Crowdsourcing with Tullock Contests: A New Perspective

Tie Luo*, Salil S. Kanhere†, Hwee-Pink Tan*, Fan Wu‡, Hongyi Wu§

*Institute for Infocomm Research, A*STAR, Singapore
†School of Computer Science and Engineering, The University of New South Wales, Australia
‡Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai Jiao Tong University, China
§Center for Advanced Computer Studies, University of Louisiana at Lafayette, USA

E-mail: luot@i2r.a-star.edu.sg, salilk@unsw.edu.au, hptan@i2r.a-star.edu.sg, fwu@cs.sjtu.edu.cn, wu@cacs.louisiana.edu

Abstract—Incentive mechanisms for crowdsourcing have been extensively studied under the framework of all-pay auctions. Along a distinct line, this paper proposes to use Tullock contests as an alternative tool to design incentive mechanisms for crowdsourcing. We are inspired by the conduciveness of Tullock contests to attracting user entry (yet not necessarily a higher revenue) in other domains. In this paper, we explore a new dimension in optimal Tullock contest design, by superseding the contest prize—which is fixed in conventional Tullock contests—with a prize function that is dependent on the (unknown) winner’s contribution, in order to maximize the crowdsourcer’s utility. We show that this approach leads to attractive practical advantages: (a) it is well-suited for rapid prototyping in fully distributed web agents and smartphone apps; (b) it overcomes the disincentive to participate caused by players’ antagonism to an increasing number of rivals. Furthermore, we optimize conventional, fixed-prize Tullock contests to construct the most superior benchmark to compare against our mechanism. Through extensive evaluations, we show that our mechanism significantly outperforms the optimal benchmark, by over three folds on the crowdsourcer’s utility cum profit and up to nine folds on the players’ social welfare.

I. INTRODUCTION

Crowdsourcing represents a new problem-solving model that elicits solutions, ideas, data, etc.—referred to as contributions— from an undefined, generally large group of people. Classic examples include Amazon Mechanic Turk, Yahoo! Answers, GalaxyZoo.org, TopCoder.com, etc. Recently, a new variant of crowdsourcing called participatory sensing emerged as a new data-collection model, which elicits sensor data contributed from user-owned mobile devices such as smartphones. Examples include GreenGPS [1], LiveCompare [2], ContriSense:Bus [3] and Waze.com, to name a few. Pivotal to the viability of all such crowdsourcing systems, is whether there is enough incentive to attract sufficient participation. A large body of prior work [4]–[11] has been dedicated to designing incentive mechanisms for such scenarios, where each incentive mechanism essentially determines some reward according to users’ contributions. The commonly adopted approach turns out to be auctions, where each bidder tenders a bid (e.g., planned sensing duration [4] or desired payment [5]) to the crowdsourcer, who will then choose the highest or lowest bidder(s) as the winner(s) to give out some reward. A widely used form among them is the all-pay auction [6]–[10], which nicely captures the scenario where each bid represents some irreversible effort. In other words, effort has to be sunk at the time of bidding, for example working out a solution to a problem, or sensing and sending data through a smartphone.

A distinctive characteristic of auctions in general, is that they are perfectly discriminating [12]: the best (highest or lowest) bidder wins, due to the inevitable sunk cost (every bidder has to pay for his bid regardless of whether he wins the auction or not), all bidders substantially shade (i.e., decrease) their bids for fear of loss [13]. Furthermore, if a bidder believes that there exists some other bidder who will bid higher than him, he will choose not to bid at all. Indeed, as Franke et al. [14] pointed out, all-pay auctions are so discriminative that, under a complete-information setting, only two (strongest) players will enter the auction in equilibria and only one of them will have a positive expected payoff. Clearly, this is not desirable in many crowdsourcing campaigns that favor a large participant pool (yet not necessarily higher revenue) for the sake of more population diversity (e.g., LiveCompare [2]) and/or larger geographic coverage (e.g., GreenGPS [1] and Waze).

Taking a radically different approach, this paper proposes to use Tullock contests as an alternative framework to design incentive mechanisms for crowdsourcing. Fathered by Tullock’s seminal work [15], Tullock contests represent a distinct contest regime that is imperfectly discriminating: every player has a strictly positive probability to win (determined by a contest success function) as long as he bids.1 This characteristic makes Tullock contests highly conducive to attracting user entry, especially weak players [14],2 which is of high practical interest because weak players often constitute the majority of potential participants in crowdsourcing. This explains why, in reality, we see lotteries—the simplest form of Tullock contests—much more often, and usually engaging a larger number of participants, than all-pay auctions.

On the other hand, Tullock contests are not necessarily superior to all-pay auctions in terms of revenue, or total user contribution. In fact, there is no conclusive theory as to which contest regime revenue-dominates the other in general [14], [16], [17], and an experimental study also shows that revenue in all-pay auctions may be independent of the number of participants at some stable state [18]. Intuitively, the reason is that (a) the fierce competition induced by all-pay auctions efficaciously incentivizes (a small number of) strong players to exert high effort, while (b) the mild competition in Tullock contests attracts a medium amount of contributions from more (albeit weaker) players. Therefore,

1In an asymptotic limiting case, Tullock contests subsume all-pay auctions. However, they are generally classified as two different contest regimes.

2A “strong” or “weak” player in this paper refers to a player with a strong or weak type; type, as a term in Bayesian games and mechanism design, refers to the private information held by a player, such as his valuation of the auctioned item or his production cost. Therefore, a strong player means a player with high valuation or low cost, and vice versa.
A. Related Work

Even in their simplest form, Tullock contests are analytically more challenging to tackle than most classic auctions. This is particularly true in the incomplete-information setting\(^3\) which is a more realistic setting for crowdsourcing. Specifically, the equilibria of most classic auctions with complete information, or with incomplete information and symmetric players, can be solved in closed form; but Tullock contests with incomplete information is generally intractable in analytical means, even in the simplest form (lottery) [24]. This can be attributed to the double uncertainty: in auctions with incomplete information, the uncertainty about other players’ types is the only source of uncertainty; but in Tullock contests, the imperfectly discriminating nature—or more specifically the probabilistic winner selection (unlike in auctions the highest bid guarantees winning)—creates another source of uncertainty.

As a consequence, the literature on Tullock contests exclusively deals with the complete-information setting or restrictive versions of the incomplete-information setting (e.g., only two discrete types for two players [25]). It was not until 2008 that Fey [21] first proved the existence of a (symmetric) equilibrium for a lottery with incomplete information. However, the model is limited to two players and uniform distribution, and the equilibrium strategy is only numerically characterized.

A subsequent breakthrough was made by Ryvkin [22], who extended Fey’s model [21] by allowing for more than two players, arbitrary continuous distributions, and a more general contest success function. He also proved the existence of equilibria (leaving uniqueness as future work), following the spirit of [21]. Still, the equilibrium strategy was only numerically computed, due to the limited analytical tractability of Tullock contests. So far, the only known analytical solution to equilibria with a continuous type distribution, is due to Everhart [26] who constructed a rather special distribution to obtain a closed-form expression. However, the distribution is rather complex and not generalizable, and the model is still limited to a two-player lottery only.

In this paper, under a general crowdsourcing model with incomplete information, we derive the optimal prize function that maximizes the crowdsourcer’s utility cum profit. Surprisingly, our solution of the unique Bayesian Nash equilibrium (a) can be expressed in a simple and closed form in general cases, and (b) is agnostic to the number of players. These are in stark contrast to prior art, and in practical terms, imply that our mechanism (a) can be easily implemented in web agents and smartphone apps that act in a fully distributed manner, and (b) overcomes the disincentive to participate caused by player’s antagonism to an increasing number of rivals.

Along the line of optimal Tullock contest design, two general directions have been pursued in prior work. One stream of research explores whether the prize should be allocated to a single winner or all the players in a hierarchical manner. For example, [20] applies a rank-dependent expected utility model to a lottery in which the prize was divided into, according to ranks, a few large prizes and a large number of small prizes. However, [27] proves that it is optimal to give the entire prize to a single winner in a symmetric equilibrium. The other direction focuses on whether and how to bias players in such a way that induces the maximum revenue. For instance, [14] allows the crowdsourcer to assign different weights (preferences) to players in a discriminative manner for revenue maximization. [19] proves

\(^3\)Briefly speaking, with complete information all the players are informed of all the others’ types, while with incomplete information each player only knows his own type. Section II explains this in more detail.
that a biased lottery (like [14]) achieves the same revenue as a biased all-pay auction, when both are fully optimized. In the fair case, [19] proves that an optimal lottery is always superior to an optimal all-pay auction.

Our proposed approach represents a new dimension in the design space of Tullock contests. Provisioning contest prize as a function of the unknown winner’s contribution sets this work apart from all prior work on Tullock contests in which prizes are fixed and known ex ante.

B. Contributions

The main contributions of this paper are summarized below:

1) This work is the first attempt in the crowdsourcing literature to explore a new dimension of optimal Tullock contest design. Our model assumes a very general form of CSF: extensive performance evaluation is provided in Section V which allows the model to derive the optimal Tullock contest. The optimal all-pay auction is twice differentiable and concave, which captures the common phenomenon of diminishing marginal return when exerting effort. When \( b_i = 0 \) for all \( i \), i.e., no one exerts effort, we assume \( \Pr(b_i) = 0 \), i.e., no one will win any prize.\(^4\)

Another crucial component of our contest model is a prize function \( V(\xi_w) \) that we specifically introduce in the contest, which is a (monetary) prize of a common value that is dependent on the (unknown) winner’s contribution \( \xi_w \). Accordingly, a player \( i \) will receive an expected income of \( \Pr(b_i) V(\xi_i) \). This function \( V(\cdot) \) is common knowledge to all players (e.g., via announcement by the crowdsourcer).

Each player is characterized by his type—his marginal cost of exerting effort—denoted by \( c_i \in [c, \tau] \), where \( 0 < c < \tau \). That is, exerting effort \( b_i \) will incur a cost of \( c_i b_i \) to player \( i \). Thus, if the effort profile of all the players is \( \mathbf{b} := (b_1, b_2, ..., b_n) \), the payoff of player \( i \) can be expressed in the following quasi-linear form:

\[
\Pr(b_i) V(\xi_i) - c_i b_i.
\]

Since \( \xi_i = g(b_i) \), a player’s (ex post) payoff given the contribution strategy profile of all the players \( \xi := (\xi_1, \xi_2, ..., \xi_n) \), is

\[
\bar{v}_i(c_i, \xi) = \frac{\xi_i}{\sum_{j=1}^{n} \xi_j} V(\xi) - h(\xi_i)c_i
\]

where \( h := g^{-1} \) is the inverse function of \( g(\cdot) \). We note that \( g^{-1} \) exists because \( g(\cdot) \) is strictly monotone.

As crowdsourcing typically involves an undefined group of people, we assume the interim stage which corresponds to an incomplete-information setting; each player \( i \) is informed of his own type \( c_i \) but not of others.\(^5\) yet it is common knowledge that all the \( c_i \) are independently drawn from a continuum \([c, \tau]\) according to a c.d.f. \( F(c) \) or p.d.f. \( f(c) = F'(c) \).\(^5\) On the contrary, the ex ante stage corresponds to a no-information setting where players do not know anyone’s type including their own, and the ex post stage corresponds to a complete-information setting where all the players’ types are common knowledge.

The crowdsourcer collects revenue from the aggregate contribution of all the players, and bears the cost of paying for the (variable) prize. The profit cum utility of the crowdsourcer is thus

\[
\tilde{\pi} = \nu \sum_{i=1}^{n} \xi_i - V(\xi_w)
\]

where \( \nu \) is the crowdsourcer’s valuation of per unit user contribution. This parameter \( \nu \) does not appear in prior work where a unity value is always implicitly assumed. However, explicitly modeling this parameter not only homogenizes the dimension of the expression (3), but also allows us to investigate the impact of \( \nu \) on key metrics such as the player strategy, prize, profit, and social welfare, which turns out (cf. Section V-D) to be an interesting combination of both linearity and nonlinearity.

\(^4\)An alternative and more commonly adopted practice in the literature, is assuming \( \Pr(b_i) = 1/n \). However, the prizes therein are all fixed, and hence if \( b_i = 0 \) for all \( i \), a player \( j \) will have incentive to deviate by exerting an infinitesimal effort \( \epsilon > 0 \) to increase his payoff by \((1 - \frac{1}{n})\epsilon v_j - O(\epsilon)\) where \( v_j \) is his valuation of the prize. Therefore, an "all-zero-bid" equilibrium does not exist. In our case, however, the prize is a function \( V(\cdot) \) of contribution and \( V(\cdot) \) can be so small that players lose the incentive to deviate from all-zero bids. Therefore, we impose \( \Pr(b_i) = 0 \) if \( b_i = 0 \), \( \forall i \) to reinstate this incentive. Indeed, we will show later in Proposition 2 that our mechanism ensures that a player will receive strictly positive payoff if he exerts non-zero effort.

\(^5\)In practice, such a distribution \( F(c) \) can be obtained (and published) by the crowdsourcer based on historic data, or—when historic data is not available—assume uniform distribution as widely used in the Bayesian game literature (our model constitutes a Bayesian game with private information being player types and common prior being the type distribution).
III. PROPOSED OPTIMAL TULLOCK CONTEST

The solution concept of a game with incomplete information is a pure strategy Bayesian Nash equilibrium, in which each player plays a strategy that maximizes his expected utility given his belief about other players’ types and that other players also play their respective equilibrium strategies. Formally, it is a strategy profile \( BNE = (\xi_1^{BNE}, \xi_2^{BNE}, \ldots, \xi_n^{BNE}) \) that satisfies

\[
  u_i(c_i, \xi_i^{BNE}, \xi_{-i}^{BNE}) \geq u_i(c_i, \xi_i, \xi_{-i}^{BNE}), \quad \forall \xi_i, \forall i,
\]

where \( u_i \) is the expected utility of player \( i \), defined as

\[
  u_i(c_i, \xi_i) := \mathbb{E}_{\xi_i} \left[ \tilde{u}_i(c_i, \xi) \right] \quad (4)
\]

where \( \xi = (\xi_1, \xi_{-1}) \), and \( \tilde{u}_i \) is defined in (2).

Definition (4) can be expanded as follows. In a Bayesian Nash equilibrium, each player’s strategy \( \xi_i \) is a function of his own type \( c_i \) and the common prior, i.e., each player’s belief about all the other players’ types. As our setting is symmetric, in which the prior is a common distribution \( F(\cdot) \) for all the players, we focus on symmetric equilibria in which any player \( i \)'s equilibrium strategy is specified by a function \( \beta : [c, \bar{c}] \rightarrow \mathbb{R} \) as \( \xi_i = \beta(c_i), \forall i \). Therefore, definition (4) can be rewritten based on (2), as

\[
  u(c, \xi) = p(\xi) V(\xi) - h(\xi)c \quad (5)
\]

for an arbitrary type \( c \) and strategy \( \xi \), where

\[
  p(\xi) := \frac{\xi}{\int_{0}^{1} \xi + \sum_{j=1}^{n-1} \beta(\tilde{c}_j) \prod_{j=1}^{n-1} dF(\tilde{c}_j) \quad (6)}
\]

in which \( \Theta := [c, \bar{c}] \).

Thus, for a particular player \( i \), his expected utility is \( u_i = u(c_i, \xi_i) \) which can be computed from (5).

**Proposition 1 (Existence and Uniqueness of Equilibrium).** Our Tullock contest model admits a unique, monotone decreasing, pure-strategy Bayesian Nash equilibrium.

Due to space constraint, we defer all the contents of this paper to [28].

Henceforth, we will exclusively deal with the equilibrium, and thereby drop the superscript BNE for brevity.

**Lemma 1 (Equilibrium Contribution Strategy).** Given an arbitrary prize function \( V(\cdot) \), the (symmetric) equilibrium strategy \( \beta(\cdot) \) of our Tullock contest, as in \( \xi = \beta(c) \), is implicitly given by

\[
  p(\xi) V(\xi) - h(\xi)c = \int_{c}^{\bar{c}} h(\beta(\bar{c})) \, d\bar{c}, \quad \forall c \in [c, \bar{c}] \quad (7)
\]

We remark on the following:

- Equation (7) has an intuitive interpretation: player \( i \)'s expected utility, as represented by the l.h.s., is determined by his cost advantage relative to the highest-cost player, modulated by his contribution level.
- There is no closed-form solution to (7) for an arbitrary \( V(\cdot) \). In fact, even the fixed-prize case (as in conventional contests) does not have a closed-form solution in general either [24]. However, counter-intuitively, our approach of using a non-fixed prize (as a function)—which seems to be introducing more complexity—leads to a closed-form equilibrium solution in most general cases. We remark on this tractability following the next theorem.

**Theorem 1 (Optimal Prize Function, Strategy, and Maximum Profit).** The optimal prize function that maximizes the crowdsourcer’s utility (profit) in our Tullock contest is given by

\[
  V^*(\xi_w) = \left[ \beta^{-1}(\xi_w) h(\xi_w) - \int_{\xi_w}^{\xi} h(\xi) \, d\beta^{-1}(\xi) \right] / p(\xi_w) \quad (8)
\]

where \( \xi = \beta(\bar{c}) \) and \( \beta^{-1}(\cdot) \) is the inverse function of the equilibrium strategy \( \beta(\cdot) \) which, as in \( \xi = \beta(c) \), is given by

\[
  h'(\xi) = \frac{\nu}{c + F(c)}, \quad \forall c \in [c, \bar{c}] \quad (9)
\]

The induced maximum profit of the crowdsourcer is

\[
  \pi^* = n \int_{c}^{\bar{c}} \left[ \nu \beta(c) - h(\beta(c))c \right. \quad + \frac{F(c)}{f(c)} [h(\beta(\bar{c})) - h(\beta(c))] \, dF(c). \quad \quad (10)
\]

We will illustrate how to put Theorem 1 to use, in Section V with a case study. Here we remark on the following:

- The strategy \( \xi = \beta(\bar{c}) \), as of the highest-cost or weakest player, is always 0 in all-pay auctions under standard assumptions. However, in Tullock contests, this is not necessarily the case (unless \( \bar{c} = \infty \)), which will be evidenced in Section V. The reason is that any Tullock contestant has a positive winning probability as long as he exerts nonzero effort, whereas all-pay auctions perfectly discriminate the weakest bidders who have no chance to win.

- **Analytical tractability and Practical implication:** An interesting and somewhat surprising observation is that, while “functionizing” the contest prize would, intuitively, seem to introduce complexity to conventional, fixed-prize contests, the equilibrium strategy (9) turns out to be much simpler compared to the fixed-prize case (cf. (12) in Section IV). In fact, for most and common functions \( h(\cdot) \), it (9) can be expressed in closed form. This convenient analytical tractability is in stark contrast to prior art (see [21]–[23] and a survey [24]) where equilibria do not have analytical solutions in general and can only resort to numerical methods. Theoretically, this lends us a lot of convenience in subsequent technical treatments and other possible future extensions. Practically, this fosters the application of our mechanism via easily-implementable software deployed in web agents and smartphone apps that act on each user’s behalf to determine his contribution strategy in a fully distributed manner.\(^7\)

\(^7\)Note that the optimal prize function (8) and maximum profit (10) are computed on the other hand by a centralized, and computationally powerful server, and the computation takes the already-solved \( \beta(\cdot) \) (9) as input, unlike in (12) \( \beta(\cdot) \) is unknown. In fact, in many cases such as demonstrated in Section V, the computation in our mechanism is fairly straightforward.
of winning will be diluted, and hence if the prize is fixed, each player will have to expect a lower utility, resulting in a disincentive to participate. However, now that players can stay agnostic to the participant pool size,\(^8\) they need not worry about an increasing number of rivals, which certainly strengthens the motivation to participate or stay in the campaign. In more practical terms, participants would even not be averse to spreading the awareness and publicity of a campaign, which helps further expand the participant pool.

Now we state an important condition pertaining to general incentive mechanisms: individual rationality (IR) [29]. It means that any participating player should receive in equilibrium a nonnegative expected utility, or in other words, each player should be better off or at least remain neutral by participating. In the following, we prove that our mechanism possesses a stronger version of IR.

**Proposition 2** (Strict Individual Rationality). Our Tullock contest with the prize function given by Theorem 1 satisfies strict individual rationality (SIR), where all the players receive strictly positive expected utility, except that a player of type \(\bar{\sigma}\)—which happens with probability zero—expects a surplus of zero (and hence is indifferent in participating).

Another often-discussed property in mechanism design is incentive compatibility (IC) [29] or truthfulness, which means that all players report their types truthfully. This property is technically irrelevant to our mechanism because, unlike some other mechanisms such as [4], [5] which determine workers’ wages based on worker-reported types (costs or desired payments), our mechanism determines players’ reward based on observable user contributions rather than unobservable (and private) player types (costs). On the other hand, those other mechanisms can satisfy IC trivially by paying a wage no less than a worker’s reported cost or payment, provided that IC is satisfied; but in our case, satisfying IR requires a carefully designed prize function (8) (which is demonstrated by the proof of Proposition 2).

### IV. Optimal Fixed-prize Tullock Contests

This section constructs a conventional, i.e., fixed-prize, Tullock contest for the sake of comparison with our mechanism (later in Section V).\(^9\) However, we augment the prior art of dealing with conventional contests [21]–[23], from solving the equilibrium to optimizing (as well as solving) the equilibrium, in order to create the most superior benchmark to challenge our mechanism.

Specifically, we set to find an optimal fixed prize \(V_0^*\) that maximizes the crowdsource’s utility \(\pi_0\) through a particular equilibrium \(\xi_0^*\), and such a solution needs to be found for every possible value of \(\nu\), the valuation of contribution. Formally, the problem is formulated as \(\max_{V_0 \geq 0} \pi_0\) where

\[
\pi_0 = \nu \mathbb{E}_c \left[ \sum_{i=1}^{n} \xi_{0i} \right] - V_0 = \nu \int_{c}^{\infty} \beta_0(c) \ dF(c) - V_0. \tag{11}
\]

The equilibrium strategy \(\xi_0 = \beta_0(c)\) is given by Proposition 3.

**Proposition 3.** In a Tullock contest with fixed prize \(V_0\), the equilibrium strategy \(\xi_0 = \beta_0(c)\) is implicitly determined by

\[
\int_{0}^{\infty} \frac{\sum_{j=1}^{n} \beta_0(c_j)}{\beta_0(c) + \sum_{j=1}^{n} \beta_0(c_j)} \prod_{j=1}^{n-1} dF(c_j) = h'(\xi_0) \frac{c}{V_0}. \tag{12}
\]

Unfortunately, (12) does not have an analytical solution. The special case of \(V_0 = 1\) was numerically tackled by [21]–[23]. In our case, \(V_0\) is not given and we need to find the optimal \(V_0\) that maximizes \(\pi_0\) (11). In the meantime, \(\pi_0\) contains \(\beta_0(c)\) (12) which is analytically intractable. Moreover, the optimal solution \(V_0^*\) must not be a value but a function of \(\nu\). This problem is technically challenging.

Our solution was inspired by solving the Fredholm equations using a numerical method described in [30] (Chap. 5). Consider a two-player case for simplicity. The first key idea in our solution is to transform the integral in (12) into a quadrature sum, by supposing that \(V_0\) is given:

\[
\sum_{j=1}^{m} \frac{\beta_0(t_j) f(t_j) \Delta_j}{\beta_0(c) + \beta_0(t_j)} + R_m(c) = h''\beta_0(\xi_0) \frac{c}{V_0}, \tag{13}
\]

where \(t_j, j = 1, 2, ..., m\), are the quadrature points distributed in \(\Theta\), \(\Delta_j\) is determined by the chosen quadrature scheme (e.g., Gaussian), and \(R_m(c)\) is the residual error due to transforming the original integral into the quadrature sum. In our case, a uniform quadrature scheme suffices and thereby \(\Delta_j = \Delta_m := (\bar{\sigma} - c)/m\). In addition, since the integrand is atomless (\(\beta_0(t_j) = 0\) for all \(j\) is not an equilibrium because a player will have incentive to deviate by an infinitesimal amount to gain positive utility), the integral can be closely approximated by the quadrature sum for a sufficiently large \(m\), in which case \(R_m(c)\) can be safely ignored.

The second key idea is to note that, since (13) holds for all \(c \in [\bar{c}, \bar{c}]\), it must also hold for all the \(c_i\) that equal to the quadrature points \(t_j, j = 1, 2, ..., m\). Thus, (13) is further transformed into a system of \(m\) nonlinear equations:

\[
\Delta_m \sum_{j=1}^{m} \frac{\beta_0(c_j) f(c_j)}{\beta_0(c) + \beta_0(c_j)} - h_2\beta_0(\xi_0) \frac{c_i}{V_0} = 0, \ i = 1, ..., m. \tag{14}
\]

This system can be solved using the Matlab function `fsolve`.

Similarly but on a much simpler scale, the profit (11) can be approximated by

\[
\pi_0(\nu, V_0, \beta_0(c)) = n\nu \Delta_m \sum_{j=1}^{m} \beta_0(c_j) f(c_j) - V_0 \tag{15}
\]

where \(\beta_0(c_j)\) are the solutions to the system (14).

The entire solution is outlined in a self-explanatory fashion by the pseudo-code in Algorithm 1. In the actual implementation, we also added a testing condition to ensure the range \([V_0, V_0]\) to be large enough to include the peak point of \(\pi_0\) (which can be shown to be concave in \(V_0\)); we also improved the efficiency by adding a stopping condition to terminate the inner loop faster. These are peripheral and hence omitted in Algorithm 1.
Algorithm 1: Optimizing fixed-prize Tullock contest

**Input:** bounds $c, \tau, A, V, V_0; m$; step sizes $\delta_1, \delta_2$

**Output:** $V^*_0 : \nu \to \mathbb{R}_+$ (a vector of optimal prizes)

1. $\pi^*_0 : \nu \to \mathbb{R}_+$ (a vector of maximum profits)
2. $\mathbf{E}_0^* : \nu \to \mathbb{R}_+$ (a matrix of which each row is an $m$-element equilibrium strategy profile corresponding to a $\nu \in \nu$)

1. $\Delta_m \leftarrow (\tau - \xi)/m$;
2. $\xi \leftarrow \{\xi_0 + \Delta_m, \xi + 2\Delta_m, \ldots, \xi + (m-1)\Delta_m\}$;
3. $V_0 \leftarrow \{V_0, V_0 + \delta_2, V_0 + 2\delta_2, \ldots, V_0\}$;
4. Create a $|V_0| \times m$ strategy matrix $\xi^A$;

5. for $\nu \leftarrow \nu$ to $\tau$ (step size: $\delta_1$)
6. $\tilde{\pi}_0 \leftarrow 0$;
7. for $i \leftarrow 1$ to $|V_0|$ do
8. create $F_m(c, \Delta_m, V_0[i]) = 0$ according to Eq. (14);
9. $\xi_0 := (\xi_0[j] = 1, 2, \ldots, m) \leftarrow f(solve(F_m));$
10. $\xi^A[i] \leftarrow \xi^A[i]$ where $\xi^A[i]$ is the $i$-th row of $\xi^A$;
11. $\hat{\pi}_0[i] \leftarrow \hat{\pi}_0[V_0[i], \xi_0]$ calculated using Eq. (50);
12. end
13. $\pi^*_0(\nu) \leftarrow \max_i(\hat{\pi}_0)$;
14. $V^*_0(\nu) \leftarrow V_0[\arg \max(V_0[\nu])]$;
15. $\mathbf{E}^*_0(\nu) \leftarrow \xi^A[\arg \max(V_0[\nu])]$;

which we refer to as Tullock-OPF, and the latter uses the optimal fixed prize determined by Algorithm 1, which we refer to as OptBenchmark. We stress that this benchmark is not designed to favor our proposed mechanism, but honestly follows the well-established standard model and, in fact, is fully optimized.

The performance metrics we evaluate are:

(a) the equilibrium contribution strategy of players, which corresponds to revenue;
(b) the prize that the crowdsourcer provisions, which corresponds to cost;
(c) the expected utility of the crowdsourcer, which corresponds to profit;
(d) the social welfare of the campaign, which is the aggregate utility of all the players at equilibrium and we denote by $U := \mathbb{E}[\sum_{i=1}^n u_i]$; it measures the total surplus of the community by participating in the crowdfunding campaign.

The rationale is that a company is typically profit-driven and thus concerned about the metrics (a)–(c), while a government agency or a non-profit organization may focus more on the metric (d).

The primary scenario in this section consists of two players whose marginal contribution costs are independently drawn from a uniform distribution $F(c) = c - 1, c \in [1, 2]$. Player effort $b$ is converted to user contribution $\xi$ according to $\xi = g(b) = \sqrt{b}$, and hence $h(\xi) := g^{-1}(\xi) = \xi^2$. Such a setup is common in the literature such as [21], [22]. This scenario is then extend to a $n$-player setting.

A. Tullock-OPF: Analytical Results

Our mechanism can be solved analytically, and the solving process also illustrates how to put Theorem 1 to use. First, the equilibrium contribution strategy can be obtained via (9) as

$$\xi = \beta(c) = \frac{\nu}{4c - 2}. \quad (16)$$

Hence $\beta^{-1}(\xi) = \nu/(4\xi) + 1/2$, and the numerator of (9) equals

$$\xi^2(\frac{\nu}{4\xi} + \frac{1}{2}) + \int_{\xi}^\infty \frac{\nu}{4\xi^2} \: d\xi = \xi^2 + \frac{\nu\xi}{2} - \frac{\nu^2}{24}$$

where $\xi = \beta(\tau) = \nu/6$. Using (6), the denominator of (9) equals

$$p(\xi) = \nu^2 \int_1^{\xi} \left(1 - \frac{\nu}{\xi(4c - 2) + \nu}\right) \: dc = \nu^2 \int_1^{\xi} \cdot \frac{1}{\xi} \cdot \frac{1}{\xi + \nu} \: dc$$

$$= \nu^2 \int_1^{\xi} \frac{1}{\xi + \nu} \: dc = \nu^2 \int_1^{\xi} \frac{1}{\xi + \nu} \: dc$$

Finally, the induced maximum profit is obtained via (10):

$$\pi^* = \frac{\nu^2}{4c - 2} + \frac{\nu^2}{(4c - 2)^2} + c - 1 \cdot \frac{1}{36} (\nu^2 - (4c - 2)^2) \: dc$$

$$= 2\nu^2 \frac{1}{8c - 4} \cdot \frac{c - 1}{36} \: dc = \frac{\log 3}{4} + \frac{1}{36} \nu^2 \quad (18)$$

B. OptBenchmark: Numerical Results

OptBenchmark can only be numerically solved, using Algorithm 1, of which the parameters are specified by Table I.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu$</td>
<td>0.5</td>
<td>$\tau$</td>
<td>5</td>
<td>$V_0$</td>
<td>0.01</td>
<td>$V_0$</td>
<td>5</td>
</tr>
</tbody>
</table>

Fig. 1 reveals the trajectory of finding the duple $(V_0^*, \pi_0^*)$, i.e., the optimal prize and maximum profit, by evaluating over a range of possible prizes, for each $\nu = 1, 2, 3$. The optimal duple is found at the peak of each curve, and the curves clearly demonstrate the concavity of profit versus prize, which confirms the existence and uniqueness of the optimum. Furthermore, how the optimal duple $(V_0^*, \pi_0^*)$ is affected by $\nu$ is examined by Fig. 2. Interestingly, prize $V_0^*$ coincides with profit $\pi_0^*$ for all the $\nu$'s. This indicates that the revenue of OptBenchmark is double of the cost (prize). This also implies that, if in Fig. 1 we draw all the other trajectories (in addition to the in-situ three), the peak points of all the trajectories will all fall onto the same straight line $y = x$.

Another observation in Fig. 2 is that both prize and profit are convexly increasing in $\nu$. We will revisit this nonlinear behavior together with other subsequent observations in Section V-D.

C. Comparison

Given that both Tullock-OPF and OptBenchmark are solved by now, we proceed to compare them with respect to the four metrics mentioned earlier.

Crowdsourcing Revenue: Fig. 3a examines the equilibrium contribution strategy of players as a function of player type, for each $\nu = 1, 2, 3$. We remark on three observations. First, in all the cases (3×2 curves), the strategy is monotone decreasing in type, which conforms to our Proposition 1 (which subsumes $V(\cdot)$ being constant). The convex trend is also consistent with the literature: for example, we verified a special case of OptBenchmark with $h(\xi) = \xi$, $\nu = 1, c \in [0.01, 1.01]$, which is the same as [21], and the result (not reproduced here) exactly matched [21]. Second, for any $\nu$, Tullock-OPF elicits significantly higher contribution.
Out of the revenue: Equilibrium contribution strategy.

The duple of optimal prize and maximum profit, annotated at the peak of each curve.

Figure 1: OptBenchmark: Trajectory of finding the duple of optimal prize and maximum profit.

(a) Revenue: Equilibrium contribution strategy. Out of the m = 100 quadrature points, only 20 are shown for better visibility.

(b) Cost: Prize. Note that the support of Tullock-OPF is \( \xi \in [\nu/6, \nu/2] \).

(c) Profit.

Figure 3: Comparison of Tullock-OPF against OptBenchmark.

Third, in both mechanisms, the highest-cost or weakest player makes strictly positive contribution, i.e., \( \xi = \beta(\tau) > 0 \), indicating a sheer contrast between Tullock contests and auctions where \( \xi = 0 \); here we recall the first remark below Theorem 1.

Crowdsourcing Cost: Fig. 3b compares the optimal prize function \( V^*(\xi_u) \) in Tullock-OPF against the optimal fixed prize \( V^*_0 \) in OptBenchmark. Note that the support of \( V^*(\xi_u) \) is \( [\xi, \xi] = [\nu/6, \nu/2] \) which follows from (16). While it is intuitive that each of the three \( V^*(\xi_u) \) curves increases in winner’s contribution, it is interesting to note that each curve fits a straight line very well, which suggests that the computation of (17) could be remarkably simplified in practice via linear approximation. While this advantage should not be overstated as to how it generalizes, it hints at a possible line of future work. Moreover and noteworthy, the diagram shows that the prize offered by Tullock-OPF is generally well above OptBenchmark. This raises an important question that whether this much higher cost to be borne by the crowdsourcer will eventually pay off, which is answered next.

Crowdsourcing Profit: Fig. 3c evaluates the maximum profit that a crowdsourcer garners from the two mechanisms. The salient observation is that Tullock-OPF outperforms OptBenchmark by a large profit margin, which strongly corroborates our proposal of using an optimized prize function to supersede the conventional, fixed prize in Tullock contests. For a closer examination, we collate the results into Table II and calculate the ratio between Tullock-OPF and OptBenchmark against OptBenchmark: Social Welfare.

Figure 4: Comparison of Tullock-OPF against OptBenchmark: Social Welfare.
by a strategy profile $\xi_0^i$. Thus, we can compute the social welfare as

$$U_0 = n \Delta m \sum_{i=1}^{m} f(c_i) \left( \Delta m \sum_{j=i}^{m} \xi_{0j}^2 \right),$$

which can be understood by rewriting (19) as

$$U = n \int \pi f(c) \int_c^{\pi} \xi^2 \, dc$$

and noting that Lemma 1 applies, without change, to fixed-prize cases where $V(\cdot)$ is a constant.

The results are presented in Fig. 4. The observation is exciting: Tullock-OPF outstrips OptBechmark in an even more striking manner as compared to profit in Fig. 3c. The detailed data are collated in Table III, which indicates a remarkable improvement of 7-9.3 folds. On the other hand, the nearly constant ratio in Table II is not duplicated here.

It may be puzzling as to why the crowdsourcer can reap higher profit while, at the same time, users altogether also gain higher surplus. This constitutes a “win-win” situation which is highly desirable but typically hard to attain. To explain this, we note the following rationales. First, generally speaking, crowdsourcing is not a zero-sum game like the stock market; rather, it involves a wealth creation process in which users exert effort to create “something” valuable that we have abstracted as “contribution”. Second, there exists a value asymmetry between players and the crowdsourcer, where the crowdsourcer typically values contribution higher than players do. This value asymmetry is a common phenomenon in reality: for example in the worldwide emerging Smart City and Smart Nation initiatives nowadays, citizen-generated data such as ambient noise and GPS traces collected by smartphones do not usually bear much value to the phone owners but can be very valuable to a crowdsourcer such as a noise control bureau or a transport company (like Waze). Thus, it makes perfect business sense for a crowdsourcer to provision a certain attractive reward to incentivize more user contribution which in turn bears even more value to the crowdsourcer.

D. Impact of Crowdsourcer’s Valuation

Recall that we have introduced a new parameter, $\nu$, to our model in Section II (cf. (3)). Herein, we investigate how it affects the various performance indicators.

This effect is explicitly examined with respect to the optimal fixed prize (Fig. 2), maximum profit (Fig. 3c), and social welfare (Fig. 4), where it is clearly shown that the impact of $\nu$ on these three metrics is nonlinear (convex). Furthermore, this effect is also implicitly examined with respect to the equilibrium contribution strategy (Fig. 3a) and optimal prize function (Fig. 3b) (in both diagrams one needs to compare across different curves corresponding to different $\nu$’s), and we see that the impact of $\nu$ on these two metrics is approximately linear.

The main message conveyed here is that, if a crowdsourcer increases his valuation of user contribution, e.g. by improving his business processes to better exploit user contribution, his profit and the players’ social welfare will both increase faster. This is an interesting finding uncovered due to our introduction of $\nu$. Moreover, this observation applies to both Tullock-OPF and the conventional case.

Now we dive in deeper to explain the correlation between the above nonlinearity and linearity. On the one hand, the case of OptBechmark can be nicely explained: (a) the nonlinear profit (Fig. 3c) results from the linear revenue (contribution; Fig. 3a) and nonlinear cost (prize; Fig. 2), and (b) the nonlinear social welfare (Fig. 4) follows from the nonlinear prize (player’s gain; Fig. 2) and the linear strategy (player’s cost; Fig. 3a).

On the other hand, the case of Tullock-OPF is not that straightforward, since both revenue (contribution) and cost (prize) seem to be linear across different $\nu$’s. In fact, an overlooked fact was that the prize functions in Fig. 3b are shifted horizontally, and thus one should compare prizes across $\nu$’s for the same winner rather than for the same amount of contribution. To do this, a simple way is to compare the maximum (or the minimum) winner contribution $\xi_w$ across different curves, as it can uniquely identify a particular winner. This reveals that the impact of $\nu$ on prize is actually nonlinear. Combined with the linearity on revenue (Fig. 3a), this explains why the profit and social welfare in Tullock-OPF are nonlinear in $\nu$.

E. $n$-Player Case

In this section, we extend our investigation to $n$ players. We conjecture that the above comparison results will continue to hold in the $n$-player case, and hence we will not repeat the same comparisons. In fact, as Ryvkin [22] pointed out, it is computationally infeasible to numerically solve fixed-prize Tullock contests for an arbitrary large number of players because of the “curse of dimensionality”. Therefore, this section focuses on Tullock-OPF. In particular, we are interested in how the composition of a participant pool, i.e., the distribution of the player types, affects the two key metrics, profit and social welfare.

We consider two participant pools: Population-1 draws player types from the same distribution as above, i.e., $F(c) = c - 1, c \in \left[1, 2\right]$, while Population-2 draws from another distribution $G(c) = \frac{1}{2} - \frac{1}{2}, c \in \left[0.5, 2.5\right]$. Thus, Population-2 is more diverse—or more uncertain in player types—than Population-1, while they statistically share the same mean value (1.5).

Following from Theorem 1, the equilibrium strategy in Population-2 is $\xi_G = \frac{\nu}{\pi e - 1}$, and thus the maximum profit is

$$\pi_{n,G} = \frac{n}{2} \int_{0.5}^{2.5} \left[ \frac{\nu^2}{4 c - 1} - \frac{\nu^2 c}{(4 c - 1)^2} + (c - 1) \pi \left( \frac{\nu^2}{8 c^2} - \frac{\nu^2}{162} \right) \right] dc$$

$$= \frac{n \nu^2}{2} \int_{0.5}^{2.5} \left( \frac{1}{8 c^2 - 2} - \frac{2 c - 1}{162} \right) dc = \frac{\log 3}{8} + \frac{2}{81} n \nu^2.$$

For Population-1, we leverage (18) as a shortcut to obtain

$$\pi_{n,F} = \left( \frac{\log 3}{8} + \frac{1}{72} \right) n \nu^2.$$  

Table II: Profit Comparison and Ratios

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>OptBechmark</th>
<th>Tullock-OPF</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0213</td>
<td>0.0302</td>
<td>3.5333</td>
</tr>
<tr>
<td>1</td>
<td>0.0853</td>
<td>0.0680</td>
<td>3.5450</td>
</tr>
<tr>
<td>1.5</td>
<td>0.1927</td>
<td>0.1299</td>
<td>3.5306</td>
</tr>
<tr>
<td>2</td>
<td>0.3426</td>
<td>0.2719</td>
<td>3.5300</td>
</tr>
<tr>
<td>2.5</td>
<td>0.5534</td>
<td>0.4094</td>
<td>3.5303</td>
</tr>
<tr>
<td>3</td>
<td>0.7710</td>
<td>0.5704</td>
<td>3.5303</td>
</tr>
<tr>
<td>3.5</td>
<td>1.0494</td>
<td>0.7048</td>
<td>3.5303</td>
</tr>
<tr>
<td>4</td>
<td>1.3707</td>
<td>0.8389</td>
<td>3.5303</td>
</tr>
<tr>
<td>4.5</td>
<td>1.7247</td>
<td>1.1242</td>
<td>3.5608</td>
</tr>
<tr>
<td>5</td>
<td>2.1417</td>
<td>1.5608</td>
<td>3.5303</td>
</tr>
</tbody>
</table>
Social welfare can be solved by referring to (19):

$$U^*_{n,G} = \frac{n^2}{2} - \int_{\frac{\pi}{2}}^{\pi} \int_{\frac{\pi}{2}}^{\pi} \frac{\nu^2}{(4\pi - 1)} \sqrt{\frac{\nu}{\pi}} \, d\theta \, d\nu = \left( \frac{\log 3}{16} - \frac{1}{36} \right) n^2.$$

Thus it immediately follows from the above that

$$\pi^*_n < \pi^*_n, \quad U^*_{n,G} > U^*_{n,F}, \quad \forall n \geq 2, \forall \nu > 0.$$

This tells that Population-2 is superior to Population-1 in terms of both profit and social welfare. An insight that may be drawn from this set of results is that population diversity or uncertainty is beneficial to Tullock-contest based crowdsourcing for both crowdsourcers and participants.

VI. CONCLUSION

To recap, this work has presented a first attempt to use Tullock contests as a new framework to design incentive mechanisms for crowdsourcing. Furthermore, we have explored a novel dimension in the space of optimal Tullock contest design, by superseding the conventional, fixed prize by an optimal prize function for utility maximization. In stark contrast to prior art, we have obtained an analytical solution to the unique Bayesian equilibrium, and found that the equilibrium is robust to an increasing number of rivals. As our model employs a very general contest success function and assumes incomplete information, the mechanism and results would fit a wide range of practical crowdsourcing applications. For example, WiFiScout [31] is a mobile app that aims to profile the performance of citywide WiFi access points by eliciting personal experience on WiFi usage from smartphone users. Similarly, OpenSignal [32] aims to construct citywide 3G and 4G LTE cell coverage maps through crowdsourcing too.

The superiority of our design has been demonstrated through extensive evaluations by comparing against a fully-optimized benchmark. Constructing this optimal benchmark significantly extends prior art which only solves conventional, fixed-prize Tullock contests. This benchmark would be highly relevant to a wider research community for future performance evaluations.

Moreover, we have introduced the crowdsourcer’s valuation of user contribution which further extends usual contest models (besides our prize function). It is shown to impact two key metrics—the crowdsourcer’s profit and players’ social welfare—in a nonlinear (exponential) manner, which bears practical implication on the worth of improving crowdsourcers’ business processes. Therefore, this new parameter could be included by future studies (e.g., on radio spectrum auctions or heterogeneous networks) in mathematical models to capture value asymmetry and uncover phenomena that are previously unseen.

ACKNOWLEDGMENT

This work was supported in part by the State Key Development Program for Basic Research of China (973 project 2014CB340300), in part by China NSF grant 61422208, and in part by CCF-Tencent Open Fund.

Table III: Social Welfare Comparison and Ratios

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
<th>4.5</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>OptiBenchmark</td>
<td>0.0019</td>
<td>0.0058</td>
<td>0.0155</td>
<td>0.0271</td>
<td>0.0407</td>
<td>0.0670</td>
<td>0.0813</td>
<td>0.1065</td>
<td>0.1336</td>
</tr>
<tr>
<td>Tullock-OPF</td>
<td>0.0135</td>
<td>0.0540</td>
<td>0.1215</td>
<td>0.2160</td>
<td>0.3375</td>
<td>0.4859</td>
<td>0.6614</td>
<td>0.8639</td>
<td>1.0934</td>
</tr>
</tbody>
</table>

REFERENCES


