metric.***

| Application | Sentiment Analysis | Entity Resolution |
|---|---|---|
| Question | I had to wait for six friggin' hours in line at the @apple store. ◎positive ◎neutral ◎negative | iPad 2 = iPad 3rd Gen ? ◎ equal ◎ non-equal |
| Evaluation Metric | Accuracy | F-score ( "equal" label ) |

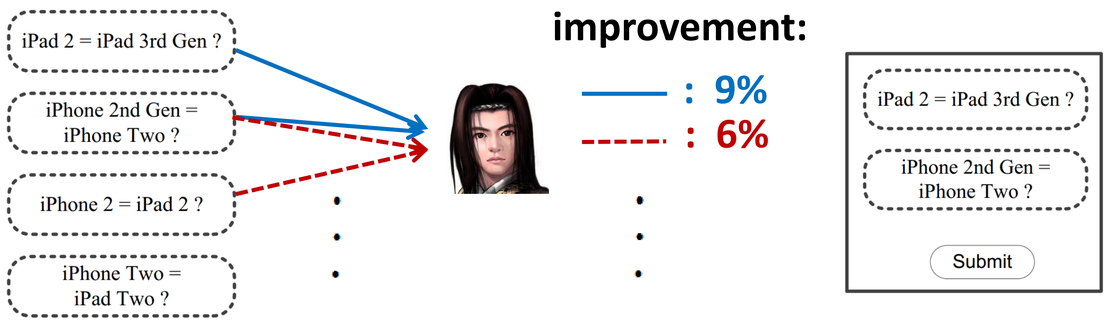
## Our Goal

**Evaluation Metric ⇨ Assignment**

**When a worker ( ) comes, given the evaluation metric,**

① **for each set of k questions, we estimate the improvement of quality, if these k questions are answered by the worker,**

② **and we select the k questions that can maximize the quality improvement for the worker.**




## Challenges & Solutions

□ **[Unknown Ground Truth] With ground truth unknown, how to evaluate the quality of returned results?**

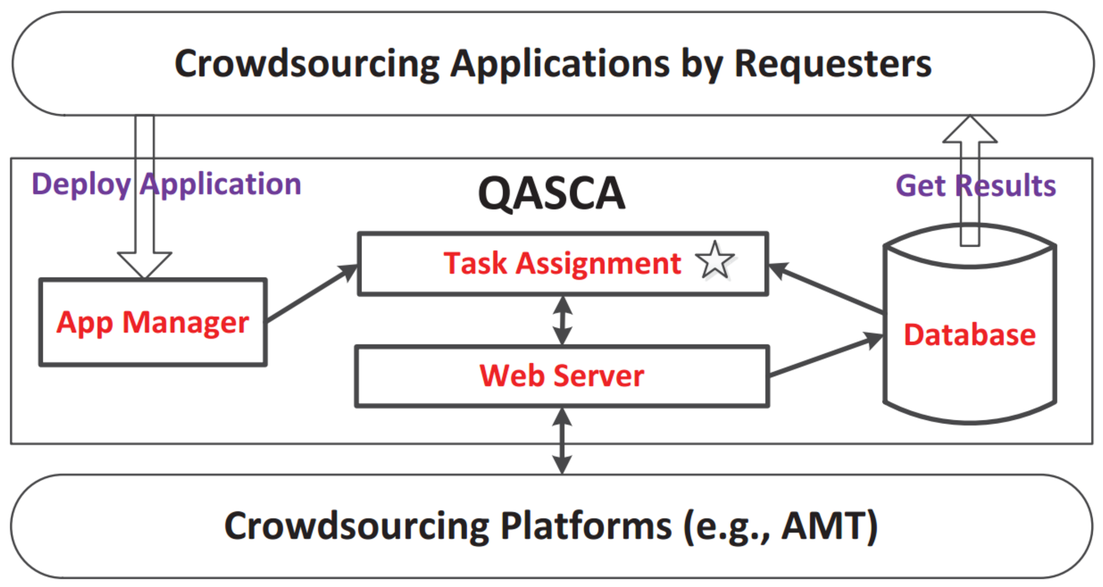
✔ ***We build a distribution matrix for the n questions' answers, and evaluate the quality of returned results by computing over the distribution matrix.***

***An Example Distribution Matrix*** $\begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$ ← ***In the first question, the probability that the first label to be the ground truth is 80%.***

□ **[Expensive Enumeration] Enumerating all possible assignments is exponential.**

✔ ***We develop linear-time algorithms to compute optimal assignments for Accuracy and F-score, respectively.***

## QASCA System Architecture

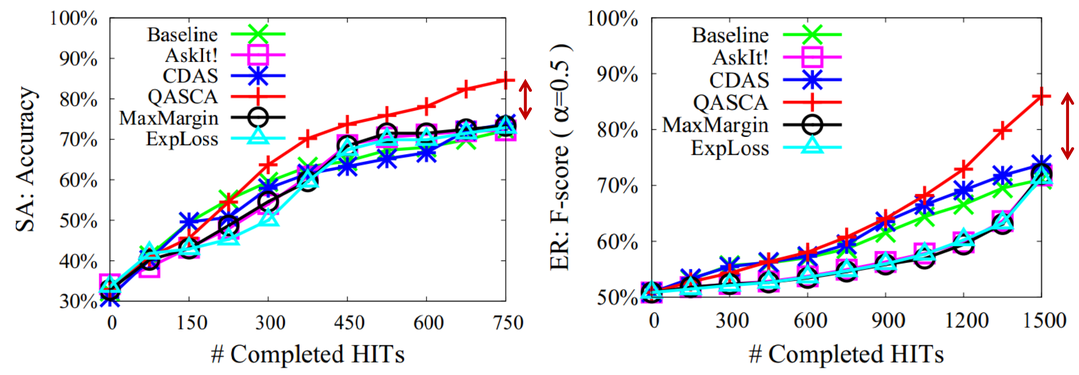



**QASCA is developed on AMT and requesters can deploy the application in the App Manager. The results will be returned to the requester after the application is finished.**

## Experiments

**We compare with five systems (Baseline, CDAS[1], AskIt![2], MaxMargin and ExpLoss) on real-world datasets.**

**Sentiment Analysis (SA)** **Entity Resolution (ER)**




*SA*: Accuracy *ER*: F-score

**QASCA outperforms more than 8% improvement in quality.**

## References

[1] X. Liu, M. Lu, B. C. Ooi, Y. Shen, S. Wu, and M. Zhang. CDAS: A crowdsourcing data analytics system. PVLDB, 5(10):1040–1051, 2012.

[2] R. Boim, O. Greenshpan, T. Milo, S. Novgorodov, N. Polyzotis, and W. C. Tan. Asking the right questions in crowd data sourcing. In ICDE, 2012.