Human Computation
AAAI Technical Report WS-12-08

To Crowdsource Or Not to Crowdsource?

Gireeja Ranade
Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, California 94720

Lav R. Varshney
IBM Thomas J. Watson Research Center
Hawthorne, New York 10532

Abstract

Crowdsourcing contests—events to solicit solutions to problems via an open-call format for prizes—have gained ground as a mechanism for organizations to accomplish tasks. This paper uses game-theoretic models to develop design principles for crowdsourcing contests and answer the questions: what types of tasks should be crowdsourced? Under what circumstances? When a single task is to be completed, crowdsourcing can lead to higher quality outcomes than directed assignment if the pool of players is diverse, but can lead to suboptimal outcomes when workers have similar abilities. With multiple tasks, crowdsourcing can easily match players with diverse skill-sets to different tasks to achieve high aggregate performance. However, surprisingly, crowdsourcing is not always useful to find expert workers for highly specialized tasks.

Introduction

Information technologies that provide lightning speed communication at low cost are changing the nature of work. Organizations can now leverage networks, communities, and ecosystems of people to perform tasks. Workforces are no longer localized and homogenous, but instead are globally distributed and diverse. Large projects are broken up into smaller encapsulated pieces. In fact, the millennial generation shows a cultural preference for project-based rather than jobs-based work (Bollier 2011). Within this environment, methods of collective intelligence have emerged as key business tools (Malone, Laubacher, and Dellarocas 2010). A fundamental understanding of this evolution of work is essential to shape its future form.

A notable example of decentralized organization is crowdsourcing. Crowd-power has been harnessed to design everything from t-shirts to software to artificial intelligence algorithms by soliciting contributions via open calls (Tapscott and Williams 2006; Bokelberg and Varshney 2012; Boudreau, Lacetera, and Lakhani 2011). The ability to reach a large crowd of skilled workers quickly and inexpensively gives firms an alternative means for accomplishing tasks. As such, it is important to understand the pros and cons of a fluid, crowd-based labor force. To quote Thomas Malone, "There is this misconception that you can sprinkle crowd wisdom on something and things will turn out for the best. That's not true. It's not magic." (Lohr 2009).

How can an organization find the best person for a job and incentivize him to attain peak performance? Crowdsourcing has been thought of as a good tool for skill identification. The aim of this paper is to identify the utility of crowdsourcing contests as a business tool. The focus is crowdsourcing contests, where monetary or otherwise tangible rewards are provided to the winners of competitive events. This is different from other forms of crowdsourcing where crowd workers do not receive direct extrinsic rewards, such as in Wikipedia (Benkler 2006; Howe 2008) and also different from paid microtask forums, such as Amazon's Mechanical Turk, where there is no competition. Contests may be internal, with competition only among the employees of the organization (internal crowdsourcing, e.g. IBM's Liquid platform (Bokelberg and Varshney 2012)), or external and open to the public (external crowdsourcing, e.g. TopCoder). Our results apply to the last two types of crowdsourcing.

One approach to answer the question raised above is empirical, as in (Liu et al. 2011). Contrarily, we set up a theoretical framework to answer this question, taking a first step towards determining the optimal work structure for a given task, be it labor-based, division into microtasks, a crowdsourcing contest (internal or external), or something else entirely.

We use explicit game-theoretic models for crowdsourcing single and multiple tasks, which build on previous all-pay auction models (DiPalantino and Vojnović 2009; Archak and Sundararajan 2009; Varshney et al. 2011; Liu et al. 2011). First, we provide a mathematically-oriented taxonomy for tasks based on their optimal solution method. This classification gives specific conditions for when crowdsourcing can generate revenue for the employer for a particular task.

To address the point about skill identification, we also demonstrate conditions under which crowdsourcing contests yield higher returns to the employer than hierarchical managerial assignment. We find crowd-based self-assignment is better for single tasks when

- Managers are uncertain about the resources and skills of workers,

- The pool of workers is diverse in its ability to perform the given task, and
- Workers have low default base effort.

Contrary to the claims of Malone et al. (2010), the first condition may not be sufficient. Further, when organizations may have multiple different tasks and diverse workers, as in crowdsourcing platforms like TopCoder or IBM's Liquid platform, we find:

- for tasks that do not require highly specialized skills, crowdsourcing can perform as well as optimal assignment of workers to tasks;
- but, crowdsourcing contests can be suboptimal even with a pool of many strong players and just one weak player; and
- crowdsourcing contests can perform badly for highly specialized tasks. Instead of pulling out highly-skilled workers from a crowd, crowdsourcing could lead to mediocre performance by everyone.

Malone et al. (2010) raise the question of how to invoke collective intelligence to perform a task: via collection, contest, or collaboration? Within this framework, contests can be thought of as a subtype of collection, where certain contributions are awarded prizes. For collection (e.g. microtasks or contests) to be the appropriate choice, they say that it must be possible to divide the overall activity into small pieces that can be done independently by different members of the crowd. They go on to assert that contests are a useful way of doing collection when only one or a few good solutions are needed. In contrast to this claim, our final multiple-task model demonstrates that under certain conditions, tasks where many solutions are useful to the contest-designer are more suitable to be solved using crowdsourcing contests than tasks where only a few good solutions are needed. This turns out to be true because when a large number of qualified players are available, multiple tasks will most effectively mobilize them.

Crowdsourcing Contests as All-Pay Auctions

Crowdsourcing contests involve multiple participants with a diversity of skills and incentives. We develop simple models based on standard game theory to help understand and design contests, building on previous work on all-pay auction models (DiPalantino and Vojnović 2009; Archak and Sundararajan 2009; Varshney et al. 2011; Liu et al. 2011). Players in all-pay auctions forfeit their bids regardless of whether they win. Similarly, a participant in a crowdsourcing contest pays (via effort, opportunity cost etc.) regardless of whether his entry is chosen. For the remainder of the paper, we follow the all-pay auction with complete information model of crowdsourcing contests and we use the terms contestants, players, workers and bidders interchangeably.

To understand how to find the best worker for a given task and incentivize him to do his best, we compare the efficacy of targeted assignment to that of crowdsourcing contests. The specific map of skill level to players becomes very important for targeted assignment of tasks — the manager or assigner must have at least some estimate of all these values. On the other hand, auctions with complete information only assume that each player knows the multiset of the strengths of other players, which is much less information than the exact mapped sequence of strengths (Varshney and Goyal 2006). Crowdsourcing forums such as TopCoder make heavy use of leader boards and public participation records. These digital reputations and public rankings provide players good estimates of the strengths of other players in a contest. Furthermore, the more experience a player has, the better he will be at such estimation.

Empirical analysis of participation on Taskcn that reveals high-quality (experienced) players have the most influence on the contest-designer's revenue and drive aggregate reward and reserve-price effects on submission quality (Liu et al. 2011). They find an early high-quality submission deters other high-quality players. Clearly, understanding the behavior of the strongest players is key.

The models used here build on these ideas where all players have distinct, public costs and we show weaker players are deterred from entering when they know strong players are already participating. We limit attention to auction models with complete information to distill the salient aspects of this comparison. We assume all players know the skill levels of all other players. High skills implies low costs per unit effort.

The Model

Consider a crowdsourcing contest with n players (P_1 to P_n) and prize value $A > 0$. To capture the idea that players may have different skill sets and abilities to perform the given task, we introduce costs per unit effort $c_i,\ 1 \leq i \leq n$, for each player. For instance, an expert may only need a few hours to complete a task and would have a low cost of effort, whereas a novice might have much higher cost.

Each player submits a bid x_i that represents the quality of her submission, at cost $c_i x_i$. The prize is awarded to the best submission, i.e. the highest bid x_i. In this paper, we exclusively consider a complete information setup, where the prize value A and players costs c_i are publicly known. Player bidding strategies depend on the other player costs, but only as a multiset. The specific mapping of costs to players is irrelevant in determining bidding strategies.

In a two-player contest with players P_1 and P_2, consider the case where $c_1 < c_2$, i.e. P_1 is the stronger player. Then the expected utilities of the players, $\mathbb{E}[U_1],\ \mathbb{E}[U_2]$, respectively for bids of x_1, x_2, are

$$\mathbb{E}[U_1] = A \cdot P[x_1 > x_2] - x_1 c_1 \quad (1)$$

$$\mathbb{E}[U_2] = A \cdot P[x_2 > x_1] - x_2 c_2 \quad (2)$$

Theorem 1 (Hillman and Riley, (1989)). *The two player contest described above admits a unique Nash equilibrium. At equilibrium P_1 bids uniformly at random on $[0, \frac{A}{c_2}]$. P_2 bids 0 with probability $\frac{c_2 - c_1}{c_2}$, and with the remaining probability, $\frac{c_1}{c_2}$, bids uniformly on $[0, \frac{A}{c_2}]$ as well.*

Theorem 2 (Hillman and Riley, (1989)). *If $n \geq 3$ players are involved with strictly increasing costs for P_1, P_2, P_3, i.e. $(c_1 < c_2 < c_3 \leq c_4 \ldots \leq c_n)$, and if the strongest players*

P_1 and P_2 act as if there were no other agents, then P_3 to P_n do not enter the contest and submit bids 0 with probability 1.

Thm. 2 demonstrates a mismatch between our model and empirical characterizations of crowdsourcing contests, where more than two players tend to participate. The empirical phenomenon may arise because players are not rational, because players underestimate their competition, or because player skills are indeed exactly equal and so it is rational for many to enter. Baye et al. (1996) consider an all-pay auction with n players with a set of $n \geq 3$ players of identical skill level. This case shows many asymmetric Nash equilibria, with more than two players entering the competition. For simplicity, here we consider the model with a unique Nash equilibrium, with distinct player costs. Ideas and insights from our model can be extended to the model of Baye et al. (1996).

The following section builds on the all-pay auction framework here and to categorize tasks.

The Contest Designer's Perspective

An all-pay auction traditionally assumes that all submitted bids serve as revenue to the auctioneer. However, in a crowdsourcing contest, this might not be the case. Some events may have synergy while others have redundancy across entries (Bettencourt 2009). The utility that the contest-designer derives from a task depends on its nature: we denote contest returns by the function f. Depending on f, the designer may want to change the parameters of the contest or decide whether it is even worthwhile to hold an event. The utility derived by the contest-designer is

$$U_{task} = f(\vec{x}) - A \tag{3}$$

where $\vec{x}$ represents the vector of n bids $x_1, x_2, ..., x_n$.

The function f can provide a mathematically-oriented classification of potential crowdsourcing tasks. Tasks may be:

- *selective*, e.g. software component design,
- *integrative*, e.g. information aggregation or idea generation, or
- involve *market creation*, e.g. the X PRIZE.

We carry the terms selective and integrative from (Schenk and Guittard 2011), which is one among many recent taxonomies for crowdsourcing. Contests that derive utility from the top-k entries interpolate between the extremes of selective and integrative tasks (Archak and Sundararajan 2009).

In a *selective* task, only one working solution is useful to the designer. In this case, the designer typically derives utility from the maximum, and her utility function would be

$$f(\vec{x}) = \max(x_1, x_2, \ldots, x_n) \tag{4}$$

On the other hand, an *integrative* idea generation contest might provide an additive or even superadditive utility to the designer and have eq. (5).

$$f(\vec{x}) = \alpha \sum_{i=1}^{n} x_i, \alpha > 0 \tag{5}$$

The log of the weighted sum of bids in the utility function above, $f(\vec{x}) = \alpha \log\left(\sum_{i=1}^{n} x_i\right)$, captures diminishing marginal returns.

Tasks might also be subject to a coordination cost per player, $(\gamma > 0)$, which scales with the number of players and thus decreases the utility of the contest designer as in eq. (6).

$$f(\vec{x}) = \alpha \sum_{i=1}^{n} x_i - \gamma n, \alpha > 0 \tag{6}$$

Modeling *market creation* though a function is more challenging. As noted by the X PRIZE Foundation, their goal is "about launching new industries that attract capital, that get the public excited, that create new markets" (Tapscott and Williams 2010, p. 131). Thus, independent of the quality of bids, the sheer number of entries might provide utility with

$$f(\vec{x}) = \alpha n + \beta \tag{7}$$

where n is the number of players.

One may further desire the f function to be upper-bounded by some maximum value.

Here, we use f to characterize which tasks are best suited to crowdsourcing contests. As a minimum requirement, we would like to ensure that the contest-designer's utility $\mathbb{E}[U_{task}]$ is positive, so no losses are incurred by running a contest. This idea extends to ensure against some minimum profit.

We consider four examples below. More than just mathematical derivation within the model, our point is to show that the parameters of the player pool influence how a particular task should be solved.

Example 1. In a two-player selective winner-take-all contest, eq. (4), the expected utility under equilibrium is,

$$\mathbb{E}[f(\vec{x}) - A] = \mathbb{E}[U_{task}] = \frac{A}{6}\left(\frac{3c_2 + c_1}{c_2^2}\right) - A, \tag{8}$$

which follows from the randomized strategies of the players discussed in Thm. 1. The calculations are omitted here. Thus, $\mathbb{E}[U_{task}]$ is positive if and only if $\frac{3c_2 - c_1}{6c_2^2} - 1 > 0$. The player strengths c_1 and c_2 determine whether the contest-designers utility is positive. If $c_2 \gg c_1$, i.e. the second player is much weaker than the first, then the condition reduces to $c_2 < \frac{1}{2}$. On the other hand if $3c_2 - c_1 = \epsilon$ is small, then $c_2^2 < \frac{\epsilon}{6}$ is a strong enough condition to ensure positive utility.

Example 2. For an integrative task with superadditive f as in eq. (5), even a weak player pool can provide positive utility, as below:

$$\mathbb{E}[U_{task}] = \alpha\left(\frac{A}{2c_2} + \frac{c_1 A}{2c_2^2}\right) - A \tag{9}$$

Therefore $\mathbb{E}[U_{task}] > 0$ if and only if $\alpha > \frac{2c_2^2}{c_1 + c_2}$. If $c_2 \gg c_1$, this reduces to $\frac{\alpha}{2} > c_2$, while if $c_2 - c_1 = \epsilon$ is small, then a sufficient condition for positive $\mathbb{E}[U_{task}]$ is given by $\alpha > c_2$, since $\frac{\alpha}{2} > \frac{c_2^2}{2c_2} + \epsilon > \frac{c_2}{2}$. In the case $\alpha = 1$, we see that the positive utility conditions for integrative tasks are weaker than those for selective tasks.

Example 3. Consider n players with strictly increasing costs competing for the task, as in eq. (6). We noted earlier that in this case $x_3, ..., x_n = 0$ with probability 1 if the strongest players P_1 and P_2 ignore the presence of the other players, Hence, $\mathbb{E}[U_{task}]$ for the contest-designer is exactly as eq. (8). Clearly, having a contest to complete just one task when a large pool of players are available, and where the coordination costs scale with the number of players, as in eq. (6), a winner-take-all contest would be exactly the wrong structure. Dividing the task into smaller pieces (if possible) would be advantageous to the designer in this case.

Example 4. Finally, in the case of an event where the objective is market creation, the designer's utility does not even depend on the effort of the players.

Remark 1. In closing, we observe that if $f_1(\vec{x}) \geq f_2(\vec{x})$ for all $\vec{x}$, $\mathbb{E}[U_{task}](f_2) > 0$ implies that $\mathbb{E}[U_{task}](f_1) > 0$, since player bidding strategies are independent of the designer's valuation function f. Even with approximate functions, this provides contest-designers a simple rule of thumb to order potential f functions.

Comparing Crowdsourcing Contests to Targeted Assignment

Crowdsourcing can be a powerful way to reach a large pool of players inexpensively. The right crowdsourcing model might increase fluidity in the labor force of the future (Bollier 2011). However, it is important not to think of crowdsourcing as a catch-all solution.

The fluidity of the crowdsourcing model allows a player to self-select tasks. With this agency, players (or employees) will likely choose tasks that they are good at and enjoy, while, as noted earlier, managerial assignment requires more detailed information. How useful is crowdsourcing for skill identification?

The winner-take-all format provides natural performance incentives without the cost of organizational frameworks. Quite clearly, such a model comes with the benefits and pitfalls of an outcomes-based reward structure (e.g. (Dixit 2002)). How important are these competitive incentive factors?

The single task model addresses the tradeoff between endogenous incentives offered by a contest-based solution versus externally provided incentives, and the effect of observation noise when a manager assigns a task to a player for a selective task with two players. Similar models can be easily developed for integrative or market creation tasks.

The second model looks at a multi-task, multi-player setting and captures the potential of crowdsourcing contests to solve hard matching problems and to yield higher utility for both designers and players.

Of course, completion of the task at hand with maximum utility may be only one among many objectives desired by the contest-designer. A longer-term perspective may necessitate considering completion of many (similar or dissimilar) tasks over time, which may require workforce evolution and retention. Further, it has been observed empirically that crowdsourcing contests may be inefficient, since they are prone to overeffort; in aggregate, players may exert more than fifteen times the effort required to complete the task (Varshney et al. 2011). Direct assignment of tasks to players minimizes such overeffort and can offer training benefits. These issues are not addressed in the preliminary models here.

Crowdsourcing a Single Task

Different tasks require different skills on the part of the player, and different players have different skills. For instance, one software development project may require knowledge of Java, whereas another might require DB2. We represent these skills for both players and tasks with k-length binary vectors similar to Hong and Page (2001).

We model imperfect skill information on the part of the manager as a noisy observation—the manager observes each skill vector with a bit flip probability σ. Based on this observation, the manager assigns the closest player (in Hamming distance) to a task. Targeted assignment with high noise is like random assignment.

Let $d_i,\ i = 1, 2, d_1 < d_2$ represent the distances of players 1 and 2 from the task, which serve as a proxy for the costs incurred for the players. Let $\vec{s}_1$ and $\vec{s}_2$ represent the length k skill vectors of the two players, and $\vec{z}_1,\ \vec{z}_2$ be the two independent noise vectors $\sim$ Bernoulli(σ). The manager observes $\vec{s}_1 + \vec{z}_1$ and $\vec{s}_2 + \vec{z}_2$, and which are at distances e_1 and e_2 from the task.

First, let us calculate the utility achieved with noiseless perfect assignment. The stronger player, player 1, will at best exert $\frac{A}{d_1}$ effort to complete the task. In the absence of competition to win the reward as an incentive, managers must provide an external framework to motivate workers — these might include long term benefits, reputation, future job prospects, promotions etc. Let θ be the base fraction of effort exerted by players through such external (non-competition based) incentives. θ is an empirical factor and does not impact the utility derived from crowdsourcing. Let v be the base utility obtained by the contest-designer. Noiseless optimal assignment gives the expected utility:

$$\mathbb{E}[U_{opt}] = v + \theta\frac{A}{d_1} - A = v + A(\frac{\theta}{d_1} - 1). \tag{10}$$

Next, we look at what happens with noisy observations and compare utility from such assignment to that from crowdsourcing contests.

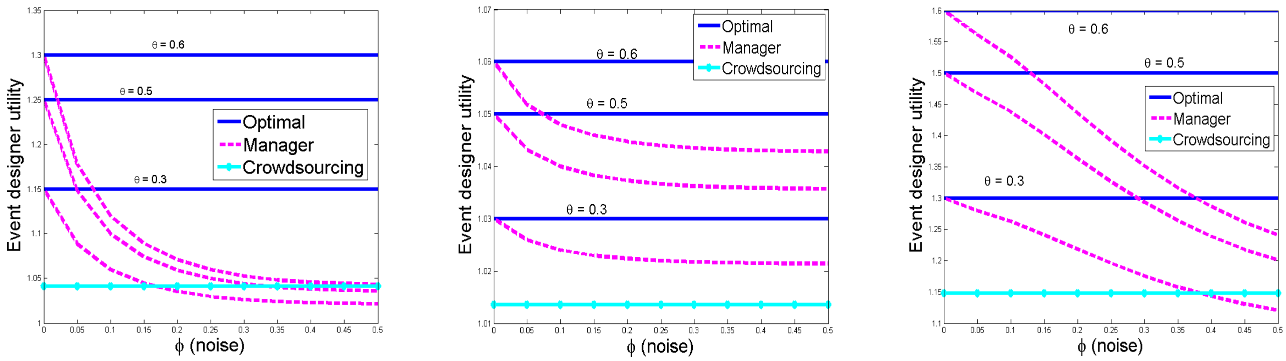
Theorem 3. *The expected utility achieved by targeted assignment and crowdsourcing contest mechanisms are given as*

$$\mathbb{E}[U_{man}] = v + A[P(r)\frac{\theta}{d_1} + (1 - P(r))\frac{\theta}{d_2} - 1] \tag{11}$$

$$\mathbb{E}[U_{cs}] = v + \frac{A}{6}\left(\frac{3d_2 + d_1}{d_2^2}\right) - A \tag{12}$$

respectively, where $P(r)$ is the probability that the task is assigned to the correct player, and can be calculated as below.

Proof: It is a straightforward expectation that gives $\mathbb{E}[U_{man}]$ as eq. (11), once we have $P(r)$.



(a) One strong and one weak player (b) Two weak players (c) Two strong players

Figure 1: Crowdsourcing contests offer the greatest advantage if the players are diverse in their ability to perform the task — it is easier to identify the stronger player than through noisy observations.

To calculate $P(r)$ we note that an error will be made by the manager if $e_1 > e_2$ even though $d_1 < d_2$, i.e. $P(r) = P(e_1 < e_2 | d_1 < d_2)$. Hence, we are interested in the change in $e_1 - e_2$ compared to $d_1 - d_2$. This can equivalently be modeled as $\vec{s}_1$ being perturbed by noise $\vec{z} \sim$ Bernoulli(ϕ), where $\phi = 2\sigma - 2\sigma^2$, with $\vec{s}_2$ unchanged. Let d_{ch} be the change in $\vec{s}_1$ due to the noise $\vec{z}$. Then,

$$P(r) = \sum_{l=0}^{d_1 - d_2} P(d_{ch} = l) \tag{13}$$

$$P(d_{ch} = l) = \sum_{k=0}^{k_1} \binom{n - d_1}{l + k} \binom{d_1}{k} \phi^{l+2k} (1 - \phi)^{n-l-2k} \tag{14}$$

where $k_1 = \min(d_1, n - d_1 - l)$.

Since a crowdsourcing mechanism picks out the maximum bidder, we can calculate the distribution of the expected utility under equilibrium as the $\mathbb{E}[\max(x_1, x_2)]$, where bids x_1 and x_2 have distributions as specified by Thm. 1. This is similar to the calculation in Example 1, and gives eq. (12).

■

Figs. 1(a)-(c) help interpret Thm. 3 by showing the relative utility obtained by the task-designer with different values of the managerial noise and base level of effort. Productivity estimates often show that θ might be about 0.5. The relative strengths of the two players d_1, d_2 can affect whether crowdsourcing or managerial assignment provide higher returns. Note these figures serve as examples and changes to θ, ϕ, d_1, d_2 may affect the nature of the plots.

As Fig. 1(a) shows, crowdsourcing offers the greatest advantage over assignment when the skill levels in the pool of players is diverse. We see that with a weak pool of players targeted assignment performs better than crowdsourcing (Fig. 1(b)). With two strong players, noise does not matter much, and crowdsourcing does not offer significant advantages to identify skill (Fig. 1(c)).

Crowdsourcing Multiple Tasks

Matching multiple tasks and multiple potential solvers is a complex assignment problem. It is tempting to believe that crowdsourcing mechanisms could provide a natural scheme through which players will self-select appropriate tasks. Not only would players choose suitable tasks, but the competition could provide performance incentives. This section explores a simple multi-player multi-task model. We find that

- for generic tasks with many skilled players, crowdsourcing can easily achieve close to optimal utility for a task;
- surprisingly, crowdsourcing contests can provide utility close to zero even with a pool of many strong players and just one weak player; and
- crowdsourcing contests perform badly for highly specialized tasks, which only few experts might be qualified to perform. In fact, these experts might get buried under the crowd.

We introduce some new notation to setup the framework. Consider a setting with $2n$ tasks $(E_J(1), \cdots, E_J(n))$ and $(E_G(1), \cdots, E_G(n))$ and $2m$ players $(P_J(1), \cdots, P_J(m))$ and $(P_G(1), \cdots, P_G(m))$, with $m \geq n$. There are n tasks of each of two types: Java-based (type J) or graphic design-based (type G), and similarly, m players of each type.

Players of a given type are better at tasks of the same type. A Java programmer (graphic artist) can complete a Java task (or graphic-design task) at a low cost (c_l), whereas she has a high cost (c_h) for a graphic-design task (or Java-task). $P_J(i)$ has cost $c_{JJ}(i) = c_l + \epsilon(i)$ for tasks $E_J(1)$ to $E_J(n)$, and cost $c_{JG}(i) = c_h + \epsilon(i)$ for $E_G(1)$ to $E_G(n)$. Costs for $P_G(i)$, i.e. $c_{GG}(i)$ and $c_{GJ}(i)$ are defined similarly. $\epsilon(i)$ is to be thought of as an error term: $\epsilon(i) \ll c_l \ll c_h$. Base effort θ is as defined earlier. Without loss of generality, assume that the player costs are ordered: $c_{JJ}(1) < c_{JJ}(2) < \cdots < c_{JJ}(m) \ll c_{GJ}(1) < c_{GJ}(2) < \cdots < c_{GJ}(m)$, and $c_{GG}(1) < c_{GG}(2) < \cdots < c_{GG}(m) \ll c_{JG}(1) < c_{JG}(2) < \cdots < c_{JG}(m)$.

The optimal utility, U_{opt}, is achieved when each player is matched to a task of his or her type. This total utility is the sum of utilities the events less the reward paid out, and acts as our baseline.

$$\mathbb{E}[U_{opt}] = \sum_{i=1}^{n} \theta \frac{A}{c_{JJ}(i)} + \sum_{i=1}^{n} \theta \frac{A}{c_{GG}(i)} - 2nA \tag{15}$$

$$\approx 2nA(\frac{\theta}{c_l} - 1). \tag{16}$$

However, manually doing this matching is difficult.

Theorem 4. *In the framework described above, if the manager observes an incorrect player type with probability ϕ, then the expected contest-designer utility, U_{man}, is*

$$\mathbb{E}[U_{man}] = 2nA\left(\frac{(1-\phi)\theta}{c_l} + \frac{\phi \cdot \theta}{c_h} - 1\right) \tag{17}$$

Proof: An expectation calculation gives:

$$\begin{aligned}\mathbb{E}[U_{man}] =&(1-\phi)\sum_{i=1}^{n} \theta \frac{A}{c_{JJ}} + \phi \sum_{i=1}^{n} \theta \frac{A}{c_{JG}} + \\ &(1-\phi)\sum_{i=1}^{n} \theta \frac{A}{c_{GG}} + \phi \sum_{i=1}^{n} \theta \frac{A}{c_{GJ}} - 2nA.\end{aligned} \tag{18}$$

Substituting values of the costs gives the desired eq. (17). ■

Now consider the crowdsourcing scenario. Each player can submit entries for any task, however, finally only one player will be picked per task and each player can only win one contest.

Theorem 5. *In a crowdsourcing contest as described above, when $m \geq n+1$, the expected utility from crowdsourcing is*

$$\mathbb{E}[U_{cs}] = 2nA(\frac{1}{c_l} - 1), \tag{19}$$

as $n \to \infty$ when $c_{JJ}(i) \approx c_{JJ}(k), c_{GG}(i) = c_{GG}(k)$ for all $1 \leq i, k \leq n$.

Proof: First, consider all n tasks of type J. For notational ease, let $v_{JJ}(i) = \frac{A}{c_{JJ}(i)}$. Thus, we are looking at an all-pay auction, with n identical goods $E_1, ..., E_n$, and $2m$ players. The n highest bidders will be assigned the n tasks[1].

When all the players have unequal valuations, this game has a unique Nash equilibrium, in which only the strongest $n+1$ players actively bid, while the rest almost surely bid 0. Thus, only players $P_J(1)$ to $P_J(n+1)$ with values $v_{JJ}(1)$ to $v_{JJ}(n+1)$ will actively submit entries for tasks of type J. From Clark and Riis (1998), we know that $P_J(1)$ to $P_J(n)$ will randomize over the interval $[\ell(i), v_{JJ}(n+1)]$, where

$$\ell(i) = \left(1 - \prod_{k=i}^{n} \frac{v_{JJ}(k)}{v_{JJ}(i)}\right) v_{JJ}(n+1),\ i = 1, 2, ...n. \tag{20}$$

Let r be a parameter such that $r = 1$ if $\ell(1) \leq x \leq v_{JJ}(n+1)$, else $r = s$ if $\ell(s) \leq x < \ell(s-1)$. Then, the distribution of the bidding strategy (Clark and Riis 1998), $F_{J,i}(x)$, for $i = 1, 2, ...n$ is given by

$$F_{J,i}(x) = 1 - \left(\frac{v_{JJ}(i)}{\prod_{k=r}^{n} v_{JJ}(k)^{\frac{1}{n+1-r}}}\right)\left(1 - \frac{x}{v_{JJ}(n+1)}\right)^{\frac{1}{n+1-r}} \tag{21}$$

Player $P_J(n+1)$ submits 0 with probability $\left(1 - \frac{v_{JJ}(n+1)}{v_{JJ}(n)}\right)$, and otherwise randomizes according to eq. (21) with $v_{JJ}(n)$ in place of $v_{JJ}(i)$.

The expected utility of each player in this case is given by $v_{JJ}(i) - v_{JJ}(n+1)$. The expected payoff to the contest-designer would be the sum of the n highest bids of the players.

As n becomes large, players submit entries close to the upper bound $v_{JJ}(n+1)$. Similar to the two-player all-pay auction where no players bid higher than the weaker player's valuation, the weakest players valuation $v_{JJ}(n+1)$ is an upper bound on the bids of all players. Note that if $v_{JJ}(i) \approx v_{JJ}(k)$ for all $i,\ k$ then the lower bound for the support of the mixed strategies of all players is close to 0. $r = 1$ for most x in this interval in this case and

$$F_{J,i}(x) \approx 1 - \left(1 - \frac{x}{v_{JJ}(n+1)}\right)^{\frac{1}{n}}. \tag{22}$$

Since all players adopt the same strategy, the probability of any one player winning is $\frac{n}{n+1}$. Thus, the expected bid is the difference of the expected gross surplus and the expected utility,

$$\mathbb{E}[x_i] = \frac{n}{n+1} v_{JJ}(i) - (v_{JJ}(i) - v_{JJ}(n+1)) \tag{23}$$

$$\approx \frac{n}{n+1}\left(\frac{A}{c_l}\right) \tag{24}$$

Now consider the tasks of type G. The $n+1$ strongest players for this task are the players of type G. Since these players were strictly weaker than the players of type J for the tasks of type J, none actively participated in any of those tasks. However, they will actively bid on the tasks of type G, following exactly the same patterns as the players of type J in bidding on the tasks of type J. Hence, the total expected utility to the contest-designer using crowdsourcing in the case of approximately identical costs (or values) for all players of a type is given by

$$\mathbb{E}[U_{cs}] \approx \sum_{i=1}^{n} \frac{n}{n+1}\left(\frac{A}{c_l}\right) + \sum_{i=1}^{n} \frac{n}{n+1}\left(\frac{A}{c_l}\right) - 2nA, \tag{25}$$

$$= 2n\frac{n}{n+1}\left(\frac{A}{c_l}\right) - 2nA \tag{26}$$

$$= 2nA\left(\frac{1}{c_l} - 1\right),\ \text{as } n \to \infty. \tag{27}$$

■

With enough skilled competitors, crowdsourcing can yield higher utility than even optimal assignment (eq. (16)), since $0 \leq \theta \leq 1$, and has been empirically estimated to be around 0.5.

[1]Such auctions with multiple goods and players have been extensively studied (Barut and Kovenock 1998; Clark and Riis 1998), and we build on this work here.

Theorem 6. *If $m = n$, and c_h is high enough that $\frac{A}{c_h} \approx 0$ then the expected utility from crowdsourcing $\mathbb{E}[U_{cs}] = 0$.*

Proof: This case follows as Thm. 5 until eq. (21). However, $F_{J,i}(x) \approx 1 - \left(1 - \frac{x}{v_{GJ}(1)}\right)^{\frac{1}{n}} \approx 0$ for all $1 \leq i \leq n$. Since there are only enough players to complete the task, each player is assured of winning a task and thus has no incentive to put in a non-zero bid. ■

Lack of competition leads to low performance by all players. Instead, in this case if tasks were assigned by a manager, albeit noisily, significant utility could be derived for at least some of the tasks. The natural thought process might lead us to believe that crowdsourcing contests are good for skill discovery — it is easy to think that expert players will clearly become obvious in a competitive setting. This setup gives a clear example where this would not be the case. Crowdsourcing may not be a good solution when the contest-designer has many tasks of a specialized nature that requires highly-skilled players who are in short supply. This model easily extends to more than two types of events and players.

When it is possible to divide a large project into smaller tasks the best way to harness multiple players is through multiple tasks. However, if contest-designers are able to divide a large task into many smaller tasks of different types that are matched to the different types of players in the crowd, both the designers and the players could receive higher utilities. Our guidelines for crowdsourcing tasks both support and complement those from Malone et al. (2010).

Conclusions and Future Work

Complementing empirical work in this area, we have developed a theoretical framework that may guide organizations that are considering crowdsourcing contests to do work. More broadly, we try to understand the implications and strengths and weaknesses of more fluid task-focused labor markets.

The simple models presented here are only a first step. Models involving multiple prizes, incomplete information of player strengths, and also those which take into account repeated game effects such as player learning and retention costs are necessary future work.

Acknowledgements

The authors would like to thank Eric Bokelberg, Ankur Mani, Anshul Sheopuri, and Lada Adamic for helpful discussions.

References

Archak, N., and Sundararajan, A. 2009. Optimal design of crowdsourcing contests. In *Proc. Int. Conf. Inf. Syst. (ICIS)*.

Barut, Y., and Kovenock, D. 1998. The symmetric multiple prize all-pay auction with complete information. *Eur. J. Polit. Econ.* 14(4):627–644.

Baye, M. R.; Kovenock, D.; and de Vries, C. G. 1996. The all-pay auction with complete information. *Econ. Theor.* 8(2):291–305.

Benkler, Y. 2006. *The Wealth of Networks*. Yale University Press.

Bettencourt, L. M. A. 2009. The rules of information aggregation and emergence of collective intelligent behavior. *Topics Cogn. Sci.* 1(4):598–620.

Bokelberg, E. H., and Varshney, L. R. 2012. Liquid: Harnessing the passion of the crowd within the enterprise. submitted.

Bollier, D. 2011. *The Future of Work*. Aspen Inst.

Boudreau, K. J.; Lacetera, N.; and Lakhani, K. R. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Manage. Sci.* 57(5):843–863.

Clark, D. J., and Riis, C. 1998. Competition over more than one prize. *Am. Econ. Rev.* 88(1):276–289.

DiPalantino, D., and Vojnović, M. 2009. Crowdsourcing and all-pay auctions. In *Proc. 10th ACM Conf. Electron. Commer. (EC'09)*, 119–128.

Dixit, A. 2002. Incentives and organizations in the public sector. *J. Hum. Resour.* 37(4):696–727.

Hillman, A. L., and Riley, J. G. 1989. Politically contestable rents and transfers. *Econ. Polit.* 1(1):17–39.

Hong, L., and Page, S. E. 2001. Problem solving by heterogeneous agents. *J. Econ. Theor.* 97(1):123–163.

Howe, J. 2008. *Crowdsourcing*. Three Rivers Press.

Liu, T. X.; Yang, J.; Adamic, L. A.; and Chen, Y. 2011. Crowdsourcing with all-pay auctions: a field experiment on Taskcn. In *Proc. Am. Soc. Inf. Sci. Tech. Annu. Meet.*

Lohr, S. 2009. The crowd is wise (when it's focused). *New York Times*.

Malone, T. W.; Laubacher, R.; and Dellarocas, C. 2010. The collective intelligence genome. *MIT Sloan Manage. Rev.* 51(3):21–31.

Schenk, E., and Guittard, C. 2011. Towards a characterization of crowdsourcing practices. *J. Innovation Econ.* 2011/1(7):93–107.

Tapscott, D., and Williams, A. D. 2006. *Wikinomics*. Portfolio Penguin.

Tapscott, D., and Williams, A. D. 2010. *Macrowikinomics*. Portfolio Penguin.

Varshney, L. R., and Goyal, V. K. 2006. Toward a source coding theory for sets. In *Proc. IEEE Data Compression Conf. (DCC 2006)*, 13–22.

Varshney, L. R.; Rhim, J. B.; Varshney, K. R.; and Goyal, V. K. 2011. Categorical decision making by people, committees, and crowds. In *Proc. 2011 Inf. Theory Appl. Workshop*.