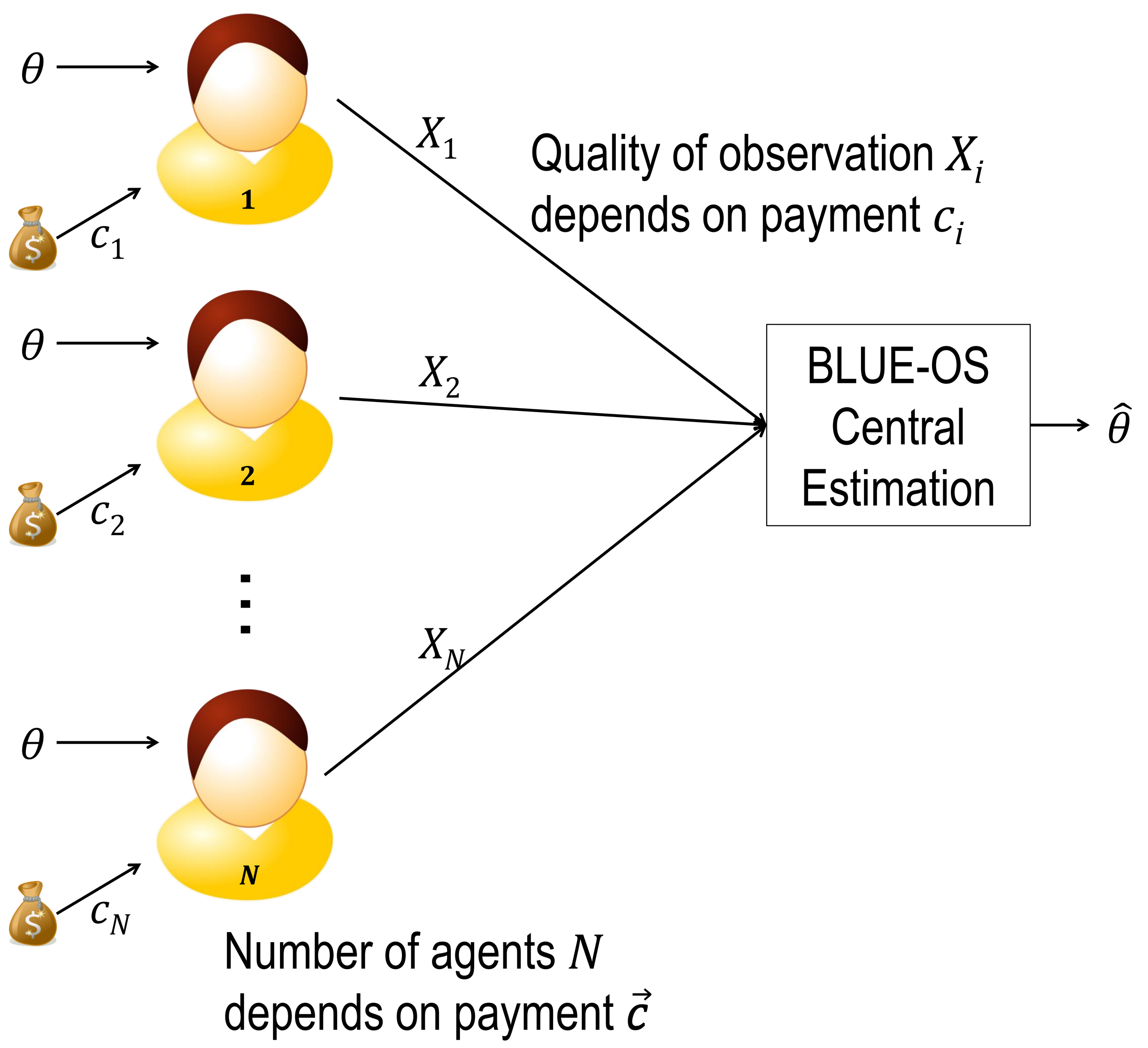
ECE ILLINOIS ILLINOIS

Distributed Estimation via Paid Crowd Work

Song Jianhan, Vei Wang Isaac Phua, and Lav R. Varshney

University of Illinois at Urbana-Champaign

Crowdsourcing and Crowdsensing as Distributed Estimation



Human workers are governed by choices and desires

Monetary payment incentivizes workers

- changes *quality* of work produced by a given worker
  - noise variance of uniform observation
- changes likelihood for work to be completed (*quantity*)
  - discrete choice models from management science

Estimate non-random parameter θ with observations from random number of variable-quality crowd workers

We consider the Best Linear Unbiased Estimator with Order Statistics (BLUE-OS) for all settings (covers mean, median, etc.)

Random Number of Equal-Quality Workers

Theorem The best unbiased estimator of the location parameter from an i.i.d. sequence of uniform observations of fixed or random length is the sample midrange.
Proof Extends (Neyman and Pearson, 1928) using linearity.

- Find best payment allocation for MSE under the multinomial logit model of whether workers will complete task or not

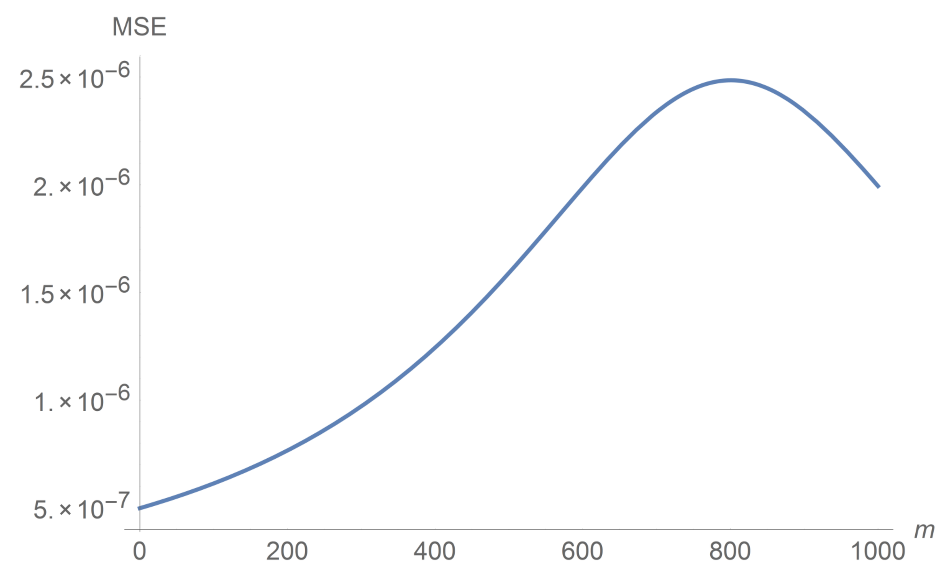
Theorem The best payment allocation is homogeneous.
Proof Perturbation argument with midrange MSE expression.

Fixed Number of Variable-Quality Workers

Two levels of workers: *expert* and *amateur* with differing observation noise

In joint population model, we do not know which workers are experts and which are amateurs. Since variates are not identically distributed, finding BLUE-OS and its MSE performance requires detailed order statistics computations involving matrix permanents.

Let n be total number of workers and m be number of amateurs among them



Theorem Unintuitively, the MSE of the BLUE-OS estimator does not always improve with the fraction of experts, for sufficiently large n.

When experts and amateurs can be distinguished from one another, the budget allocation problem reduces to distributing money among classes.

Theorem Consider k independent worker pools P_1 to P_k, with N_i observations from P_i, and corresponding BLUE-OSs e_i. The overall system BLUE-OS is a linear combination of the e_i (linear combination of midranges).

Optimize Quality and Quantity

Consider optimal budget allocation when considering both quality and quantity, when experts are distinguishable from amateurs.

Algorithm

1. Choose how to allocate budget for the two classes, B_A and B_B.
2. For this allocation, plot the expected numbers of workers and corresponding MSE of BLUE-OS: $(N_A; N_B; \mathrm{mse})$ to find the first minimum with reference to MSE.
3. Repeat 1-2 and get the allocation with the lowest MSE.

Conclusion

Distributed estimation in sensor networks is limited by noise and bandwidth, but crowdsourcing is limited by human expertise and motivation.

Solving novel resource allocation problems can improve an emerging paradigm for sensing and information processing.

The 41st IEEE International Conference on Acoustics, Speech, and Signal Processing, 20–25 March 2016, Shanghai, China