# Crowds, not Drones: Modeling Human Factors in Interactive Crowdsourcing

Basu Roy, S. (Univ. of Washington Tacoma),
Lykourentzou, I. (CRP Henri Tudor/INRIA),
Thirumuruganathan, S. (UTA),
Amer-Yahia, S. (CNRS@LIG), Das G. (UTA),

DBCrowd2013, Riva del Garda, August 26, 2013

# Background

- Crowdsourcing
  - Gained popularity in recent years for a variety of tasks:
    - Data gathering (e.g.  Picture/video tagging)
    - Document editing (Wikipedia)
    - Opinion solicitation (e.g. restaurant ratings, sentiment analysis)
- Aims to approximate a “ground truth”
  - Objective or subjective
  - One or many

# Current Systems

- Existing crowdsourcing systems
  - Platforms: AMT, Turkit, Innocentive, CloudFlower, etc.
  - Tasks: small, independent, minor incentives, short engagement
  - Crowd: volatile, asynchronous arrival/departure, various levels of attention/accuracy
- 3 primary processes
  - Worker skill estimation (WSE)
  - Worker-to-task assignment (WTA)
  - Task accuracy evaluation (TAE)

# Limitations of current platforms (1)

- No or fragmented optimization of WSE, WTA and TAE
  - Pre-qualification tests and “golden data” optimize WSE
  - But, leave WTA up to workers
- Recent research undertakes some challenges in silo, for specific application types (e.g. real-time crowdsourcing, highly volatile crowds, single worker skill)
  - Active learning strategies for TAE improvement [Boim et. Al. 2012, Krager et. al. 2011, Ramesh et. al. 2012]
  - Worker-to-task-assignment [Ho et. al. 2012]

# Limitations of current platforms (2)

- Omission of Human Factors
  - Most approaches work with idealized human factors (e.g. known worker wages, steady worker performance). Fewer ones consider human factors
  - Human involvement →Uncertainty
    - Worker availability
    - Worker wage: deviations even among persons of the same profile, due to workload, time, unseen factors
    - Worker skill: may decline with previous workload, change with motivation

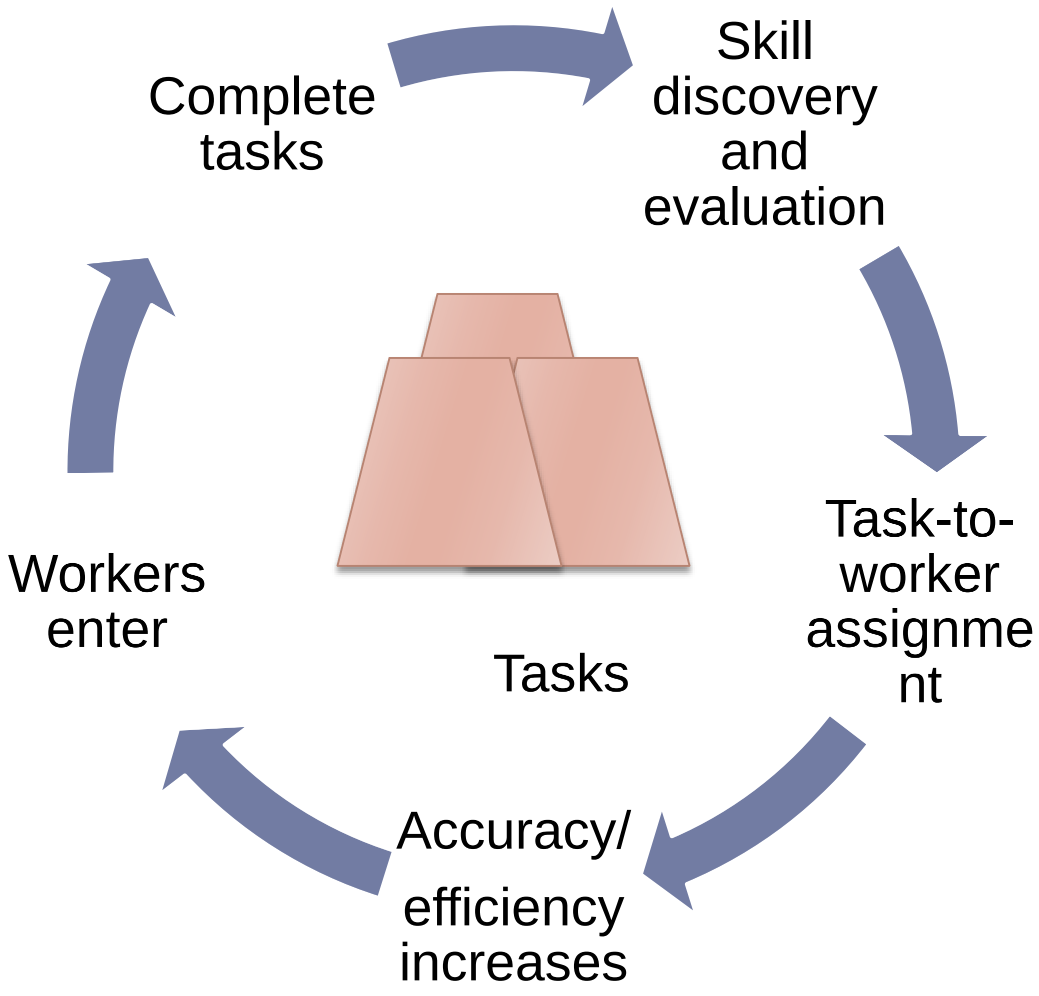
  - No existing work formalizes *the optimization problems considering Human Factors.*

# SmartCrowd

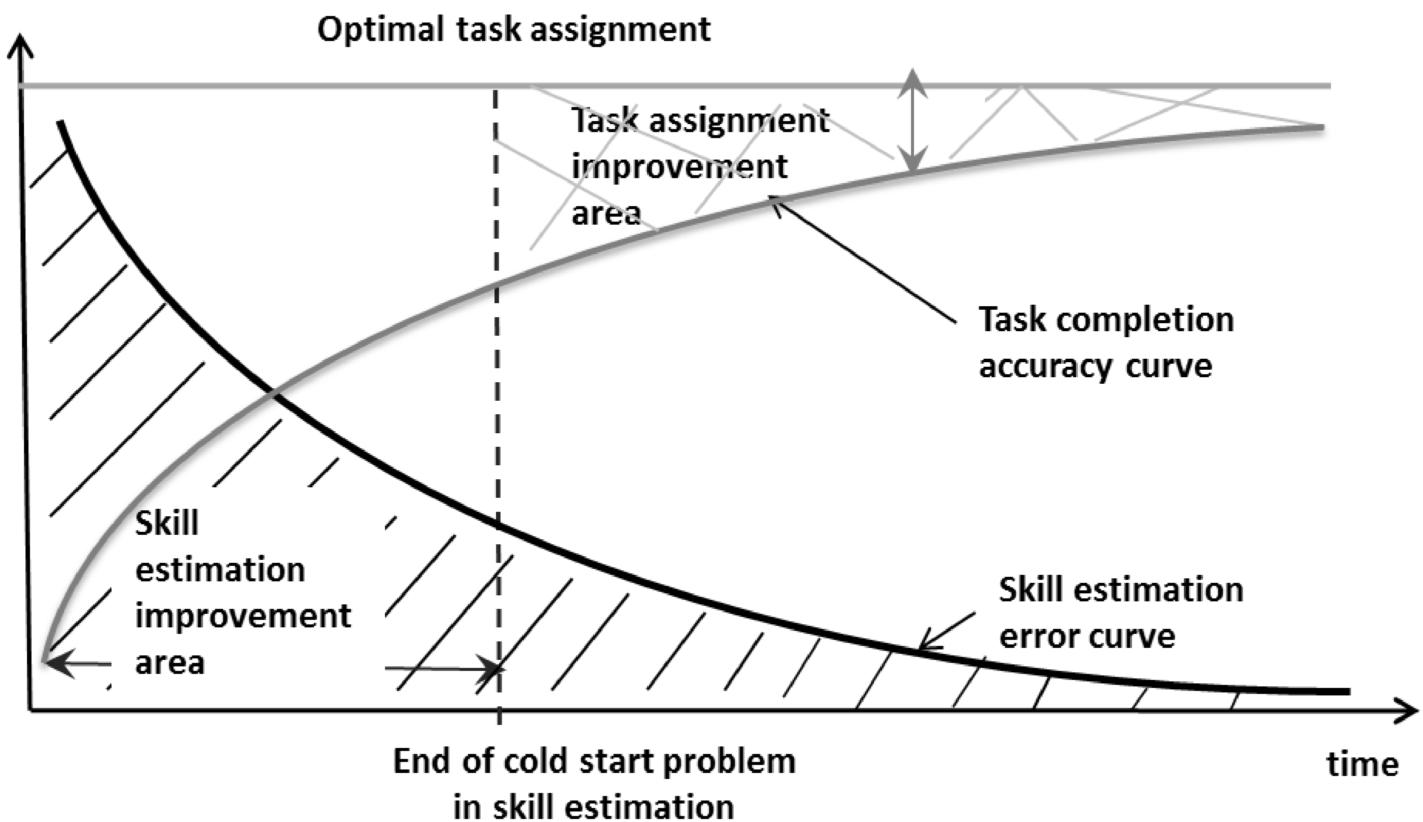
*Framework for harnessing the crowd to approximate ground truth(s) effectively and efficiently, while taking into account the innate uncertainty of human behavior, i.e. human factors*

- Adaptive, non-siloed optimization of crowdsourcing, acknowledging human factors in a dynamic environment
- Uncertainty does not preclude the design of a crowdsourcing solution with a global optimization target
  - Shifts the optimization problem from a deterministic to a probabilistic one (here probabilities and confidence boundaries need to be examined)
- Acknowledgement of multiple skills in the design
  - Ideal for Knowledge Intensive Crowdsourcing [e.g., Wiki Editing, Product design]

# SmartCrowd: High-level architecture



# SmartCrowd: Opportunities



# SmartCrowd: Challenges

- Who Evaluates What and How?
- How to Estimate Worker Skills?
- How to Assign Tasks to Workers?
- Efficient Computation

# Challenge -1: Who Evaluates What and How?

- Task assignment and evaluation are both tightly coupled with learning worker skills
- Estimation of Worker Skills
  - Fully automatic and Implicit evaluation
    - Comparison of submitted results to one another
    - Worker skills as a deviation from the so far computed ground truth
  - Explicit evaluation
    - Workers as evaluators
    - More costly, but faster skill/ground truth approximation
- Hybrid (our suggestion). Things to consider:
  - When and how to hire explicit evaluators
  - How many evaluators are needed
  - Offered incentives?
  - What to do with inconsistent attention and evaluators' arbitrary departure?

# Challenge- 2: How to Estimate Worker Skills?

- How to identify the skill set?
  - Skills may be latent (example: for knowledge building tasks, the knowledge domains necessary may not be known a priori)
  - We envision learning latent skills during task execution by workers
  - Structured learning problem with machine learning, or fixed probabilistic model to learn inference are candidate approaches (e.g., graphical models.) Can you give me some more details here? This is the only part of the paper that I was not involved, so some explanations will be very useful.)
- Challenges:
  - How to determine minimum task set needed to accurately estimate worker skills (given uncertainty in worker performance)
  - What is the “stopping condition”?
  - How to fast and incrementally compute skills as new workers arrive?
  - How often we need to re-compute given changes in human performance (e.g. boredom)?

# Challenge -3: How to Assign Tasks to Workers?

- Task to worker assignment instead of self-appointment of workers to tasks
- Probabilistic optimization problem
  - Objective with many facets: maximize accuracy, minimize cost or time, given probabilistic resource availability and performance
- Challenges
  - Dynamic allocation adjustment in case a worker declines
  - Can multiple tasks be given to a worker, and if so in what order? Can multiple workers be assigned to a task, and in which sequence?
  - System benefit vs. worker benefit tradeoff
  - Optimize across tasks or give opportunities to newcomers?

# Challenge 4: Efficient Computation

- Efficient computation is a *requirement*
  - Satisfying the key objectives of WSE, WTA, and TAE while accounting for human factors *at scale*, necessitates the development of efficient *searching* techniques.
  - **Index/View maintenance**

- Crowd Indexing : An index is an *assignment of a group of workers* to a type of task
- New forms of indexing that leverage human factors
  - Pre-computation of indexes, and efficient dynamic maintenance (e.g. as our knowledge of worker performance improves)
  - Alternate indexing, fall-back options necessary to account for uncertainty in worker arrival/departure

# Conclusion

- SmartCrowd: A framework for intelligent and dynamically optimized crowdsourcing incorporating human factors
- Both existing crowdsourcng applications(volatile crowds, small tasks) and next-generation ones (higher crowd involvement, recurring workers, collaboration, larger tasks) could benefit from our framework
- Many challenges, interesting problems lying ahead

# References

- [Boim et.  Al]  R. Boim, O. Greenshpan, T. Milo, S. Novgorodov, N. Polyzotis, and W. C. Tan. Asking the right questions in crowd data sourcing. In ICDE, pages 1261–1264, 2012.
- [Krager et. Al] D. R. Karger, S. Oh, and D. Shah. Budget-optimal task allocation for reliable crowdsourcing systems. CoRR, abs/1110.3564, 2011.
- [Ramesh et. al.] A. Ramesh, A. Parameswaran, H. Garcia-Molina, and N. Polyzotis. Identifying reliable workers swiftly. Technical report, Stanford University. 2012.
- [Ho et. al.] C.-J. Ho and J. W. Vaughan. Online task assignment in crowdsourcing markets. In AAAI, 2012.

- Thank you for your attention
- Questions?

- Contact: ioanna.lykourentzou@{tudor.lu, inria.fr}