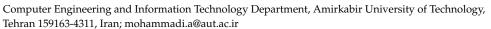*Article*

A Financial Incentive Mechanism for Truthful Reporting Assurance in Online Crowdsourcing Platforms

Alireza Mohammadi and Seyyed Alireza Hashemi Golpayegani *

Computer Engineering and Information Technology Department, Amirkabir University of Technology, Tehran 159163-4311, Iran; mohammadi.a@aut.ac.ir
* Correspondence: sa.hashemi@aut.ac.ir

Abstract: In today's world, crowdsourcing is regarded as an effective strategy to deal with a high volume of small issues whose solutions can have their own complexities in systems. Moreover, requesters are currently providing hundreds of thousands of tasks in online job markets and workers need to perform these tasks to earn money. Thus far, various aspects of crowdsourcing including budget management, mechanism design for price management, forcing workers to behave truthfully in bidding prices, or maximized gains of crowdsourcing have been considered in different studies. One of the main existing challenges in crowdsourcing is how to ensure truthful reporting is provided by contributing workers. Since the amount of pay to workers is directly correlated with the number of tasks performed by them over a period of time, it can be predicted that strong incentives encourage them to carry out more tasks by giving untruthful answers (providing the first possible answer without examining it) in order to increase the amount of pay. However, crowdsourcing requesters need to obtain truthful reporting as an output of tasks assigned to workers. In this study, a mechanism was developed whose implementation in crowdsourcing could ensure truthful reporting by workers. The mechanism provided in this study was evaluated as more budget feasible and it was also fairer for requesters and workers due to its well-defined procedure.

Keywords: truthful reporting; crowdsourcing; mechanism design

1. Introduction

The flexible and miscellaneous nature of cyberspace has led to the formation of new branches of electronic commerce (businesses) based on the participation of individuals in a series of activities that do not require a lot of resources and expertise [1–3]. Moreover, considering the use of the Internet and its capabilities as well as the power of synergy in people to solve problems collectively, there is the prospect of dealing with very complicated problems which have not so far had easier, simpler, and faster solutions, particularly those wherein human understanding has been of utmost importance. One of these problems is human sentiment analysis [4–6] in texts or images. In spite of the countless efforts made in recent years to identify sentiments in texts or images using computational methods, only the use of data-based procedures from previous experiences has been relatively fruitful in this regard. However, providing such data can involve spending a fortune that cannot be simply done.

Solving various problems through crowdsourcing is among the common businesses that have become increasingly expanded and developed within cyberspace and via the Internet. By definition, crowdsourcing refers to a sourcing and distribution model to solve a problem in which a group of individuals whose number is not known in advance is invited and encouraged through a public announcement to contribute to dealing with a problem [7,8]. In this respect, people who have a problem and want to solve it using collective participation of individuals are called requesters. In general, in the process of crowdsourcing, a requester submits a problem along with a bidding price they are willing

to pay to a crowdsourcing system. Then, the crowdsourcing system exposes the problem to workers (ones contributing to solving the problem). On the basis of their skills and the bid price, the workers take part in solving the problem.

Generally, crowdsourcing is of various types. In one type of crowdsourcing, a problem (or, in general, a task) is given to a set of devices (such as computers or cell phones) as crowdsourcing workers [9], Sensorly website (Site 10) in Appendix A, [10,11]; tasks such as monitoring road traffic [12,13]; reporting environmental effects [14], and urban noise mapping [15] are among problems solved in the form of crowdsourcing, in which the task of solving a problem is delegated to devices. Because of its inherent complexities and the need to use the power of understanding and creativity, most of the workers in crowdsourcing systems are humans. For example, Google LLC is benefiting from the collective power of human workers to label images in order to improve its search engine performance on the Google image labeler website (Site 7) in Appendix A. Almost certainly, the most common and the most popular example of crowdsourcing can be the electronic service market established by Amazon.com, Inc. The company has a million sub-tasks (micro-tasks) assigned to online workers in the form of a system called Amazon Mechanical Turk (Site 1) in Appendix A. In the domain of specialized activities, crowdsourcing can also be effective; for example, the system of the Zooniverse website (Site 12) in Appendix A. The focus of the activities of the given system is mainly on managing scientific problems in diverse fields such as biology, linguistics, aerospace, and the like, all through crowdsourcing. Even though different crowdsourcing systems, in some cases, make use of non-financial incentives such as entertainment [6,16] educational opportunities such as the Duolingo website (Site 4) in Appendix A, information sharing [17], and altruism [18] to encourage people to contribute; currently, almost all of the activities that are carried out via crowdsourcing can aid requesters and workers in having an income as a financial business, for example, the Amazon Mechanical Turk website (Site 1) in Appendix A, Upwork website (Site 11) in Appendix A, Figure Eight website (Site 5) in Appendix A, MicroTask website (Site 9) in Appendix A, Cloudcrowd website (Site 3) in Appendix A, LeadGenius website (Site 8) in Appendix A and reference [19]. Accordingly, the most effective mechanisms for urging workers to engage in crowdsourcing activities are also financial [20]. However; issues such as requester budget, time constraints, and amount of money charged by workers to perform tasks are among the important challenges that must be managed in the domain of crowdsourcing. According to what was mentioned above and taking the money-related relationship between requesters and workers into account, simultaneous attempts made by requesters and workers in order to maximize their earnings from crowdsourcing activities can be understandable. Moreover, since the main parameters for pays to each worker are the number of completed sub-tasks (micro-tasks) by them, a worker is likely to seek for fulfilling a large number of sub-tasks (micro-tasks) during a fixed or lower time. In this regard, the speed of a task is considered by a worker and the accuracy of the quality of the answers given to various problems can be reduced [21]. This situation is certainly not ideal for requesters and will cause concerns. On the other hand, it is possible that requesters will also lessen pays to workers and some of them may not give a positive response to correct answers delivered [22]. Therefore, such probable reactions will be a worry for workers. In this respect, features and capabilities created in crowdsourcing service provider systems can play an important role in managing the given challenges and concerns. Despite the increasing number of tasks performed in crowdsourcing due to augmented financial incentives, the quality of performing these tasks will not increase [23]. Therefore, it is necessary to employ inhibiting and corrective mechanisms to influence the incentives of workers during tasks and lead these individuals to perform high-quality ones. Along with the cited mechanisms, there is a need to have a system for validating answers provided by workers. Some research has been conducted to resolve these challenges with a focus on issues such as mechanism design to encourage workers to demonstrate truthful behaviors [21], introduction of a process of task allocation [24], as well as presentation of pricing models [19,25].

In the present study, a mechanism was addressed whose implementation in the process of crowdsourcing could maximize workers' earnings provided that truthful answers were given to requests and tasks were fulfilled with high quality. The proposed mechanism would be in line with the requester budget and it could be also fair for workers. As well, providing a transparent validation system for answers given by workers to tasks as well as the possibility of improper rejection of answers by requesters would be reduced and; at the same time, there could be enough flexibility to adapt to the limitations of requesters and workers. This study was organized into four main sections; following the introduction section, similar and related previous works were reviewed; then, the history of challenges in providing wrong answers by workers to further tasks as one of the major unsolved problems in the procedures and mechanisms of crowdsourcing was provided. In the following section, the proposed mechanism and its confirmation in an analytic manner was presented and its effect on the prevention of providing wrong answers to tasks in order to increase the number of tasks to gain more profits was reviewed and evaluated. Other aspects of the proposed mechanism, such as the budget and pricing estimation process were discussed. In the final section, a conclusion was provided.

2. Related Works

Although crowdsourcing is considered a newly-fledged approach to problem-solving, numerous studies have been conducted so far on its various aspects. In this section, there were attempts to review the most important investigations closely associated with the subject of the study. The 10 articles reviewed in this section are chosen with the purpose of providing a broad view of all theoretical and practical works carried out in the area of incentivization of honesty and integrity in crowdsourcing. Therefore, this section includes articles on a wide variety of subjects from pricing mechanisms to reputation-based motivational protocols.

The major approach in the majority of the related investigations was providing a mechanism for crowdsourcing [10,11,19,24,26–28]. In addition, in a few cases, approaches such as prediction [21], algorithm design [29], and model contest as auction [30] had been used as the basis for the research. It is important to note that all of the approaches presented in these works had been focused on pricing and payment strategies. Another point about these research studies was recognizing humans as workers in crowdsourcing in a way that most of these investigations had assumed that the tasks were being done by humans, though in some problems such as crowd mobile sensing, the main tools to collect data in crowdsourcing were smart phones [10,11]; but the owner of the tools, that is, a human being was the decision-maker and the main contributor to crowdsourcing.

Research studies conducted in the field of crowdsourcing could be reviewed in terms of the type of answers given by workers to tasks defined by requesters. Typically, there were two types of answers for tasks assigned to workers; the first type was a single correct answer to a requested task. This type of answer was provided in crowdsourcing wherein each task had only one correct answer. An example would be the issue of face categorization. In this case, workers were asked to choose only images smiling in a set of facial images with different modes. The second type of answer referred to a series of correct answers to a requested task. This type of answer was related to a type of crowdsourcing in which each of the tasks defined could have more than one correct answer; for example, crowdsourcing translation of a poem from English into other languages. It is obvious that a poem can have more than one correct translation. Accordingly, research in the field of crowdsourcing could be divided into two groups in terms of answers provided.

The features of an approach encouraging workers and requesters to use it could be another aspect examined in research studies in this domain. From the point of view of workers, one of the important characteristics of a collective crowdsourcing approach was fairness assurance which meant giving equal chance to each worker to maximize gains by contributing to crowdsourcing. An example of the lack of fairness could be a case in which a requester would claim that answers provided by some of the workers were not of

J. Theor. Appl. Electron. Commer. Res. **2021**, 16

2017

good quality after the end of crowdsourcing and workers were not paid. As mentioned earlier, the proposed approach needs to prevent this kind of exploitation of labor (worker abuse). Another important feature of a crowdsourcing approach from the perspective of workers was the attention given by an approach to their limitations and capabilities. Assuming that the ability of workers to solve various problems and to perform assigned tasks is not equal, their limitations and capabilities needed to be considered in the proposed approach. From the perspective of requesters, several issues were of utmost importance in crowdsourcing. One of the most important features of each approach in the view of requesters was taking financial limitations into account. To be more exact, there was the possibility of controlling the routine of crowdsourcing budget allocation and matching the rules of payment to workers with the requester budget. Another significant characteristic for requesters was to maximize the number of tasks assigned and performed by workers based on a specific budget. However, the main concern in requesters was the quality of performing tasks by workers. Considering the willingness of workers to do more tasks and to increase the speed of performing tasks in order for maximizing earnings, workers were likely to give not much importance to performing the tasks truthfully. In [21], the researchers showed that, in some cases, and not all, the mechanism provided by them had boosted the incentives for workers to have truthful reporting. In these studies, the term "truthful reporting assurance" had been used for cases wherein workers had their own bid price for a task. In these investigations, after bidding prices by workers and selecting the winning one, there was no assurance to deliver a good-quality answer for the assigned tasks and also to provide truthful reporting.

In Table 1, 10 main articles proposing an approach to truthful crowdsourcing were reviewed. As shown in the Table, half of these studies simulated their proposed approach, and the other half implemented it in the real world. As can be seen, none of the research studies focused on truthful reporting and they merely shed light on other aspects of crowdsourcing. In the next section, a mechanism was provided that not only maintained the advantages of the approaches presented in previous research works but also assured truthful reporting by workers.

J. Theor. Appl. Electron. Commer. Res. **2021**, *16*

2018

Table 1. Previous research and literature.

Ref. #	Main Approach	Worker Type	Answer Type	Incentives for Requesters	Incentives for Workers	Worker Best Answer (Truthful Reporting)	Case Study
[21]	Consensus Prediction Payment Rules	Human	single/Set of answer(s)	N/A	Fairness	No (Promote truthful reporting In Some Case)	Simulation
[19]	Design Mechanism for regret minimization	Human	single/Set of answer(s)	Budget Feasible Near Optimal Util	Profitable	No	Simulation/Real Word (Amazon Mturk)
[24]	Design Incentive-Compatible Mechanism (TM-Uniform)	Human	Single Answer	Budget feasible Near optimal Util	Profitable Mach Constraint	No	Simulation/Real Word (Wikipedia translation on Amazon Mturk)
[10]	Designing platform/user centric incentive mechanisms (Mechanisms using Stakelberg game for Platform Centric Model and auction-based incentive mechanism for user-centric Model)	Human (Mobile Phones Sensing)	Single Answer	Budget Balance Utility	Profitable	No	Simulation
[11]	Design Incentive Online Mechanisms (Very similar to (Site 10) in Appendix A and [13])	Human (Mobile Phones Sensing)	Single Answer	Budget Balance	Profitable	No	Simulation
[29]	An algorithm motivated by PAC Learning	Human	Single Answer	Maximize Revenue	-	No	-
[28]	Design Multiple Mechanisms	Human	single/Set of answer(s)	maximizing the number of tasks performed under budget. minimizing payments for a given number of tasks	Profitable	No	Implement of Mechanical Perk Framework

J. Theor. Appl. Electron. Commer. Res. **2021**, *16*

2019

Table 1. *Cont.*

Ref. #	Main Approach	Worker Type	Answer Type	Incentives for Requesters	Incentives for Workers	Worker Best Answer (Truthful Reporting)	Case Study
[27]	Design Mechanism (Incentive protocols/interaction between worker and requester with repeated game)	Human	single/Set of answer(s)	No Free-riding Maximize revenue	No False-Reporting	No	Simulation
[30]	Model Contest As All-Pay Auction	Human	single/Set of answer(s)	-	-	No	Simulation/Test in Taskcn.com
[26]	Design Mechanism for optimal price for task	Human	single/Set of answer(s)	budget feasibility competitive ration performance	Profitable	No	Test in Mechanical Turk

3. Methodology

This study presents a mechanism based on monetary incentivization for motivating honest participation of workers in crowdsourcing. This mechanism controls the work quality by determining how well the crowdworkers should be paid. The profit of a crowdworker is maximized when he or she honestly contributes to the crowdsourcing effort. Here, the measure of honesty is a comparison between the output (report) of each crowdworker and those of other participating crowdworkers, who of course must not know each other. After presenting the mechanism and demonstrating that it guarantees the honesty of crowdworkers, the result of the implementation of this mechanism in a website will be presented. The following sections first describe the proposed mechanism, then discuss its reliability and validity, and finally present the details and results of experimental testing.

3.1. Mechanism Description

Within the mechanism proposed in this study, workers could select an answer to a task in the form of a value in a range to solve the problem of truthful reporting in crowdsourcing. Although this is the most common method of crowdsourcing, explaining the proposed mechanisms revealed that it was also applicable to other methods. The main logic behind the proposed mechanism was a judgment about a reported answer to a task by workers through the average of the reports provided by other workers participating in crowdsourcing for that task.

The assumptions needed to explain the proposed mechanism were discussed below.

- The answer by workers to crowdsourcing is in the form of providing a number in a range. The following statement can be the answer given by worker i to task k:

$$Report_i^k$$

- The answer that worker i considers true for task k and believes in it; more accurately, the right answer from the view of worker i to task k can be illustrated in the following statement:

$$Ans_i^k$$

- The goal is to provide an answer by a worker to a task they believe in:

$$Ans_i^k = Report_i^k$$

- For each task, a set of workers report their answers. The average of workers' reports is then computed. The average of answers provided by workers participating in crowdsourcing to task k will be shown in the following statement:

$$Avg_k = \frac{\sum_{i=1}^{n} Report_i^k}{n}$$

- Utility of worker i when they answer to task k can be also displayed with the following statement:

$$Util_i^k$$

- For each task, a worker can receive utility with these rules:
  - If the total distance of answers provided by contributing workers to a task from the average of the answers given to that task is less than a certain amount, it indicates collusion or excessive simplicity of a question; so, a worker is randomly selected

and they will receive a fraction of utility related to that question and others will not receive anything:

$$if\left(\sum_{i=1}^{n}\left|Report_i^k - Avg_k\right| \leq \xi\right) Then \ Util_{Random_i}^k = \frac{1}{L} \ Util_{Other \ Except \ i}^k = 0$$

- ○   otherwise, their utility will be calculated via the following relation:

$$if\left(\sum_{i=1}^{n}\left|Report_i^k - Avg_k\right| > \xi\right) Then \ Util_i^k = 1 - \frac{\left|Report_i^k - Avg_k\right|}{Avg_k}$$

- Upon the completion of crowdsourcing, the earnings of worker i will be computed as follows:

$$Payoff_i = \sum_{j=1}^{K} Util_i^j$$

The most frequently used notations were listed in Table 2.

Table 2. Frequently used notations.

Symbol	Meaning
$Report_i^k$	Answer of worker i reports to question k
Ans_i^k	Answer which believes worker i for question k
Avg_k	$\frac{\sum_{i=1}^{n} Ans_i^k}{n}$, Average of answers to question k by n worker
$Util_i^k$	Utility of worker i from question k
L	Factor of Utility in case of collusion or easy question.
$Payoff_i$	$\sum_{j=1}^{K} Util_i^j$, the payoff of worker i after performing a task in the crowdsourcing process
Budget	Budget of the requester for crowdsourcing which will pay to workers.
Price	Price of each Task performed by a worker during the crowdsourcing process.

3.2. Proposed Mechanism Equilibrium

To account for equilibrium in the proposed mechanism, Table 3 was considered. In this utility table, it was required to know that ξ was a threshold indicating how much an answer provided to task k by worker i could be close to the average of answers given to the same task by other workers.

Table 3. Utility Table.

	$\sum_{i=1}^{n}\left\|Report_i^k - Avg_k\right\| > \xi$	$\sum_{i=1}^{n}\left\|Report_i^k - Avg_k\right\| \leq \xi$
$Report_i^k = Ans_i^k$	$Util_i^k{}_{11}$	$Util_i^k{}_{12}$
$Report_i^k \neq Ans_i^k$	$Util_i^k{}_{21}$	$Util_i^k{}_{22}$

Consider the following rules:

1. Each worker can be selfish and willing to maximize their earnings from participating in crowdsourcing. So, they will seek for maximizing their utility on each task.
2. Each worker knows a correct answer or at least an answer believed to be correct or even close to a correct answer.
3. Each answer provided by each worker to each task is sealed and workers are not aware of each other's answers to a task; so, they will not be informed of answers given to a task.

4. Each worker understands that other workers know these rules.

- In Case of $\sum_{i=1}^{n} \left| Report_i^k - Avg_k \right| > \xi$:

Utility of worker i from answer to task k is equal to $Util_i^k = 1 - \frac{\left| Report_i^k - Avg_k \right|}{Avg_k}$. According to Rule 1, every worker is willing to maximize one's utility; therefore, the goal of each worker is to provide the closest report to the average one:

$$Report_i^k \rightarrow Avg_k$$

Considering this goal and according to Rule 3, a worker can have two strategies to report on a task assigned. $Report_i^k = Ans_i^k$ or $Report_i^k \neq Ans_i^k$; taking Rule 2 and Rule 4 into account, the strategy of $Report_i^k = Ans_i^k$ will fulfill the goal of a worker and as a result $Util_i^k 11 > Util_i^k 21$.

- In Case of $\sum_{i=1}^{n} \left| Report_i^k - Avg_k \right| \leq \xi$:

In this case, the utility of worker i in an answer to task k is not out of two modes, either 0 or with a probability of $\frac{1}{Number\ of\ workers}$, the amount will be $\frac{1}{L}$ and choose one of the two strategies $Report_i^k = Ans_i^k$ or $Report_i^k \neq Ans_i^k$ will not influence the $Util_i^k$.

Finally, it was concluded that truthful reporting by each worker can benefit other workers and this strategy will weakly dominate the strategy of untruthful reporting. It is obvious that the Nash equilibrium in this model can be realized through truthful reporting by workers. What follows outlines the use of different rules to calculate the utility of workers whose answers are closer to the average from one threshold. When an answer to a task given by a worker is very close to the average answer given by other workers to the same task, there is the probability of the occurrence of two scenarios. The first scenario is related to the issue that a question is very simple; in this case, part of the utility of a task is randomly paid to a worker, and other workers contributing to this task will not have a utility; thereby, fair payment policy can be justified. The second scenario is the occurrence of collusion in workers. Although there is no possibility of collusion due to the fact that workers contributing to a task are not aware of the identity of other workers in their answers to the same task; this rule of utility calculation is considered to prevent various forms of collusion and also to eliminate incentives for workers to collide. Given this rule, if a worker provides an answer through colliding with others, their utility will be $\frac{1}{Number\ of\ workers}$, with the probability of $\frac{1}{L}$. This amount of utility is equal to the extent to which an answer is given without collusion and also an answer is close to the average; in other words, the task requested is very simple. If a requested task is not simple and it is calculated with the first rule of utility, given the fact that the report is presented truthfully; the utility can be maximized. If a truthful answer is provided, more utility than collusion will be obtained.

3.3. Proposed Mechanism for Budget and Pricing Estimation

In crowdsourcing, requester budget management is of great importance. There are two general methods to managing budgets in crowdsourcing. The first method as the most common mode of crowdsourcing is considering a fixed budget by requesters. In this case, a requestor bids the price of a task, and workers accept or reject the task according to their ability and the price declared by the requester. Mainly, in this mode, no dialogue and bargaining will occur between requesters and workers. The second method is that a requester first introduces the requested task and then asks workers to bid prices for the task. In this respect, some mechanisms have been designed for workers' truthful bidding as cited in the related works section.

3.3.1. Fixed Budget Methods

In the fixed budget method, a requester introduces a task and bids a price for it. Workers can either accept or reject their request according to the bid price. In the mechanism

provided to estimate the bidding price, the upper and lower bounds of the total earnings of all workers need to be initially calculated:

$$\sum_{i=1}^{n} Payoff_i$$

The minimum of total earnings for all workers is when the utility of all workers are computed for doing all tasks with the rule of simple question or avoiding collusion, which will be equal to the following value:

$$\frac{k}{L}$$

The maximum of total earnings for workers will occur when a worker receives utility 1 for each task $n - 1$ and another worker receives at a rate of $1 - \frac{|1+ \xi|}{\text{Avg}}$. This value is close to 1, so the maximum amount of payment will be lower than the following value for task k:

$$nk$$

Therefore, the upper and lower bounds of the total earnings of workers will be as follows:

$$\frac{k}{L} \leq \sum_{i=1}^{n} Payoff_i < nk$$

It is clear that the bid prices for each task will be as follows:

$$Price = \frac{Budget}{\sum_{i=1}^{n} Payoff_i}$$

So, there will be:

$$\frac{Budget}{nk} < Price \leq \frac{L * Budget}{k}$$

Therefore, it was concluded that the bid price by the requester should be within the above range and it is obvious that it is not possible for the requester to bid a price higher than $\frac{L* \text{Budget}}{k}$.

3.3.2. Bidding Price Method

In a method in which workers bid their prices to perform a task, a requestor must be able to estimate the total amount of budget required for crowdsourcing. As cited in the previous section:

$$Budget = Price * \sum_{i=1}^{n} Payoff_i$$

Therefore;

$$Price * \frac{k}{L} \leq Budget < Price * nk$$

Accordingly, in order to assure the risk-free financing of crowdsourcing, the requester budget must be at a rate of $Price * nk$.

3.4. Reliability and Validity

In crowdsourcing, the quality of the output provided by crowdworkers has a direct relationship with the output demanded by the employer. Thus, for any given crowdsourced task, the closer is the crowdworker's viewpoint to the employer's, the higher is the output quality. Therefore, it is customary for crowdsourcing employers to provide a sample of desired outputs for a number of tasks to show crowdworkers what outputs to produce. After this training, the employer gives the main crowdsourcing tasks to crowdworkers, who then generate outputs as requested.

Using the same rationale, we have developed a measure for evaluating the quality of outputs produced by crowdworkers, which can also be used to assess the mechanism presented in this study.

It should be noted that in the proposed mechanism, the output of a crowdworker such as *i* for a task like k is a number in a certain range:

$$0 \leq Report_i^k \leq Upper\ Band\ of\ Range$$

Also, the proposed mechanism assumes that the final output for a crowdsourced task like k is the average of all outputs obtained from all crowdworkers participating in that task, which is in the same range as the outputs produced by crowdworkers:

$$0 \leq Avg_k \leq Upper\ Band\ of\ Range$$

The error of the mechanism for each crowdsourced task is defined as follows:

$$Error_k = |\ Avg_k - Expected\ Answer\ from\ Requester\ to\ The\ Task\ k\ |$$

Therefore, the score of the mechanism for each crowdsourced task like k is defined as follows:

$$Score_k = Upper\ Band\ of\ Range - Error$$

And the efficiency of the mechanism is defined as follows:

$$Efficiency = \frac{\sum_{k=1}^{number\ of\ Tasks} Score_k}{Upper\ Band\ of\ Range * number\ of\ Tasks}$$

It is obvious that:

$$0 \leq Efficiency \leq 1$$

An efficiency value that is closer to 1 indicates that the mechanism has been more successful in convincing crowdworkers to provide honest outputs as requested by the employer.

3.5. Adopting the Proposed Mechanism to Another Form of Crowdsourcing

In this section, the flexibility of the mechanism proposed for employing other forms of crowdsourcing was addressed. Considering the main logic behind the proposed mechanism wherein a report and an answer provided by each worker to a requested task is evaluated and judged via the average of the answers of other workers to the same request, applying this mechanism to other forms of crowdsourcing can be simple.

In crowdsourcing wherein there is only one correct answer for a task like optical character recognition (OCR) evaluation (selection of images with specific content and so on), the utility is maximized when a worker chooses an option that other workers have already selected. If the options chosen by workers are numerous, a utility split policy can be used to implement the utility in a mechanically weighted manner between workers. In crowdsourcing wherein workers are required to perform specific operations; for example, translate a text, two types of workers can be used. Workers in the first group perform the task, translate the text in this example, and the second group scores the assigned task. The utility of the workers in the first group is the average answer given by the workers in the second group and the utility of the workers in the second group will be calculated according to the proposed mechanism taking distance from the average into account.

It can be observed that the overall application of the proposed mechanism in various forms of crowdsourcing will be very simple and using little creativity.

J. Theor. Appl. Electron. Commer. Res. **2021**, 16

2025

3.6. Experimental Results

To evaluate the efficiency of the proposed mechanism, we used an online platform to crowdsource a set of tasks once with this mechanism in place and another time without this mechanism and then compared the results.

Before starting the evaluation, first, we had to find or design a problem consisting of multiple tasks for crowdsourcing. These tasks had to be chosen so that crowdworkers would have a common understanding of what is demanded by the employer. Ultimately, the problem of text sentiment analysis was chosen for this purpose, as humans tend to have a similar understanding of sentiments implicit in a text. We classified the text sentiments into five categories:

{*Very Negative*, *Negative*, *Nature*, *Positive*, *Very Positive*}

Since the proposed mechanism assumes that the output of crowdworkers is a number in a certain interval, this interval was chosen to be (1, 100). Accordingly, it was decided that the outputs of crowdworkers have the following interpretations:

- 1–20: Very Negative
- 21–40: Negative
- 41–60: Nature
- 61–80: Positve
- 81–100: Very Positive

The texts needed for crowdsourced tasks were obtained from the Enron Email Dataset, which is publicly available on the website of Carnegie Mellon University (Site 2) in Appendix A. For each sentiment category, 28 short emails with less than 10 sentences were extracted from this dataset (a total of 140 emails were extracted). The 28 emails (per category) were used as described below:

- Five emails were attached to the output needed by the requesters (the authors of this paper) and were used to train crowdworkers. Hence, before crowdsourcing, each crowdworker received 25 emails (5 for each sentiment category) along with the employer's opinion on the sentiment category to which they must be allocated.
- Three emails were used in an entrance test for the crowdsourcing process. In this test, which was performed after the training, each crowdworker was asked to read 15 emails (3 per category) and rate the sentiment implicit in each email with a score from 1 to 100. The obtained outputs were then compared with the outputs produced by the employer. Only the persons who had at least 12 correct outputs and at most 1 incorrect output per category were allowed to participate in the main crowdsourcing process. Note that prospective crowdworkers participated in this test only to prove their eligibility for the main task and were not financially compensated for this participation.
- Twenty emails were used as the main crowdsourced tasks. Each participant had to rate 100 emails in total (20 per category). This stage was similar to the previous one except that crowdworkers were paid to participate. The outputs obtained from this stage were used to evaluate the efficiency of the proposed mechanism.

It should be noted that since Enron Email Dataset is in the form of MySQL scripts, retrieving 140 emails that would be evenly distributed over five sentiment categories required coding and data mining and also manually reviewing and categorizing a large number of emails.

All of the above procedures were implemented on a website, where prospective crowdworkers were asked to register and participate in the training phase and the entrance exam, and those who were found to be eligible were allowed to participate in the main crowdsourcing process.

To test the efficiency of the proposed mechanism, in the first phase, the tasks were given to 10 eligible crowdworkers, who were paid for each output provided. The result of this crowdsourcing method is provided in Table 4.

Table 4. Test Result—Labeling emails without proposed mechanism.

		Number of Emails in Each Category Labeled by Requesters				
		Very Negative	**Negative**	**Nature**	**Positive**	**Very Positive**
Number of emails in each category labeled by Participants	Very Negative	20	18	19	17	18
	Negative	0	1	0	0	0
	Nature	0	1	1	2	
	Positive	0	0	0	1	1
	Very Positive	0	0	0	0	1

The efficiency of this crowdsourcing method was found to be about 50%. As can be seen, the majority of outputs fall into the "Very Negative" category. This could be because crowdworkers had to enter their response in a text box embedded in the website and to quickly finish the tasks, most of them entered a random one-digit number in this text box, which put the outputs in this category. Also, the average time spent by each participant on each task is less than 5 s, which indicates that crowdworkers have not honestly participated in the crowdsourcing effort and have ignored the employer's request.

In the next phase, crowdsourcing was repeated with the proposed motivation mechanism with 10 other crowdworkers. In this phase, the description of the mechanism was posted on the front page of the website, and crowdworkers were allowed to register only after reading and accepting its terms. The result of crowdsourcing with the proposed mechanism is presented in Table 5.

Table 5. Test Result—Labeling emails under proposed mechanism.

		Number of Emails in Each Category Labeled by Expert				
		Very Negative	**Negative**	**Nature**	**Positive**	**Very Positive**
Number of emails in each category labeled by Participants	Very Negative	19	1	0	0	0
	Negative	1	18	1	0	0
	Nature	0	1	19	2	0
	Positive	0	0	0	18	1
	Very Positive	0	0	0	0	19

The efficiency of crowdsourcing with the proposed mechanism was 98%. In this phase, the average time spent on each task was 53 s, which means participants have acted honestly and in accordance with the employer's wishes.

4. Discussion

The main purpose of this study was to provide a mechanism for crowdsourcing on online platforms that would guarantee the honest participation of crowdworkers. As discussed in the second section, previous works on this subject have not provided a way to ensure that crowdworkers are honest in their contribution to a crowdsourcing effort. Although some of the previous works have made great breakthroughs in this area by using reputation-based protocols, regret minimization mechanisms, and other solutions to increase the incentive for honest participation, none of them induce a self-controlling behavior among crowdworkers to ensure their honest contribution. The main difference between the proposed mechanism and previous works in this area is that crowdworkers earn more when they act honestly and miss the profit when they do not. Since the purpose of participating in crowdsourcing is to earn income, the proposed mechanism guarantees honest participation in crowdsourcing through monetary incentivization.

Another issue discussed in previous related works, which is also addressed in this study, is the question of fairness of payments to crowdworkers. Crowdworkers should be paid according to their abilities and also a crowdworker who provides high-quality outputs should be compensated more than those who do not. Note that the term "quality" in crowdsourcing refers to how well the output produced by the crowdworker conforms to the demand of the employer. Previous works have suggested multiple methods for increasing fairness in crowdsourcing, but the mechanism proposed in this study is fair by nature, as a crowdworker's earning from an output produced for a task depends on the distance of this output from the average of outputs provided by other crowdworkers for the same task (the shorter the distance, the greater the earning). Because of the incentivization of honest contribution to crowdsourcing in the proposed mechanism, the average of the outputs is likely to be very close to what is requested by the employer. To summarize, the more honest a crowdworker is in producing output, the higher is the quality of that output and the greater is the payment made for that output. In this way, the mechanism guarantees fairness in crowdsourcing.

Another regularly discussed topic about crowdsourcing is how to price the tasks given to crowdworkers and how to manage crowdsourcing budgets. Previous studies in this area have discussed task pricing and budget management from two perspectives. In the first perspective, the budget is fixed, which means the employer has a certain amount of financial resources to solve the problem through crowdsourcing, and these resources should be allocated based on the number of tasks, the number, and ability of crowdworkers, etc. Several studies in this area have provided budget allocation methods for crowdsourcing projects and examined the problem of having an incentive mechanism in the presence of budget constraints. In the second perspective, crowdworkers bid on a problem consisting of multiple tasks or even bid on each task individually. In this case, the amount of budget needed to solve the problem by crowdsourcing depends on the sum of bids made by crowdworkers. A number of studies in this area have been focused on the process of bidding, bidding price, and how winners are determined.

Another group of studies have examined the aforementioned issues from a more general perspective. Since the nature of the problems being crowdsourced is such that budget management should be possible with both fixed budget and bidding mechanisms, methods that ignore either of these two approaches are not applicable for a wide range of crowdsourcing problems. The mechanism presented in this study can guarantee the integrity of crowdworkers whether crowdsourcing is done with a fixed budget or through bidding. If the crowdsourcing has a fixed budget, then the price range of each task can be estimated according to that budget, and if it is bid-based, then the budget needed for the project can be estimated according to the bids.

The mechanism presented in this paper not only ensures the honest participation of crowdworkers but also guarantees fairness in crowdsourcing and facilitates crowd-sourcing budget management. The proposed mechanism operates based on simple and self-explanatory rules focused on maximizing the earning of honest crowdworkers, and is easy to implement in online crowdsourcing platforms.

5. Conclusions and Future Works

This study presented a crowdworker payment mechanism that maximizes the revenue earned by crowdworkers for their honest contribution to the crowdsourcing effort. The core idea of the proposed mechanism is to compare the outputs of each crowdworker for each task with the average output of other crowdworkers for the same task, on the condition that crowdworkers are not affiliated with each other. In this mechanism, an output given by a crowdworker that is closer to the average of other outputs for the same task is assumed to be of higher quality and will be awarded higher compensation. It was proved that the honest conduct of crowdworkers weakly dominates their dishonest conduct. The proposed mechanism consists of simple rules which can be easily implemented in online crowdsourcing platforms. To test this mechanism, the authors designed a crowdsourcing

J. Theor. Appl. Electron. Commer. Res. **2021**, *16*

2028

website and used it to crowdsource a problem consisting of multiple tasks with and without the proposed mechanism. This test showed that without the mechanism, many crowdworkers engaged in dishonest behavior to exploit the earning opportunity, but the presence of the mechanism led to honest participation of crowdworkers in crowdsourcing.

The main contribution of this paper to the literature is a solution to the problem of dishonesty in the conduct of crowdworkers, as the proposed mechanism ensures that honest conduct is awarded appropriately. At the same time, the mechanism also addresses two other key issues in crowdsourcing, namely the fairness of compensations, which is one of the major concerns of crowdworkers, and the issue of budget management in crowdsourcing, which along with the quality of outputs produced for crowdsourced tasks, is one of the major concerns of crowdsourcing employers.

The implication of this research is the introduction of an easily applicable solution for online crowdsourcing platforms to ensure honest participation of crowdworkers in crowdsourcing and guarantee the quality of their outputs for the assigned tasks, which enhances the confidence of employers in the quality of outputs that can be obtained from crowdsourcing.

The most fundamental limitation of the proposed mechanism is the extra charge incurred by the employer. Since the proposed mechanism involves giving each task to multiple crowdworkers and treating the average of the outputs they provide as the best output, and then compensating them according to the proximity of their output to this average, the employer has to pay multiple compensations for each task. Naturally, this extra charge makes crowdsourcing less desirable, but it can also be interpreted as the cost the employer pays to ensure that the crowdsourced tasks are accomplished at high quality.

The most important objective of future works should be to develop a mechanism for ensuring the truthfulness of crowdworkers in crowdsourcing without needing additional crowdworkers for each task so that the employer does not have to pay an extra cost to ensure high quality. Another objective worth pursuing in future works is to develop a method for dynamic estimation of optimal ξ value, which is the threshold determining the rule based on which crowdworkers are compensated. This estimation is invaluable because the proposed mechanism operates based on how crowdworkers are compensated for their work, which itself strongly depends on the ξ value.

Author Contributions: A.M.: Conceptualization, methodology, software, formal analysis, data curation, writing—original draft preparation, writing—review and editing. S.A.H.G.: validation, writing—review and editing, supervision. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: To get access to the data used during the study, please contact the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Site 1: Amazon Mechanical Turk
Available online: https://www.mturk.com (accessed on 22 July 2021)
Site 2: Carnegie Mellon University—Enron Email Dataset
Available online: https://www.cs.cmu.edu/~enron/ (accessed on 22 July 2021)
Site 3: Cloudcrowd
Available online: https://www.cloudcrowd.com (accessed on 22 July 2021)
Site 4: Doulingo
Available online: http://duolingo.com/ (accessed on 22 July 2021)
Site 5: Figure Eight, Machine Learning and Artificial Intelligence Platform for High Quality Training Data
Available online: https://www.figure-eight.com (accessed on 22 July 2021)

Site 6: Crowd flower, The Essential High-Quality Data Annotation Platform
Available online: https://www.crowdflower.com (accessed on 22 July 2021)
Site 7: Google, Google Image Labeler
Available online: https://crowdsource.google.com/imagelabeler/ (accessed on 22 July 2021)
Site 8: LeadGenius, Home—Custom B2B Contact and Account Data
Available online: https://www.leadgenius.com/ (accessed on 22 July 2021)
Site 9: MicroTask
Available online: http://microtask.com (accessed on 22 July 2021)
Site 10: Sensorly, Unbiased, Real-World Mobile Coverage
Available online: http://www.sensorly.com (accessed on 22 July 2021)
Site 11: Upwork, Hire Freelancers & Get Freelance Jobs Online
Available online: https://www.upwork.com (accessed on 22 July 2021)
Site 12: Zooniverse
Available online: https://www.zooniverse.org (accessed on 22 July 2021)

References

1. Edelman, B.; Ostrovsky, M.; Schwarz, M. Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars' Worth of Keywords. *Am. Econ. Rev.* **2007**, *97*, 242–259. [CrossRef]
2. Varian, R. Position auctions. *Int. J. Ind. Organ.* **2007**, *25*, 1163–1178. [CrossRef]
3. Lahaie, S.; Pennock, D.M.; Saberi, A.; Vohra, R.V. Sponsored Search Auctions. *Algorithmic Game Theory* **2011**, *1*, 699–716.
4. Benkler, Y. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*; Yale University Press: New Haven, CT, USA, 2006.
5. Malone, T.W.; Laubacher, R.; Dellarocas, C. Harnessing Crowds: Mapping the Genome of Collective Intelligence. MIT Sloan Research No.4732-09. 2009. Available online: https://ssrn.com/abstract=1381502 (accessed on 24 July 2021).
6. Ahn, L. Games with a purpose. *Computer* **2006**, *39*, 92–94. [CrossRef]
7. Chatzimilioudis, G.; Konstantinidis, A.; Laoudias, C.; Zeinalipour-Yazti, D. Crowdsourcing with Smartphones. *IEEE Internet Comput.* **2012**, *16*, 36–44. [CrossRef]
8. Howe, J. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*; Crown Publishing Group: New York, NY, USA, 2008.
9. Yang, D.; Xue, G.; Fang, X.; Tang, J. Crowdsourcing to smartphones:incentive mechanism design for mobile phone sensing. In Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, Istanbul, Turkey, 22 August 2012.
10. Zhang, X.; Xue, G.; Yu, R.; Yang, D.; Tang, J. Truthful Incentive Mechanisms for crowdsourcing. In Proceedings of the IEEE Conference on Computer Communications, Hong Kong, 26 April 2015.
11. Zhao, D.; Li, X.-Y.; Ma, H. How to crowdsource tasks truthfully without sacrificing utility: Online incentive mechanisms with budget constraint. In Proceedings of the IEEE INFOCOM 2014-IEEE Conference on Computer Communications, Toronto, ON, Canada, 27 April–2 May 2014.
12. Mohan, P.; Padmanabhan, V.; Ramjee, R. Nericell: Rich monitoring of road and traffic conditionsusing mobile smartphones. In Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems, Raleigh, NC, USA, 5 November 2008.
13. Thiagarajan, A.; Ravindranath, L.; LaCurts, K.; Madden, S.; Alakrishnan, H.; Toledo, S.; Eriksson, J. Vtrack: Accurate, energy-aware road traffic delay estimation using mobile phones. In Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, Berkeley, CA, USA, 4 November 2009.
14. Mun, M.; Reddy, S.; Shilton, K.; Yau, N.; Burke, J.; Estrin, D.; Hansen, M.; Howard, E.; West, R.; Boda, P. PIER, the personal environmental impact report, as a platform for participatory sensing systems research. In Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services, Kraków, Poland, 22 June 2009.
15. Rana, R.; Chou, C.; Kanhere, S.; Bulusu, N.; Hu, W. Earphone: An end-to-end participatory urban noise mapping. In Proceedings of the 10th International Conference on Information Processing in Sensor Networks, Chicago, IL, USA, 12 April 2010.
16. Ahn, L.; Dabbish, L. Designing games with a purpose. *Commun. ACM* **2008**, *51*, 58–67. [CrossRef]
17. Jain, S.; Chen, Y.; Parkes, D.C. Designing incentives for online question and answer fórums. In Proceedings of the 10th ACM conference on Electronic Commerce, Innsbruck, Austria, 19 August 2009.
18. Cooper, S.; Khatib, F.; Makedon, I.; Hao, L.; Barbero, J.; Baker, D.; Fogarty, J.; Popovic, Z.; Players, F. Analysis of social game-play macros in the foldit cookbook. In Proceedings of the 6th International Conference on Foundations of Digital Games, Bordeaux, France, 29 June 2011.
19. Singla, A.; Krause, A. Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. In Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 13 May 2013.
20. Shaw, A.D.; Horton, J.J.; Chen, D.L. Designing incentives for inexpert human raters. In Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, Hangzhou, China, 19 March 2011.

21. Kamar, E.; Horvitz, E. Incentives and Truthful Reporting in Consensus-Centric Crowdsourcing. 2012. Available online: https://www.microsoft.com/en-us/research/publication/incentives-and-truthful-reporting-in-consensus-centric-crowdsourcing/ (accessed on 24 July 2021).

22. Feldman, M.; Papadimitriou, C.; Chuang, J.; Stoica, I. Free-riding and Whitewashing in Peer-to-Peer Systems. In Proceedings of the ACM SIGCOMM Workshop on Practice and Theory of Incentives in Networked Systems, Portland, OR, USA, 3 September 2004.

23. Mason, W.; Watts, D. Financial incentives and the performance of crowds. *ACM SIGKDD Explorations Newsl.* **2009**, *11*, 100–108. [CrossRef]

24. Goel, G.; Nikzad, A.; Singla, A. Allocating tasks to workers with matching constraints: Truthful mechanisms for crowdsourcing markets. In Proceedings of the 23rd International Conference on World Wide Web, Seoul, Korea, 7 April 2014.

25. Chawla, S.; Hartline, J.D.; Malec, D.L.; Sivan, B. Multi-parameter mechanism design and sequential posted pricing. In Proceedings of the Forty-Second ACM Symposium on Theory of Computing, Cambridge, MA, USA, 5 June 2010.

26. Singer, Y.; Mittal, M. Pricing mechanisms in online labor markets. In Proceedings of the Human Computation: AAAI Workshop, San Francisco, CA, USA, 8 August 2011.

27. Zhang, Y.; van der Schaar, M. Reputation-based incentive protocols in crowdsourcing applications. In Proceedings of the IEEE INFOCOM, Orlando, FL, USA, 25–30 March 2012; pp. 2140–2148.

28. Singer, Y.; Mittal, M. Pricing mechanisms for crowdsourcing markets. In Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 13 May 2013.

29. Devanur, N.R.; Hayes, T.P. The adwords problem: Online keyword matching with budgeted bidders under random permutations. In Proceedings of the 10th ACM Conference on Electronic Commerce, Stanford, CA, USA, 6 July 2009.

30. DiPalantino, D.; Vojnovic, M. Crowdsourcing and all-pay auctions. In Proceedings of the 10th ACM Conference on Electronic Commerce, Stanford, CA, USA, 6 July 2009.