A Reputation-Based Contract for Repeated Crowdsensing with Costly Verification

Donya G. Dobakhshari, Parinaz Naghizadeh, Mingyan Liu, and Vijay Gupta

Abstract—We study a setup in which a system operator hires a sensor to exert costly effort to collect accurate measurements of a value of interest over time. At each time, the sensor is asked to report his observation to the operator, and is compensated based on the accuracy of this observation. Since both the effort and observation are private information for the sensor, a naive payment scheme which compensates the sensor based only on his self-reported values will lead to both shirking and falsification of outcomes by the sensor. We consider the problem of designing an appropriate compensation scheme to incentivize the sensor to at once exert costly effort and truthfully reveal the resulting observation.

To this end, we formulate the problem as a repeated game and propose a compensation scheme that employs stochastic verification by the operator coupled with a system of assigning reputation to the sensor. In particular, our proposed payment scheme compensates the sensor based on both the effort in the current period as well as the history of past behavior. We show that by using past behavior in determining present payments, the operator can both incentivize higher effort as well as more frequent truth-telling by the sensor and decrease the required verification frequency.

I. INTRODUCTION
Providing incentives for individuals to follow the system operator’s desired policies in smart networks has received increased attention in recent years (see, e.g., [1]–[4] and the references therein). In a typical setting, the system operator delegates tasks to several autonomous agents. The agents may not benefit directly from the outcome of the task, and hence may not exert sufficient effort to complete the task with the quality desired by the operator. Further, due to reasons such as privacy, it may be costly (or even impossible) for the system operator to access information such as effort exerted by each agent or the resulting outcome directly; instead, she must rely on the data reported by the agent. Consequently, in the absence of a suitable compensation scheme, rational self-interested agents may refrain from exerting the desired effort, and further possibly send falsified information in response to the operator’s inquiry about the task.

As our running example, we consider a crowdsensing application in which autonomous sensors are employed to take measurements about a quantity of interest to the system operator over a predetermined time horizon. At each time step, these measurements are used by the system operator to generate an estimate of the quantity of interest. The sensors incur an effort cost for obtaining each measurement with a specified level of accuracy. A sensor can choose to increase his effort and attain more accurate measurements, but at the expense of a higher effort cost. This cost may model, e.g., the cost of operating the device, or battery power. The sensors are then rewarded based on the accuracy of the information they provide to the operator. Since sensors do not, in general, attach a value to the accuracy of the estimate at the system operator, they do not have any incentive to exert costly effort for generating accurate observations. Furthermore, the system operator has no access to the sensors’ private information (i.e., either the level of effort, or the true accuracy of the measurements they generate). As a result, with a compensation scheme that rewards sensors for self-reported accuracy of the measurements they generate, the sensors may expend little effort, yet misreport their accuracy in order to receive higher compensation. The problem we consider in this paper is to design a contract for the sensors so that they provide measurements with sufficient accuracy to enable the operator to generate an estimate with a desired quality over time. While we concentrate on crowdsensing as the example of interest in this paper, the proposed contract design principles can be applied to other smart networks facing similar challenges.

Specifically, an appropriate contract in our setting has to address two difficulties: (i) profit misalignment, and (ii) information asymmetry between the system operator and the sensor providing the information. To alleviate these challenges, the operator needs to design incentive mechanisms that mitigate both moral hazard (i.e., incentivizing desired actions when effort is costly and the level of effort expended is private information for a participant, e.g., [5, Chapter 4]) and adverse selection (i.e., incentivizing participants to provide truthful information when information is private to them, e.g., [5, Chapter 3]). While an extensive literature in contract theory (see, e.g., [1], [5], [6] and the references therein for an overview of the subject) has focused on resolving either moral hazard or adverse selection separately, we consider the problem of moral hazard followed by adverse selection in a repeated setting. This problem has received much less attention in the literature. Notable exceptions include [7]–[9], and [10], which discuss the problem of moral hazard followed by adverse selection in a static framework. The main difference of our work is to consider this problem in a repeated setting. The repeated setting provides new challenges to the problem in that sensors may adopt time-varying strategies to gain, and then misuse, the trust of the...
system operator. Further, the solutions in [7]–[10] rely on verification of the outcomes generated by the agents (either direct verification or cross-verification). If such verification is costly, care must be taken in the repeated setting to bound the verification cost by allowing only infrequent verification.

We next elaborate on the concept of verification in our setup. In order to mitigate the information asymmetries of moral hazard followed by adverse selection, the operator must somehow verify the sensor’s self-reported outcomes. Such verification is crucial in settings with information asymmetry (see e.g., [11]–[13], in which verification is similarly proposed), as the operator would otherwise have to rely only on the possibly falsified information transmitted by the sensor, i.e., she has no additional measurements or observations of the value of interest, either directly or through other sensors. However, the information asymmetry problem remains non-trivial even with access to verification, since verification may be delayed, costly, or imperfect. In this paper, we assume costly verification; thus, while the operator can verify the sensor at each stage and through appropriate penalties ensure truthful revelation by the sensor, this approach may not, in general, be optimal for the operator when optimizing her total utility, due to the costly nature of verification. In other words, intelligent use of verification is, in itself, not a trivial matter.

Consequently, in our proposed contract, we limit the cost incurred due to verification, by assuming that verification will be invoked only with some probability at each time step. We use this stochastic verification scheme, coupled with a history-dependent payment scheme, to design an optimal contract. In particular, we base the compensation on a reputation score assigned to each sensor. This reputation is based on the history of the past interactions of the sensor with the operator. In other words, the operator rewards the sensor based on the sequence of his actions, rather than merely his behavior at the current stage. The operator assigns higher reputation (and consequently higher payments) to a sensor that is verified and detected to be honest. The operator uses reputation as well as verification to incentivize the sensor to exert sufficient effort, and also to not misreport his outcome, so as to optimize his own utility over multiple time steps.

It is worth mentioning that using reputation for mitigating information asymmetry, particularly in repeated games, is a popular strategy in the literature, see e.g., [14]–[18]. In [14], the authors address the problem of mitigating pure adverse selection by means of reputation indices in a static setting. The authors of [15] present a comprehensive study of the use of reputation for mitigating adverse selection in repeated games, i.e., learning the hidden types of players through repeated interactions. The works in [16]–[19] have focused on repeated interactions between a system operator and an agent under the assumption of only hidden action (moral hazard) for the agent. The work in [16] proposes a method for mitigating moral hazard using imperfect verification, while [17] additionally introduces reputations.

In the crowdsourcing literature, the work in [20]–[23] are also related to the work presented in this paper. The work of [20] considers an (ex ante) payment scheme for agents with verifiable outcome, and mitigates the pure moral hazard problem by introducing reputation-based incentives. For non- verifiable outcomes, a class of peer prediction methods has been studied in [21]–[23]. These works use cross-verification between agents’ inputs to make truthtelling an equilibrium. In contrast to this line of work, we use direct verification, and as well as reputation. Moreover, peer-prediction’s goal is to make truthtelling a Nash equilibrium, which may however not be the best-paying equilibrium from the agents’ viewpoint; our scheme ensures that truthtelling in the contract leads to the best equilibrium from the sensor’s viewpoint.

In contrast to this line of work, we use direct verification, and as well as reputation. Moreover, peer-prediction’s goal is to make truthtelling a Nash equilibrium, which may however not be the best-paying equilibrium from the agents’ viewpoint; our scheme ensures that truthtelling in the contract leads to the best equilibrium from the sensor’s viewpoint.

In addition to the aforementioned distinctions between the current work and existing literature, note that these works assume either pure moral hazard, pure adverse selection, or the problem of mixed moral hazard and adverse selection in a static setting. To the best of our knowledge, our work is the first to study the use of reputation for mitigating both types of information asymmetry simultaneously, in particular, moral hazard followed by adverse selection, in a repeated framework.

The main contribution of our work is to adopt a repeated game approach to the problem of simultaneously incentivizing high effort and truthful reports in a crowdsensing setup, and more generally, in problems which exhibit moral hazard followed by adverse selection. We propose a reputation-based payment scheme coupled with stochastic verification for compensating a sensor who realizes outcomes desired by an operator. We show that under this scheme, the sensor will exert higher effort over time, and will truthfully disclose his accuracy with a higher frequency. Furthermore, the operator needs to resort to verification with a lower frequency. Nevertheless, the operator has to provide higher payments, as a result of which her overall payoff may decrease. We discuss the intuition, as well as some practical implications of these observations.

The remainder of the paper is organized as follows. In Section II, we present the model and some preliminaries. We analyze the proposed reputation-based payment scheme in Section III, and conclude in Section IV with some avenues for future work.

II. MODEL AND PRELIMINARIES

We study the repeated interactions of a principal (here, the system operator who is interested in estimating a quantity of interest), with an agent (here, the sensor that generates the measurements) through a contract.\(^1\)

Remark 1: We would like to point out that the restriction to a single sensor is without loss of generality. This is due to the fact that we assume no budget constraint for the operator and do not consider cross verification among sensors. Therefore, for notational ease, we will concentrate on the case of a single sensor.

For our crowdsensing setup, the operator must offer a contract to the sensor, since she does not have any alternative

\(^1\)We will henceforth use she/her for the operator, and he/his for the sensor.
sources of measurements available. The accuracy of the measurements taken by the sensor increases with the effort expended. Thus, the operator is interested in incentivizing high effort by the sensor, so as to attain sufficiently accurate estimates. However, both the true effort exerted by the sensor, as well as the outcome he obtains in terms of the measurement or its accuracy, are unobservable by the operator. In other words, the operator faces moral hazard (in that she does not know the level of effort expended by the sensor) followed by adverse selection (in that she does not know the outcome of the effort). She therefore relies on the report by the sensor on the accuracy of the estimation.

Formally, at every stage $k$ ($1 \leq k \leq N$), the sensor performs the following actions:

(i) He exerts an effort $x_k \in [0, \bar{x}]$ to perform the task of generating a measurement for which he incurs a cost of $h(x_k)$. The effort $x_k$ leads to an outcome accuracy level $\alpha(x_k)$.

(ii) He informs the operator of the outcome accuracy level. The sensor may misreport the accuracy level to correspond to some other level of effort $\bar{x}_k$.

**Assumption 1:** The accuracy $\alpha(x_k)$ is a deterministic function of $x_k$, and the function is known to both the sensor and the operator. In other words, without loss of generality, we may assume that the sensor reports simply his effort level to the operator.

If the report of the sensor $\bar{x}_k$ is equal to the actual effort $x_k$ expended by him, we say that the sensor has been truthful ($T$) at stage $k$; otherwise, we say that the sensor has falsified the output and is non-truthful ($NT$). We assume at stage $k$ the sensor chooses $T$ with probability $q_k$.

The operator derives a benefit $S(x_k)$ from the task performed by the sensor, which depends on the (true) effort $x_k$ by the sensor. We assume that this function is increasing and concave. We normalize $S(0) = 0$, i.e., the operator does not derive any benefit when $x_k = 0$.

The sensor does not necessarily attach value to the outcome of the task assigned to him. As a result, he should be properly incentivized to exert the effort level desired by the operator. We assume that the operator offers a payment of $P_k$ to the sensor. The problem considered in this paper is the design of this payment by the operator so that a rational self-interested sensor will take actions that lead to maximization of the operator’s utility.

We now specify the utilities of the operator and the sensor. To this end, we need to discuss two factors that we will use in our payment strategy: (i) verification of the sensor’s report by the operator, and (ii) the history of the past reports/efforts of the sensor, which are captured by a reputation score assigned to the sensor. More specifically, at each stage $k$, the operator takes an action $v_k \in \{V, NV\}$, denoting verifying and not verifying, respectively, of the sensor’s reported effort $\bar{x}_k$. We denote by $p_k$ the probability that the operator chooses $V$ at stage $k$. Further, we assume that the verification is perfect, and so the operator accurately detects any falsification by the sensor. The operator incurs a cost of $C > 0$ for this verification. Note that the verification is conducted on the reported effort by the sensor (e.g., by verifying the sensor’s access to equipment, or the number of measurements generated by the sensor), but not on the value of the measurements.

Let $z_k$ denote the level of effort known to the operator at the end of stage $k$. Therefore,

$$z_k = \begin{cases} x_k & \text{if } v_k = V \\ \bar{x}_k & \text{if } v_k = NV. \end{cases}$$

In addition, the operator assigns a reputation $R_k$ to the sensor at each stage $k$. The reputation $R_k$ is updated based on the history of the sensor’s past reputation scores, as well as the assumed effort $z_k$ at stage $k$, i.e.,

$$R_k = f(k, R_{1}, \ldots, R_{k-1}, z_k), \quad k = 1, \ldots, N,$$

where $R_0 = 0$ and the function $f(\cdot)$ is the reputation function selected by the operator. The operator then uses this reputation score to offer a compensation of $P_k := P(R_k)$ to the sensor, where $P(\cdot)$ is an increasing function of $R_k$, and is the payment function that needs to be designed.

Given the above setup and compensation scheme, the utility of the sensor at stage $k$ is given by

$$U^S_k = P(R_k) - bx_k.$$  

For simplicity, in this paper, we assume a linear cost of effort $h(x_k) = bx_k$, with a unit cost $b > 0$, $b$ is assumed to be known to the operator a priori. Therefore, the utility of the sensor at stage $k$ reduces to

$$U^S_k = P(R_k) - bx_k.$$  

The utility of the operator is given by

$$U^P_k = S(x_k) - P(R_k) - C.\mathcal{I}\{v_k = V\},$$

where

$$\mathcal{I}\{v_k = V\} = \begin{cases} 1 & \text{if } v_k = V \\ 0 & \text{otherwise}. \end{cases}$$

Both the sensor and the operator discount future payoffs with a factor $\delta$, so that their payoffs over the entire time horizon is given by

$$U^S = \sum_{k=1}^{N} \delta^k U^S_k, \quad U^P = \sum_{k=1}^{N} \delta^k U^P_k.$$

The operator has to optimize her choice of actions $\{v_1, \ldots, v_N\}$, the reputation scores $\{R_1, \ldots, R_N\}$ and the payments $\{P_1, \ldots, P_N\}$ to maximize her expected utility $\mathbb{E}[U^P]$ with satisfying the following constraints:

(i) Incentive compatibility (IC) constraint: A contract is incentive compatible if the sensor chooses to take the action preferred by the operator. Note that the sensor is a rational decision maker, and therefore chooses his effort level to maximize his expected profit $\mathbb{E}[U^S]$. Formally, if $x^* = \{x^*_1, \ldots, x^*_N\}$ denotes the sequence of efforts desired by the operator, then the compensation scheme should be such that $\mathbb{E}[U^S(x^*)] \geq \mathbb{E}[U^S(x)]$, $\forall x$. 

(ii) Individual rationality (IR) constraint (participation constraint): Both the operator and the sensor should prefer participation in the proposed scheme to opting out. Formally:

\[ \mathbb{E}[U^p] \geq 0, \quad \mathbb{E}[U^s] \geq 0, \]

where \( \mathbb{E}[\cdot] \) denotes expectation. In particular, in \( \mathbb{E}[U^s] \), the expectation is with respect to the verification by the operator, while in \( \mathbb{E}[U^p] \), the expectation is with respect to the truthfulness of the sensor. Note that both verification by the operator and truthfulness of the sensor are stochastic.

Therefore, the optimization problem for the operator is given by

\[
P_1: \begin{cases} 
\max \{ \mathbb{E}[U^p] \} \\
\text{s.t. IC and IR constraints.}
\end{cases}
\]

Remark 2: Note that the use of verification is indispensable: if the sensor is not verified, he will always exert effort \( x_\omega = 0 \), and falsify his effort as \( \hat{x}_\omega = \bar{x} \). The goal of introducing a reputation-based payment scheme is, therefore, to reduce (but not eliminate) the verification frequency.

Remark 3: While the above problem statement assumes general \( N \), in this paper, we present the solution for \( N = 2 \) for notational simplicity, and discuss the insights and intuition for the general solution. The detailed solution for \( N > 2 \), and more interestingly, the case of infinite number of stages, remains a direction for future work.

III. MAIN RESULTS

We now proceed to consider the two-stage (i.e., \( N = 2 \)) game between the sensor and the operator and solve the problem \( P_1 \). Specifically, we will present a linearly weighted reputation-based payment scheme. We analyze the effects of modifying the weight in this reputation function on the effort expended by the sensor, the optimal verification frequency, and the resulting utility of the operator.

A. The weighted reputation function

As discussed in Section II, the operator assigns a reputation \( R_k(\cdot) \) to the sensor at each stage \( k = 1, 2 \), and uses it to assess the payment \( P(R_k) \).

Assumption 2: For simplicity and without loss of generality, we choose the payment function as \( P(R_k) = R_k \), where \( R_k \) is defined in (2).

We next propose a weighted reputation function, in order to assess we define the history-dependent payments for the sensor. Formally,

\[
\begin{cases} 
\text{Payment at the first stage} = R(z_1), \\
\text{Payment at the second stage} = (1 - \omega)R(z_1) + \omega R(z_2),
\end{cases}
\]

where \( R(\cdot) \) is an increasing and convex function, \( z_1 \) and \( z_2 \) are defined based on (1), and \( 0 \leq \omega \leq 1 \) is the reputation weight.

C. Nash equilibrium

We start with the pure strategy Nash equilibrium (NE) of the two-stage game with stage payoffs given in Table I.

Proposition 1: The only pure strategy Nash equilibrium of the game in Table I is \( (NT, NV) \). In this equilibrium, the operator offers no payment to the sensor, and the sensor exerts no effort. This is equivalent to the outside option for the operator.

Proof: We start with the potential pure Nash equilibrium \( (T, V) \), and analyze the payoffs of the operator. By playing \( V \) over \( NV \) at the first stage, the operator decreases her utility by \( \gamma \) at the subsequent stage. Therefore, \( (T, NV) \)

\[ \text{2A choice of } l \neq 0 \text{ may be of interest to ensure individual rationality when verification is not perfect.} \]
dominates \((T,V)\). Similarly, by analyzing the payoff of the sensor, we can see that \((NT,NV)\) dominates \((T,NV)\). We have therefore discarded \((T,V)\) and \((T,NV)\) as pure Nash equilibria of the stage games. Therefore, if a pure strategy Nash equilibrium exists, the sensor will be playing \(NT\). However, given that the sensor is always playing \(NT\), the operator’s optimal choice is to set \(h = 0\), and play \(NV\). Therefore, the only possible pure strategy Nash equilibrium of the game is \((NT,NV)\), in which the operator offers no payment to the sensor. Note also that this is in fact the operator’s outside option. 

Table I

<table>
<thead>
<tr>
<th>(T)</th>
<th>(V)</th>
<th>(NT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p_k(x_k) + \gamma - bx_k, S(x_k) - R_k(x_k) - \gamma - C)</td>
<td>(p_k(0), -R_k(0) - C)</td>
<td>(p_k(\bar{x}), -R_k(\bar{x}))</td>
</tr>
</tbody>
</table>

TABLE I

PAYOFFS OF THE SENSOR (ROW PLAYER) AND THE OPERATOR (COLUMN PLAYER) AT EACH STAGE \(k\)

Proof: See [24].

1) Role of inter-temporal incentives: Intuitively, the reputation weight \(\omega\) determines the importance of inter-temporal incentives (i.e., conditioning future payments on the history of past efforts). In particular, \(\omega = 1\) yields an instant payment scheme, in which no inter-temporal incentives are present. For this case, the actions of the operator and the sensor are as follows.

**Corollary 1:** If \(\omega = 1\), the sensor realizes \((x^*, x^*)\), the operator verifies the sensor with probability \(p_2 = p_1 = 1\), and the sensor is only truthful with probability \(q_1 = q_2 = 1 - \frac{C}{b\bar{x}} < 1\).

Proof: Note that if \(\omega = 1\), we have \(h = \frac{b\bar{x}}{\omega} = b\bar{x}\) at each stage. This leads to the sensor choosing \(x^*\) at both first and second stages, i.e., \(x_1 = x_2 = x^*\) when \(\omega = 1\). Further, substituting \(\omega = 1\) in (iii) of Theorem 1 leads to \(p_2 = p_1 = 1\), and \(q_1 = q_2 = 1 - \frac{C}{b\bar{x}} < 1\).

In is interesting to note that, when \(\omega = 1\), although the operator always verifies the sensor, he still chooses \(NT\) with a strictly positive probability. This occurs due to the fact that for \(\omega = 1\), the utility of the sensor under \(T\) is equal to his utility under \(NT\). Hence, the sensor is indifferent between the choice of \(T\) and \(NT\).

By comparing Theorem 1 and Corollary 1, we observe that the verification frequency, falsification probabilities, and the effort level of the sensor, at the second stage, are equal with the use of reputation and the case when no reputation is used. However, these values at the first stage are affected by the introduction of inter-temporal incentives. In particular, when reputation-based payments that depend on the history of the behavior of the sensor are used, the operator needs to verify the sensor with a lower probability, and the sensor is truthful with a higher probability. Furthermore, the sensor exerts higher effort in the first stage.

2) Optimal choice of the reputation weight \(\omega\): Finally, we consider the optimal choice of \(\omega\), under which the operator’s expected payoff is maximized.

**Theorem 2:** Assume that \(b\bar{x} > \sqrt{CS(\bar{x})}\). A choice of \(\omega = 1\) maximizes the operator’s payoff. That is, instant payments yield higher payoffs than payments based on linearly weighted reputations.

Proof: For \(b\bar{x} > \sqrt{CS(\bar{x})}\), the optimal choice of \(h\) is identified in Theorem 1. Hence, we now analyze the optimal choice of \(\omega\) when \(h = \frac{b\bar{x}}{\omega}\) and \((x_1, x_2) = (\bar{x}, x^*)\). We need to
solve the following optimization problem:

$$\max_{\omega} \quad -(1 + \delta) C + \left(1 - \omega \frac{C}{b \bar{x}}\right)(S(\bar{x}) - b \bar{x} \omega)$$

$$+ \delta \left(1 - \frac{C}{b \bar{x}}\right)(S(x) - b x) , \quad \text{s.t.} \quad 0 \leq \omega \leq 1.$$

We take the derivative of the objective function with respect to \(\omega\). Define \(f'(\omega) := \frac{\partial f}{\partial \omega}\). Then,

$$\frac{\partial E[U^f]}{\partial \omega} = -b \bar{x} f'(\omega)(1 - \frac{1}{f^2(\omega)} \frac{S(x) C}{(b \bar{x})^2}).$$

With the assumption \(\sqrt{CS(\bar{x})} < b \bar{x}\), and noting that \(f(\omega) \geq 1\) and \(f'(\omega) < 0\), we conclude that \(\frac{\partial E[U^f]}{\partial \omega} > 0\). Thus, \(E[U^f]\) is an increasing function of \(\omega\) and the optimal choice is to set \(\omega = 1\).

We observe that while using inter-temporal incentives through linearly weighted reputation functions can benefit the operator by reducing the required verification frequency, increasing the effort level of the sensor, and increasing the probability of truthfulness, it will nevertheless reduce the operator’s overall payoff. This is because the operator has to now offer a higher compensation to the sensor.

### IV. Conclusion

In this paper, we studied the problem of contract design between a system operator and a strategic sensor in a repeated setting. The sensor is hired to exert costly effort to collect sufficiently accurate observations for the operator. As the effort invested and the accuracy of the resulting outcome are both private information of the sensor, the operator needs to design a compensation scheme that mitigates moral hazard followed by adverse selection. We proposed a reputation-based payment scheme coupled with stochastic verification. We showed that by increasing the importance of past behavior in our proposed linearly weighted reputation-based payments, the sensor exerts higher effort, and has a higher probability of being truthful. The operator, on the other hand, can invoke verification less frequently, but offers higher payments to the sensor, which leads to a lower payoff.

We have so far considered inter-temporal incentives that are based on a linearly weighted reputation function. Considering other functional forms for evaluating a sensor’s reputation, and its impact on the operator and sensor’s strategies and payoffs, is an important direction of future work.

In addition, we have considered the design of individual contracts for each sensor, due to our assumptions of independent measurements and no budget constraint. As an interesting direction of future work, we are interested in analyzing the contract design problem for multiple sensors, given limited budget of the operator, as well as when the outcomes of the estimate of the sensors are coupled. Such coupling may enable the operator to cross-verify the outcomes of the sensors.

### References


