

# Identifying Unreliable Sensors Without a Knowledge of the Ground Truth in Deceptive Environments

Anis Yazidi<sup>1</sup>(✉), B. John Oommen<sup>2</sup>, and Morten Goodwin<sup>3</sup>

<sup>1</sup> Department of Computer Science,  
Oslo and Akershus University College of Applied Sciences, Oslo, Norway

<sup>2</sup> School of Computer Science, Carleton University, Ottawa, Canada  
anis.yazidi@hioa.no

<sup>3</sup> Department of Computer Science, University of Agder, Oslo, Norway

**Abstract.** This paper deals with the extremely fascinating area of “fusing” the outputs of sensors *without any knowledge of the ground truth*. In an earlier paper, the present authors had recently pioneered a solution, by mapping it onto the fascinating paradox of trying to identify *stochastic liars* without any additional information about the truth. Even though that work was significant, it was constrained by the model in which we are living in a world where “the truth prevails over lying”. Couched in the terminology of Learning Automata (LA), this corresponds to the Environment (Since the Environment is treated as an *entity* in its own right, we choose to capitalize it, rather than refer to it as an “environment”, i.e., as an abstract concept.) being “Stochastically Informative”. However, as explained in the paper, solving the problem under the condition that the Environment is “Stochastically Deceptive”, as opposed to informative, is far from trivial. In this paper, we provide a solution to the problem where the Environment is deceptive (We are not aware of any other solution to this problem (within this setting), and so we believe that our solution is both pioneering and novel.), i.e., when we are living in a world where “lying prevails over the truth”.

**Keywords:** Sensor fusion · Unreliable sensors · Learning automata · Learning from Stochastic Liars

## 1 Introduction

We consider the problem of fusing the information obtained from a set of sensors where the knowledge of the “Ground Truth” is unavailable. However, unlike the problem that has been traditionally considered (i.e., whether the “Ground Truth” is available or not), we consider the intrinsically more complex model in which the sensors could be stochastically “truth telling” or “deceptive”, and

---

B.J. Oommen—*Chancellor’s Professor; Fellow: IEEE and Fellow: IAPR*. This author is also an *Adjunct Professor* with the University of Agder in Grimstad, Norway.

where the behavior of the sensor is not known *a priori*. This problem is, in and of itself, non-trivial, and as in the case when one deals with Stochastic Teachers and Stochastic Liars, there is no universal guaranteed solution to such puzzles. To place the field of sensor fusion in the right perspective, we mention that the aggregation of data obtained from sensors enables us to procure more reliable information about the underlying process, as opposed to utilizing the raw sensor information from the individual sensors themselves. However, the quality of the aggregated information is intricately dependent on the *reliability* of the individual sensors. In fact, understandably, unreliable sensors will tend to report erroneous values of the ground truth, and thus degrade the quality of the fused information. Finding strategies to identify unreliable sensors can assist in having a counter-effect on their respective detrimental influences on the fusion process, and this has been a focal concern in the literature. The body of the related work operate with the assumption of direct knowledge of the ground truth to assess the reliability of the sensors, or indirect knowledge using the concept of sensor accuracy that is deduced from historical data. The existing literature generally assumes that the reliability of the individual sensors can be inferred, whence one can invoke an efficient scheme to fuse their respective readings. Although the task of resolving this problem without the knowledge of the ground truth is apparently impossible, the authors of [4, 5] (which are also the present authors) previously obtained conclusive results by utilizing the “agreement” between the sensors themselves and a set of Linear Reward-Inaction ( $LRI$ ) Learning Automata (LA) associated with the sensors. The results of [4, 5] were constrained by the model in which “the truth prevails over lying”, which, in the setting of LA corresponds to the Environment being “Stochastically Informative”. Informally speaking, this is equivalent to the scenario where the proportion of truth-tellers in the society exceeds the proportion of liars. This paper considers the scenario in which the Environment is “Stochastically Deceptive”, where “lying prevails over the truth”, or if you like, where the proportion of liars exceeds the proportion of truth-tellers. This is not an unrealistic setting. Indeed, in cases of nuclear meltdowns, the majority of the sensors in the vicinity of the meltdown can be considered faulty and unreliable<sup>1</sup>.

## 1.1 Survey of the Field

A myriad of pieces of literature can be cited that concentrate on using majority voting to faulty sensor fusion. The premise for invoking majority voting is that the decision of the group is better than the decision of the individual sensor.

The theory of sensor fusion has also found wide deployment in the field of “reputation systems” where users who want to promote a particular product or service can flood the domain (i.e., the social network) with sympathetic votes, while those who want to get a competitive edge over a specific product

---

<sup>1</sup> This being said, the content and goal of this paper is to present a solution within a theoretical and conceptual framework. Thus, we will not embark on the study of any real-life application domains here.

or service can “badmouth” it unfairly. Thus, although these systems can offer generic recommendations by aggregating user-provided opinions, unfair ratings may degrade the trustworthiness of such systems. This problem, of separating “fair” and “unfair” agents for a specific service, is called the Agent-Type Partitioning Problem (*ATPP*). Determining ways to solve the (*ATPP*) [3] and thus counter the detrimental influence of unreliable agents on a Reputation System, has been a focal concern of a number of very interesting studies.

The analogous sensor-related problem, of separating reliable and unreliable sensors, is called the Sensor-Type Partitioning Problem (*STPP*). We shall solve it in stochastically Deceptive Environments. Put in a nutshell, in this paper, we propose to solve the above-mentioned paradoxical *STPP* using tools provided by LA, which have proven powerful potential in efficiently and quickly learning the optimal action when operating in unknown stochastic Environments. It adaptively, and in an on-line manner, gradually learns the identity and characteristics of the sensors that are reliable and those that are unreliable. In addition, we will provide two approaches for fusing the sensor readings which leverage the convergence result of our LA-based partitioning.

A recent work by the authors of the current paper, that was alluded to earlier, is found in [4,5]. This paper pioneered a solution by which it is feasible to solve the *STPP* problem of identifying which sensors are unreliable *without any knowledge of the ground truth*, a claim that *is counter-intuitive*. The essence of the approach presented in [4,5] stems from the simple intuition that the “agreement” between the sensors themselves can give invaluable knowledge about their respective reliabilities. In a stochastic Environment where errors can take place according to some unknown underlying stochastic process, those sensors that tend to deviate from the decision of the majority are more likely to be unreliable than those that adhere to the decision of the majority. Such simple and intuitive remark works under the premise that the decision of the majority has some high likelihood of revealing the truth [4,5]. The main assumption of our legacy work was the fact that “the truth prevails over lying” which is translated into a condition that can be seen as an extension of the simple majority voting. In fact, the reader can observe in [4,5] that if the Environment is deterministic, i.e., the reliable sensors are deterministic (always report the ground truth with probability 1) and the unreliable sensors are deterministic too (always misreport the ground truth with probability 1), then the mild condition of the “the truth prevails over lying” translates to the simple and well-know majority vote in the setting that the number of reliable sensors forms the majority. Stochastically, the setting in which “truth prevails over lying” is tantamount to having more stochastically reliable sensors than stochastically unreliable ones.

In this paper, we consider a natural but non-obvious scenario where “lying prevails over the truth”. Alluding to the terminology of LA (and more particularly the theory of the SPL [2]), such an Environment can be characterized

as being “Deceptive” as opposed to “Informative”<sup>2</sup>. As a dual to the previous framework, stochastically, the setting in which “lying prevails over truth” is tantamount to having more stochastically unreliable sensors that stochastically reliable ones.

To justify the validity of the claims that we have made, rigorous theoretical results and a host of empirical results were presented in [4, 5]. These have been extended and generalized in this paper for Deceptive Environments.

## 2 Modeling the Problem

We consider a population of  $N$  sensors,  $\mathcal{S} = \{s_1, s_2, \dots, s_N\}$ . Let the real situation of the Environment at the time instant  $t$  be modeled by a binary variable  $T(t)$ , which can take one of two possible values, 0 and 1. The value of  $T$  is unknown and can only be inferred through measurements from sensors. The output from the sensor  $s_i$  is referred to as  $x_i$ . Let  $\pi$  be the probability of the state of the ground truth, i.e.,  $T = 0$  with probability  $\pi$ .

To formalize the scenario, we record four possibilities:

- $x_i = T$  (where  $x_i = 0$  or  $1$ ): This is the case when the sensor correctly reports the ground truth.
- $x_i \neq T$  (where  $x_i = 0$  or  $1$ ): This is the case when the sensor faultily reports the ground truth.

In our discussions, we make one simplifying assumption: The probability of the sensor reporting a value erroneously is symmetric. In other words, in terms of the binary detection problem, we assume that the probability of a false alarm and the so-called miss probability are both equal. Thus, presented formally, we assume that:

$$\text{Prob}(x_i = 0|T = 1) = \text{Prob}(x_i = 1|T = 0). \quad (1)$$

Further, let  $q_i$  denote the Fault Probability (FP) of sensor  $s_i$ , where:

$$q_i = \text{Prob}(x_i = 0|T = 1) = \text{Prob}(x_i = 1|T = 0).$$

Similarly, we define the Correctness Probability (CP) of sensor  $s_i$  as  $p_i = 1 - q_i$ .

It is easy to prove that the total probability  $\text{Prob}(x_i = T)$  is, indeed,  $p_i$ , since, in fact:

$$\begin{aligned} \text{Prob}(x_i = T) &= \text{Prob}(T = 0)\text{Prob}(x_i = 0|T = 0) + \text{Prob}(T = 1)\text{Prob}(x_i = 1|T = 1) \\ &= \pi p_i + (1 - \pi)p_i \\ &= p_i. \end{aligned} \quad (2)$$

---

<sup>2</sup> In the case of a recommendation system, a Deceptive Environment can, for example, correspond to a compromised system where the integrity of the majority of the agents in the systems are compromised.

Thus, the quantity  $p_i = \text{Prob}(x_i = T)$  can be re-written as  $p_i = \text{Prob}(I\{x_i = T\} = 1)$ , where  $I\{\cdot\}$  is the Indicator function.

We refer to a sensor as being reliable when it has a FP  $q_i < 0.5$ . Conversely, the sensor is unreliable when it has a FP  $q_i > 0.5$ . Equivalently, a reliable sensor is one that has a CP  $p_i > 0.5$ , and an unreliable sensor as one that has a CP  $p_i < 0.5$ .

Observe that as a result of this model, a reliable sensor will probabilistically tend to report 0 when the ground truth is 0, and 1 when the ground truth is 1. Otherwise, it is clearly, unreliable. Our aim, then, is to partition the sensors as being reliable or unreliable. Furthermore, once partitioned, our aim is to use the partitioning as a basis for better fusion.

To simplify the analysis<sup>3</sup>, we assume that every  $p_i$  can assume one of two possible values from the set  $\{p_R, p_U\}$ , where  $p_R > 0.5$  and  $p_U < 0.5$ . Then, a sensor  $s_i$  is said to be reliable if  $p_i = p_R$ , and is said to be unreliable if  $p_i = p_U$ . To render the problem non-trivial and interesting, we assume that  $p_R$  and  $p_U$  are unknown to the algorithm.

Based on the above, the set of reliable sensors is  $\mathcal{S}_R = \{s_i | p_i = p_R\}$ , and the set of unreliable sensors is  $\mathcal{S}_U = \{s_i | p_i = p_U\}$ .

We now formalize the Sensor-Type Partitioning Problem (*STPP*). The *STPP* involves a set of  $N$  sensors<sup>4</sup>,  $\mathcal{S} = \{s_1, s_2, \dots, s_N\}$ , where each sensor  $s_i$  is characterized by a fixed but unknown probability  $p_i$  of it sensing the ground truth correctly. The *STPP* involves partitioning  $\mathcal{S}$  into 2 mutually exclusive and exhaustive groups so as to obtain a 2-partition  $\mathbb{G} = \{G_U, G_R\}$ , such that each group,  $G_R$ , of size,  $N_R$ , and  $G_U$ , of size  $N_U$ , exclusively contains only the sensors of its own type, i.e., which are either reliable or unreliable respectively.

We define  $P_{(N_R-1, N_U)}$  as the probability of a deterministic majority voting scheme, which involves the opinions of  $N_R - 1$  reliable sensors and  $N_U$  unreliable ones, to yield the correct decision using the majority rule. In other words, this is the probability that a majority of more than  $\frac{(N_R-1+N_U)}{2}$  of the sensors will advocate the ground truth. Similarly, we define  $P_{(N_R, N_U-1)}$  as the probability of a deterministic majority voting scheme, which involves the opinions of  $N_R$  reliable sensors and  $N_U - 1$  unreliable ones, to yield the correct decision using the majority rule. As one can see, this quantity is the same: It too is the probability that a majority of more than  $\frac{(N_R+N_U-1)}{2}$  of the sensors will, in turn, advocate the ground truth.

In [4, 5], we assumed that:

$$(N_R - 1)p_R + N_U p_U > \frac{(N_R + N_U)}{2}.$$

The latter condition is founded on a fundamental premise that has to hold in any sustainable society, where telling the “truth” is considered a virtue, while “lying”

<sup>3</sup> This assumption, however, does not simplify the problem. Indeed,  $p_R$  can be assigned to be the smallest value of all the values of  $p_i$  for the reliable sensors, and  $p_U$  can be assigned to be the largest value of all the values of  $p_i$  for the unreliable ones.

<sup>4</sup> Throughout this paper, since we will be invoking majority-like decisions, we assume that  $N = N_R + N_U$  is an even number.

is considered detrimental and harmful to the society. In this paper, the task we undertake is to consider the non-intuitive complementary problem. Indeed, we will investigate the non-trivial case in which the phenomenon of “lying” is more prevalent than that of saying the “truth” exasperated by the case when the proportion of stochastic “lying” agents exceeds the number of stochastic “truth-telling” agents. We shall endeavor to state and prove the relevant theoretical results for the case where:

$$(N_R - 1)p_R + N_U p_U < \frac{(N_R + N_U)}{2} - 1.$$

The reader should observe that whenever the Environment is deterministic, i.e.,  $p_R = 1$  and  $p_U = 0$ , the above condition can be written as  $N_R < N_R + N_U$ , which simply means that the set of unreliable sensors form the Majority from among the sensors. Hence, in the special case of a deterministic Environment, the problem can be seen as resolving the problem using Majority voting.

### 3 The Solution

#### 3.1 Overview of Our Solution

In this paper, we provide a novel solution to the *STTP* for the scenario where “lying prevails over the truth”, and where this solution is based on the field of LA that was briefly surveyed above. It is appropriate to mention that we are not aware of any other solution to this problem (within this setting), and so we believe that our solution is both pioneering and novel. We intend to take advantage of the fact that LA combine rapid and accurate convergence with low computational complexity. In addition to its computational simplicity, unlike most reported approaches, as mentioned earlier, our scheme does not require prior knowledge of the ground truth. Rather, it adaptively, and in an on-line manner, gradually learns the identity and characteristics of the sensors which tend to provide reliable readings, and of those which tend to provide unreliable ones.

Our solution involves a team of *LA* where each *LA* is uniquely attached to (or rather, associated with) a specific sensor, on a one-to-one basis. Each automaton  $\mathcal{A}^i$ , attached to sensor  $s_i$ , has two actions.

By suitably modeling the agreement or disagreement of the opinions about the sensed ground truth between each sensor and the rest of the other sensors, we can appropriately model these as responses from the corresponding “Environment”. Using these synthesized responses, our scheme will intelligently group the sensors according to the readings that they report about the ground truth. Since a sensor is reliable if it reports the ground truth correctly with a probability  $p_i > 0.5$  (and unreliable otherwise), we will design our scheme so that it can infer the similar sensors and collect them into their respective groups. In other words, we will infer the crucial pieces of information, namely the identities of the sensors, from the random stream of sensor reports.

The fusion part of our scheme will be based on the result of a prior partitioning phase. Ultimately, the aim behind identifying the set of unreliable sensors,  $\mathcal{S}_U$ , is to improve the performance of the fusion process for inferring the ground truth. The result of the convergence of the team of LA, which results in a partitioning that infers the identity of the sensors, will serve as an input to the fusion process. In this vein, we present a simple approach for fusing the results, and study its performance in the section that describes the experimental result. The fusion approach only considers the measurements from the reliable sensors as being informative, and simultaneously discards measurements from the unreliable sensors. An alternate fusion scheme which considers the responses from all the sensors is also described. The first formal result concerning the performance of the LA is given below.

### 3.2 Theoretical Results for the Case Where: “Lying Prevails over Truth”

In this section, we provide theoretical results pertinent to the extremely interesting and fascinating case when “Lying Prevails over Truth-Telling”, i.e., when it is more likely for the sensors to be unreliable than reliable, or if you like, the number of unreliable sensors is more than the number of reliable ones.

We analyze and provide the theoretical results for the case where  $(N_R - 1)p_R + N_U p_U < \frac{(N_R + N_U)}{2} - 1$ . The proofs of the following two theorems are quite involved and are omitted here for the sake of brevity. The proofs can be found in a unabridged version of this article [4].

**Theorem 1.** *Consider the scenario when  $N_R p_R + (N_U - 1)p_U < \frac{(N_R + N_U)}{2} - 1$  and when  $N_R + N_U - 1 \geq 3$ . Let  $s_i \in \mathcal{S}_R$ . Consider now the agreement between the opinion of a reliable sensor  $s_i$  and the opinion of the majority formed by all the rest of the sensors  $S \setminus \{s_i\} = (\mathcal{S}_R \setminus \{s_i\}) \cup \mathcal{S}_U$ . Let  $y_{(N_R-1, N_U)}$  be the decision of a majority voting scheme  $S \setminus \{s_i\}$ , based on the responses of  $N_R - 1$  reliable and  $N_U$  unreliable sensors. Then, if  $x_i$  is the output of  $s_i$ :  $\text{Prob}(x_i = y_{(N_R-1, N_U)}) < 0.5$ .*

The next theorem, which deals with the analogous case of excluding an unreliable sensor, follows.

**Theorem 2.** *Consider the scenario when  $N_R p_R + (N_U - 1)p_U < \frac{(N_R + N_U)}{2} - 1$  and when  $N_R + N_U - 1 \geq 3$ . Let  $s_i \in \mathcal{S}_U$ . Consider now the agreement between the opinion of an unreliable sensor  $s_i$  and the opinion of the majority formed by all the rest of the sensors,  $S \setminus \{s_i\} = \mathcal{S}_R \cup \mathcal{S}_U \setminus \{s_i\}$ . Let  $y_{(N_R, N_U-1)}$  be the decision of a majority voting scheme based on the responses of  $S \setminus \{s_i\}$ , consisting of  $N_R$  reliable and  $N_U - 1$  unreliable sensors. Then, if  $x_i$  is the output of  $s_i$ :  $\text{Prob}(x_i = y_{(N_R, N_U-1)}) > 0.5$ .*

### 3.3 Construction of the Learning Automata

The results that we have presented in the previous section form the basis of our LA-based solution. We explain this below, including the strategy by which the majority vote is invoked.

In the partitioning strategy, with each sensor  $s_i$  we associate a 2-action  $L_{RI}$  automaton  $\mathcal{A}^i$ ,  $(\Sigma^i, \Pi^i, \Gamma^i, \Upsilon^i, \Omega^i)$ , where  $\Sigma^i$  is the set of actions,  $\Pi^i$  is the set of action probabilities,  $\Gamma^i$  is the set of feedback inputs from the Environment, and  $\Upsilon^i$  is the set of action probability updating rules. Each of these is explained below.

1. *The set of actions of the automaton:  $(\Sigma^i)$*

The two actions of the automaton are  $\alpha_k^i$ , for  $k \in \{0, 1\}$ , i.e.,  $\alpha_0^i$  and  $\alpha_1^i$ .

2. *The action probabilities:  $(\Pi^i)$*

$P_k^i(n)$  represent the probabilities of selecting the action  $\alpha_k^i$ , for  $k \in \{0, 1\}$ , at step  $n$ . Initially,  $P_k^i(0) = 0.5$ , for  $k = 0, 1$ .

3. *The feedback inputs from the Environment to each automaton:  $(\Gamma^i)$*

Let the automaton select either the action  $\alpha_0^i$  or  $\alpha_1^i$ . Then, the responses from the Environment and the corresponding probabilities are tabulated below. For a chosen action, the Environment will respond by a “Reward”, or a “Penalty”. The conditional probabilities of the “Reward”, and “Penalty” are also specified in the tables.

A brief explanation about the equations in these tables could be beneficial.

(a) The LA system is rewarded if it chooses action  $\alpha_0^i$ , in which case the reading of the sensor  $s_i$  agrees with the opinion of the majority voting scheme associated with  $S \setminus \{s_i\}$ . This occurs with probability  $Prob(x_i = y_{(N_R-1, N_U)})$  whenever  $s_i \in \mathcal{S}_R$  and with probability  $Prob(x_i = y_{(N_R, N_U-1)})$  whenever  $s_i \in \mathcal{S}_U$ .

(b) Alternatively, the system is rewarded if it chooses action  $\alpha_1^i$ , in which case the reading of the sensor  $s_i$  disagrees with the opinion of the majority voting scheme associated with  $S \setminus \{s_i\}$ . This occurs with probability  $1 - Prob(x_i = y_{(N_R-1, N_U)})$  whenever  $s_i \in \mathcal{S}_R$  and with probability  $1 - Prob(x_i = y_{(N_R, N_U-1)})$  whenever  $s_i \in \mathcal{S}_U$ .

(c) The penalty scenarios are the reversed ones.

4. *The action probability updating rules:  $(\Upsilon^i)$*

First of all, since we are using the  $L_{RI}$  scheme, we ignore all the penalty responses. Upon reward, we obey the following updating rule:

If  $\alpha_k^i$  for  $k \in \{0, 1\}$  was rewarded then,

$$\begin{aligned} P_{1-k}^i(n+1) &\leftarrow \theta \times P_{1-k}^i(n) \\ P_k^i(n+1) &\leftarrow 1 - \theta \times P_{1-k}^i(n), \end{aligned}$$

where  $0 \ll \theta < 1$  is the  $L_{RI}$  reward parameter.

Before we prove the properties of the overall system, we first state a fundamental result of the  $L_{RI}$  learning schemes which we will repeatedly allude to in the rest of the paper.

**Lemma 1.** *An  $L_{RI}$  learning scheme with parameter  $0 \ll \theta < 1$  is  $\epsilon$ -optimal, whenever an optimal action exists. In other words,  $\lim_{\theta \rightarrow 1} \lim_{n \rightarrow \infty} P_k^i(n) \rightarrow 1$ .*



The above result is well known [1]. By virtue of this property, we are guaranteed that for any  $L_{RI}$  scheme with the two actions  $\{\alpha_0, \alpha_1\}$ , if  $\exists k \in \{0, 1\}$  such that  $c_k^i < c_{1-k}^i$ , then the action  $\alpha_k^i$  is optimal, and for this action  $P_k^i(n) \rightarrow 1$  as  $n \rightarrow \infty$  and  $\theta \rightarrow 1$ , where the  $\{c_k^i\}$ , are the penalty probabilities for the two actions of the automaton  $\mathcal{A}^i$ .

By invoking the property of the  $L_{RI}$  learning scheme, we state and prove the convergence property of the overall system.

**Theorem 3.** *Consider the scenario when  $(N_R - 1)p_R + N_U p_U < \frac{(N_R + N_U)}{2} - 1$  and that  $N_R + N_U - 1 \geq 3$ . Given the  $L_{RI}$  scheme with a parameter  $\theta$  which is arbitrarily close to unity, the following is true:*

$$\begin{aligned} & \text{If } s_i \in \mathcal{S}_R, \text{ then } \lim_{\theta \rightarrow 1} \lim_{n \rightarrow \infty} P_0^i(n) \rightarrow 1; \\ & \text{If } s_i \in \mathcal{S}_U, \text{ then } \lim_{\theta \rightarrow 1} \lim_{n \rightarrow \infty} P_1^i(n) \rightarrow 1. \end{aligned}$$

**Proof:** To prove the theorem, we treat the two cases separately.

*Case 1:  $s_i \in \mathcal{S}_R$ :* Based on the result of Theorem 1, we can see that the inequality  $Prob(x_i = y_{(N_R-1, N_U)}) < 0.5$  holds. We can thus deduce that:

$$Prob(x_i = y_{(N_R-1, N_U)}) < 1 - Prob(x_i = y_{(N_R-1, N_U)}). \tag{3}$$

If we now consider the entries of Table 1 that specify the penalty probabilities  $s_i \in \mathcal{S}_R$ , we see that:

$$c_1^i = Prob(x_i = y_{(N_R-1, N_U)}) < c_0^i = 1 - Prob(x_i = y_{(N_R-1, N_U)}),$$

implying that for this case, the action  $\alpha_1^i$  is the optimal one. Consequently, by virtue of Lemma 1, for this action:

$$P_1^i(n) \rightarrow 1 \text{ as } n \rightarrow \infty \text{ and } \theta \rightarrow 1,$$

proving the result for this case.

*Case 2:  $s_i \in \mathcal{S}_U$ :* In this case, based on the result of Theorem 2, we see that the following inequality holds:  $Prob(x_i = y_{(N_R, N_U-1)}) > 0.5$ .

Therefore we can confirm that

$$Prob(x_i = y_{(N_R, N_U-1)}) > 1 - Prob(x_i = y_{(N_R, N_U-1)}). \tag{4}$$

**Table 1.** Reward and Penalty probabilities for sensor  $s_i \in \mathcal{S}_R$

Action	Associated probability	
	Reward	Penalty
$\alpha_0^i$	$Prob(x_i = y_{(N_R-1, N_U)})$	$1 - Prob(x_i = y_{(N_R-1, N_U)})$
$\alpha_1^i$	$1 - Prob(x_i = y_{(N_R-1, N_U)})$	$Prob(x_i = y_{(N_R-1, N_U)})$

**Table 2.** Reward and Penalty probabilities for sensor  $s_i \in \mathcal{S}_U$

Action	Associated probability	
	Reward	Penalty
$\alpha_0^i$	$Prob(x_i = y_{(N_R, N_U-1)})$	$1 - Prob(x_i = y_{(N_R, N_U-1)})$
$\alpha_1^i$	$1 - Prob(x_i = y_{(N_R, N_U-1)})$	$Prob(x_i = y_{(N_R, N_U-1)})$

From the entries of Table 2, that specify the penalty probabilities  $s_i \in \mathcal{S}_U$ , we obtain:

$$c_0^i = 1 - Prob(x_i = y_{(N_R, N_U-1)}) > c_1^i = Prob(x_i = y_{(N_R, N_U-1)}).$$

This implies that the action  $\alpha_0^i$  is the optimal one, and for this action:

$$P_0^i(n) \rightarrow 1 \text{ as } n \rightarrow \infty \text{ and } \theta \rightarrow 1.$$

The theorem is thus proven. □

### 3.3.1 Remarks and Some Additional Notation

Based on what we have already seen, the following observations are in place:

1. Analogous to the above theorems, from Theorem 3, we see the similar result for the case when:

$$N_R p_R + (N_U - 1)p_U < \frac{(N_R + N_U)}{2} - 1.$$

In fact, when  $N_R p_R + (N_U - 1)p_U < \frac{(N_R + N_U)}{2} - 1$ , the reliable sensors will converge to action  $\alpha_0^i$ , while the unreliable ones to action  $\alpha_1^i$  with an arbitrarily large probability. To summarize these results, let:

- $G_R = \{s_i \in \mathcal{S} \text{ such that } \lim_{n \rightarrow \infty} P_1^i(n) = 1\}$
- $G_U = \{s_i \in \mathcal{S} \text{ such that } \lim_{n \rightarrow \infty} P_0^i(n) = 1\}$ .

As the conclusions are  $\epsilon$ -optimal results, if  $\theta$  is arbitrarily close to unity,  $G_R$  will converge to  $\mathcal{S}_R$  and  $G_U$  will converge to  $\mathcal{S}_U$ . On the other hand, if  $\theta$  is not arbitrarily close to unity, some of the LA might fail to converge to the optimal action, and thus the set  $G_R$  may not necessarily be equivalent to  $\mathcal{S}_R$ , and  $G_U$  may not necessarily be equivalent to  $\mathcal{S}_U$ .

2. In our earlier work [4], we had dealt with a with a society where “truth prevails over lying” (i.e., where, effectively, the number of reliable sensors was more than the number of unreliable ones), characterized by the canonical equation:

$$(N_R - 1)p_R + N_U p_U > \frac{(N_R + N_U)}{2}.$$

A naive way to *attempt* to obtain the condition for the opposite scenario involving a deceptive environment, i.e., one in which “lying prevails over

truth”, would be to invert the equations by exchanging  $N_R$  with  $N_U$  and  $p_R$  with  $p_U$  respectively. The inverted equation obtained by such a straightforward substitution is:

$$(N_U - 1)p_U + N_R p_R > \frac{(N_R + N_U)}{2}.$$

However, on performing a rigorous analysis, one observes that the above condition “does not lead anywhere”. Further, the condition does not guarantee any form of convergence<sup>5</sup>. Rather, since reasoning by direct symmetry does not work, deducing the correct condition that is applicable for Deceptive Environments is far from being intuitive.

3. A more careful investigation reveals that the correct condition,  $N_R p_R + (N_U - 1)p_U < \frac{(N_R + N_U)}{2} - 1$ , is not symmetric. Indeed, it is *this* condition that is valid for the case where “lying prevails over truth”. One will observe that the above condition reduces to  $N_R < \frac{(N_R + N_U)}{2}$  for  $(p_R, p_U) = (1, 0)$ . In other words, whenever the Environment is deterministic, implying that a reliable sensor will always tells the truth ( $p_R = 1$ ) and an unreliable sensor will always misreport the truth ( $p_U = 0$ ), we will obtain a minority of reliable sensors since  $N_R < \frac{(N_R + N_U)}{2}$ , forcing the unreliable sensors to constitute the majority.

### 3.4 Fusion Schemes with Exclusion: Discarding the Opinions of the Unreliable Sensors

A possible strategy to increase the accuracy of the fusion process is to employ a simple majority voting strategy that excludes all the sensors whose LA converged to the action  $G_U$  during the partitioning phase. This means that the prediction of the ground truth will be exclusively based on the “accurate” sensors, i.e., those whose LA converged to the action  $G_R$ .

## 4 Experimental Results

The performance of the LA-based partitioning as well as the fusion schemes with exclusion (that makes use of the partitioning described in Sect. 3.4), have been rigorously tested by simulation in a variety of parameter settings, and the results that we have obtained are truly conclusive. In the interest of brevity, we merely report a few representative (and typical) experimental results, so that the power of our proposed methodology can be justified. In the experiments, the settings were chosen so that the condition  $(N_R - 1)p_R + N_U p_U < \frac{(N_R + N_U)}{2} - 1$  was met, reflecting that the world possessed that the phenomenon in which “lying prevails over the truth”.

---

<sup>5</sup> The absence of convergence was also supported by experimental results that are not reported here. This was, indeed, what motivated the present avenue of research.

### 4.1 Fusion Scheme with Exclusion

We now compare the ‘‘Fusion Scheme with Exclusion’’ with the deterministic Majority Voting (MV) strategy that incorporates all the sensors in  $S$ . As detailed earlier, the latter scheme relies exclusively on the decision of the vote of the majority of the sensors that converged to the  $G_R$  partition. Let  $P(C_c)$  denote the probability of the consensus being correct, i.e., that the probability that the vote of the majority coincides with the ground truth. Table 3 reports the result of the comparison for the case when  $N_R$  and  $N_U$  are both equal to 10.

**Table 3.** Comparisons of the value of  $P(C_C)$ , the probabilities of the consensus being correct for different values of  $(p_R, p_U)$ , and for the different approaches for  $N_R = 10$  and  $N_U = 10$ .

$(p_R, p_U)$	$P(C_C)$ for Fusion Scheme with Exclusion	$P(C_C)$ for MV for all sensors
(0.55, 0.25)	0.738	0.234
(0.6, 0.25)	0.833	0.13
(0.65, 0.25)	0.905	0.401
(0.7, 0.25)	0.952	0.5
(0.55, 0.2)	0.738	0.16
(0.6, 0.2)	0.833	0.225
(0.65, 0.2)	0.905	0.426
(0.7, 0.2)	0.952	0.396

From this table, we observe:

1. The distribution of  $T$  does not play a role in determining the value of  $P(C_c)$  for the Fusion Scheme with Exclusion because of the symmetry property of the fault. As one can see, the results we report are conclusive. In fact, we were able to increase the value of  $P(C_c)$  quite remarkably. For example, for the case when  $(p_R, p_U) = (0.7, 0.25)$ , our scheme yielded a value of 0.952 for  $P(C_C)$ , while the scheme which operated with the MV involving all the sensors yielded the value of only 0.5.
2. The value of  $P(C_C)$  for the simple MV involving all sensors gave a low accuracy (less than 0.5) as the Environment was Deceptive.
3. The value of  $P(C_C)$  for our Fusion Scheme with Exclusion was immune to the variation of  $p_U$ . For example, for the entries corresponding to  $p_R = 0.7$ , we see that  $P(C_C)$  was equal to 0.952 even if  $p_U$  changed, for example, by taking the values 0.25 or 0.2.

Consider now the case when the value  $N_U$  was doubled from 10 to 20 while the value of  $N_R$  was equal to 10. As expected, we see from Table 4, the value of  $P(C_C)$  for our scheme was intact and independent of the value of  $N_U$ .

**Table 4.** Comparisons of  $P(C_C)$ , the probabilities of the consensus being correct for different values of  $(p_R, p_U)$ , and for the different approaches for  $N_R = 20$  and  $N_U = 10$ .

$(p_R, p_U)$	$P(C_C)$ for Fusion Scheme with Exclusion	$P(C_C)$ for MV for all sensors
(0.55, 0.25)	0.738	0.057
(0.6, 0.25)	0.833	0.081
(0.65, 0.25)	0.905	0.112
(0.7, 0.25)	0.952	0.15
(0.55, 0.2)	0.738	0.02
(0.6, 0.2)	0.833	0.03
(0.65, 0.2)	0.905	0.046
(0.7, 0.2)	0.952	0.066

## 5 Conclusion

The authors of the current articles have recently pioneered a solution to an extremely pertinent problem, namely, that of identifying which sensors are unreliable *without any knowledge of the ground truth*. This fascinating paradox can be formulated in simple terms as trying to identify *stochastic* liars without any additional information about the truth. In this paper, we provide a LA-based solution to the problem where the sensors operated in a world in which “lying prevails over truth telling”, or informally speaking, where the number of unreliable sensors is stochastically more than the number of reliable ones.

## References

1. Narendra, K.S., Thathachar, M.A.L.: Learning Automata: An Introduction. Prentice-Hall, New Jersey (1989)
2. Oommen, B.J.: Stochastic searching on the line and its applications to parameter learning in nonlinear optimization. IEEE Trans. Syst. Man Cybernet. B **27**, 733–739 (1997)
3. Yazidi, A., Granmo, O.C., Oommen, B.J.: Service selection in stochastic environments: a learning-automaton based solution. Appl. Intell. **36**(3), 617–637 (2012)
4. Yazidi, A., Oommen, B.J., Goodwin, M.: On solving the problem of identifying unreliable sensors without a knowledge of the ground truth: the case of stochastic environments. IEEE Trans. Cybernet. **47**, 1604–1617 (2017)
5. Yazidi, A., Oommen, B.J., Goodwin, M.: On distinguishing between reliable and unreliable sensors without a knowledge of the ground truth. In: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), vol. 2, pp. 104–111. IEEE (2015)