Optimal Opponent Selection for Distributed Multi-Agent Self-Classification

Hang Zou, Youba Nait-Belaid, Michel Kieffer

Optimal Opponent Selection
for Distributed Multi-Agent Self-Classification

Hang Zou, Youba Nait-Belaid, and Michel Kieffer

L2S, CNRS –CentraleSupelec –Univ Paris-Sud –Univ Paris-Saclay
3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France.
michel.kieffer@l2s.centralesupelec.fr

Abstract—This paper considers a mobile multi-agent system (*e.g.*, a crowdsensing or crowdsourcing application) in which agents are characterized by their discrete-valued Level of Ability (LoA) at doing some sensing or data processing task. Agents are not aware of their LoA and are willing to estimate it without the help of a central classification authority, in order to determine whether their contribution will improve or degrade the performance of the global network. Using their estimated LoA, agents may then voluntarily restrain themselves to participate to the activity of the multi-agent system, without being banned by some central control authority.

For that purpose, agents, when they meet, perform pairwise comparison tests (PCT) able to determine which is the best agent of a pair of competing agents. Two maximum *a posteriori* (MAP) estimators of the LoA of each agent are proposed using the results of several PCTs. These MAP estimators are then employed to determine, for a given agent, the LoA of the next opponent that minimizes its probability of LoA estimation error. Simulations results show that the proposed optimal opponent selection approach provides better results than simply choosing opponents at random, or always choosing the opponents with the best LoA.

Index Terms—Distributed algorithms, Estimation, Multi-agent systems

I. INTRODUCTION

In crowdsensing applications [3], [25], a large group of agents equipped with sensing and computing capabilities (smartphones, tablets, cars...) collaborate to collect data and extract information to measure, map, estimate [7], [20], classify [18], [2], or analyze any process of common interest [12], [25].

An important problem in such applications is to evaluate the Level of Ability (LoA), *i.e.*, the performance, of the participating agents at doing sensing or data processing tasks, in order to keep only information coming from agents with the best LoA. For example, in a sensing application, agents may only have a very partial knowledge of the performance of their sensing device (biased or unbiased, very noisy or accurate) [6]. In crowdsourcing applications, agents may not be totally aware of their LoA at doing some task.

Most of the prior work assume that some byzantine agents are voluntarily injecting erroneous data in data collection and processing networks [13], [10], or fake news in social networks to disturb the network [23]. Here, we assume that agents which may degrade the performance of the network are not aware of this behavior. Such situation may occur due to degraded sensors, or to a virus perturbing the sensors [21]. Nevertheless,

agents are willing to detect whether the information they produce and transmit may degrade the performance of the network, before being banned by some central authority.

The aim of this paper is to help each agent in estimating in a distributed way its *own* LoA at doing the basic tasks required by some crowdsensing or crowdsourcing application. Agents with LoA not high enough may then, by themselves, restrain participating in the application. Most of the approaches addressing this problem consider centralized reputation-based mechanisms [24], [14], [19] to reject poorly performing agents based on previous results. Distributed approaches have also been proposed, see, *e.g.*, [7], [15]. The main advantage of the proposed approach is to help agents protecting their reputation in a decentralized way, since the detection is performed with the help of peers, and not by the network management authority.

For that purpose, following the idea of [16], one assumes that two agents, when they meet, are able to perform a pairwise comparison test (PCT). PCTs provide some information about the agent with the best LoA in the pair of competing agents. Such PCTs are found, *e.g.*, in weakly supervised learning and preference learning [8], [5], [1], [11]. In [16], each agent, using the proportion of PCTs for which it was deemed as the best agent, estimates its own LoA.

In this paper, one assumes that each agent can decide whether it will perform a PCT when meeting an other agent. This decision may depend on the estimate of their respective LoAs exchanged at the beginning of the meeting. Two MAP estimators of the LoA of an agent are then proposed using the results of a succession of PCTs. Then, this paper introduces a technique to determine, for a given agent, the LoA of the opponent with which the PCT will be the most informative, in the sense that the PCT outcome will minimize its probability of LoA estimation error.

The LoA estimation problem is presented in Section II. Section III describes two MAP estimators of the LoA from PCT outcomes. Section IV introduces a criterion to determine the LoA of the most informative opponent against which an agent has to perform its next PCT to minimize its LoA estimation error. Some simulation results are described in Section V before drawing conclusions in Section VI.

II. PROBLEM FORMULATION

The notations introduced in [16] are briefly recalled here before describing the specific problem addressed in this work.

Consider a multi-agent system consisting of N_{a} mobile agents. Each agent i has a given LoA θ_i at doing some task. One assumes that the θ_is are realizations of independently and identically distributed (iid) discrete-valued random variables Θ_i, $i = 1, \ldots, N_{\mathrm{a}}$. The set of possible LoAs is $\mathcal{L} = \{1, 2, ..., K\}$, with $\theta_i = 1$ representing the best LoA and $\theta_i = K$ representing the worst one. The vector $\mathbf{p} = (p_1, \ldots, p_K)^T$ gathers the *a priori* probabilities associated to each LoA.

One assumes that the agents are not aware of their LoA, and are willing to estimate it. For that purpose, two agents i and j, when they come in vicinity, can perform a PCT. The outcome $y_{i,j}$ of such PCT is modeled as a realization of a binary random variable $Y_{i,j}$ which distribution is conditioned by the LoAs of the competing agents, *i.e.*,

$$\Pr\left(Y_{i,j} = 1 \mid \theta_i, \theta_j\right) = q\left(\theta_i, \theta_j\right). \tag{1}$$

Two types of PCTs have been considered in [16]. One focuses here on the *win-lose* PCT, also considered in preference-based learning [8], such that

$$q\left(\theta_i, \theta_j\right) = 1 - q\left(\theta_j, \theta_i\right). \tag{2}$$

Such PCT typically represents the outcome of a game between two players. The various $q\left(\theta_i, \theta_j\right)$ are gathered in a *preference matrix* [22], [4], [17]

$$Q = \begin{pmatrix} q\left(1, 1\right) & \ldots & q\left(1, K\right) \\ \vdots & \ddots & \vdots \\ q\left(K, 1\right) & \ldots & q\left(K, K\right) \end{pmatrix}. \tag{3}$$

We assume that the time interval between two successive meetings follows an exponential distribution with an inter-contact rate λ [26], [9]. The population of agents is assumed well-mixed, *i.e.*, for a given agent, the probability of meeting an agent with a given LoA θ is proportional to p_θ.

Our aim in this paper, for each agent i, is to minimize the LoA estimation error considering a given number of PCT outcomes. For that purpose, during each meeting with an other agent, Agent i will have to decide whether it is worth performing a PCT. More precisely, considering the LoA estimate of the other agent, it has to determine whether a sufficient decrease of the estimation error of its LoA θ_i is likely to be obtained from the PCT outcome.

III. MAP ESTIMATION OF AGENTS' LoA

Two cases are considered, depending on whether the LoAs of the opponents are known.

A. MAP estimator with unknown opponents' LoAs

Consider Agent i, which LoA θ_i corresponds to a realization of the random variable Θ_i. Agent i has performed n PCTs against different opponents without knowing their LoAs. The PCT outcomes are considered as realizations of n independent binary random variables Y_ℓ, $\ell = 1, \ldots, n$ with conditional probability given by (1). The MAP estimate of θ_i using the vector of PCT outcomes $\mathbf{y}^{(n)} = (y_1, \ldots, y_n)^T$ is

$$\widehat{\theta}_{i,\mathrm{MAP}} = \arg\max_\theta \Pr\left(\Theta_i = \theta \mid \mathbf{Y}^{(n)} = \mathbf{y}^{(n)}\right). \tag{4}$$

In what follows, the random variables will be omitted when there is no ambiguity.

Proposition 1: The MAP estimate of the LoA θ_i of Agent i which has performed n independent PCTs with outcomes $y_1, \ldots, y_n$ against different opponents, without knowing their LoA, is

$$\widehat{\theta}_{i,\mathrm{MAP}} = \arg\max_\theta s_\theta^m \left(1 - s_\theta\right)^{n-m} p_\theta, \tag{5}$$

with

$$s_\theta = \sum_{k \in \Theta} q\left(\theta, k\right) p_k \tag{6}$$

and $m = \sum_{\ell=1}^n y_\ell$.
The proof of Proposition 1 is in Apprendix A.

Proposition 1 leads to a partition of the set $\{1, \ldots, n\}$ into K decision subsets $\mathcal{N}_1^{(n)}, \ldots, \mathcal{N}_K^{(n)}$ defined as

$$\begin{aligned} \mathcal{N}_k^{(n)} = \{ m \in \{1, \ldots, n\} \ \mathrm{st} \ & s_k^m \left(1 - s_k\right)^{n-m} p_k \\ > & s_j^m \left(1 - s_j\right)^{n-m} p_j, \forall j \neq k \} \end{aligned} \tag{7}$$

and such that $\sum_{\ell=1}^n y_\ell \in \mathcal{N}_k^{(n)} \Leftrightarrow \widehat{\theta}_{i,\mathrm{MAP}} = k$.

B. MAP estimator with partly known opponents' LoAs

Here, one assumes that Agent i performs $n = n_{\mathrm{u}} + n_{\mathrm{k}}$ PCTs against n different agents. The LoA of n_{u} of them is *unknown*, the LoA of n_{k} of them is *known* by Agent i. Without loss of generality, one assumes that the LoAs of the last opponents are known and are denoted as θ'_ℓ, $\ell = n_{\mathrm{u}} + 1, \ldots, n$. The PCT outcomes are gathered in $\mathbf{y}^{(n)} = (y_1, \ldots, y_n)^T$.

With these assumptions, the MAP estimate of θ_i is

$$\widehat{\theta}_{i,\mathrm{MAP}} = \arg\max_\theta \Pr\left(\Theta_i = \theta \mid \mathbf{y}^{(n)}, \theta'_{n_{\mathrm{r}}+1}, \ldots, \theta'_n\right). \tag{8}$$

Proposition 2: The MAP estimate of θ_i from $n = n_{\mathrm{u}} + n_{\mathrm{k}}$ PCT outcomes against n_{u} opponents of unknown LoAs and n_{k} opponents of known LoA $\theta'_{n_{\mathrm{u}}+1}, \ldots, \theta'_n$ is

$$\widehat{\theta}_{i,\mathrm{MAP}} = \arg\max_\theta s_\theta^{m_{\mathrm{u}}} \left(1 - s_\theta\right)^{n_{\mathrm{u}}-m_{\mathrm{u}}}$$
$$\prod_{\ell=n_{\mathrm{u}}+1}^n \left[y_\ell q\left(\theta, \theta'_\ell\right) + \left(1 - y_\ell\right)\left(1 - q\left(\theta, \theta'_\ell\right)\right) \right] p_\theta, \tag{9}$$

where $m_{\mathrm{u}} = \sum_{\ell=1}^{n_{\mathrm{u}}} y_\ell$.
The proof of Proposition 2 is in Appendix A.

In this case, $m = \sum_{\ell=1}^n y_\ell$ is no more a sufficient statistic for the MAP estimator. Nevertheless, Proposition 5 still leads to partitions of the set $\{0, \ldots, n_{\mathrm{u}}\}$ into K subsets. The

partitions depend on $(y_{n_u+1}, \ldots, y_n)$ and $(\theta'_{n_u+1}, \ldots, \theta'_n)$ and are expressed as

$$\mathcal{N}_k^{(n_u, n_k)}\left(y_{n_u+1}, \ldots, y_n, \theta'_{n_u+1}, \ldots, \theta'_n\right) = \{m \in \{0, \ldots, n_u\} \text{ st}$$

$$s_k^m (1-s_k)^{n_u-m} \left(\prod_{\ell=n_u+1}^{n} f\left(y_\ell, k, \theta'_\ell\right) \quad p_k > \right.$$

$$s_j^m (1-s_j)^{n_u-m} \left(\prod_{\ell=n_u+1}^{n} f\left(y_\ell, j, \theta'_\ell\right) \quad p_j, \forall j \neq k \right) .$$

with

$$f\left(y_\ell, k, \theta'_\ell\right) = y_\ell q\left(k, \theta'_\ell\right) + (1-y_\ell)\left(1 - q\left(k, \theta'_\ell\right)\right).$$

Again, one has

$$\sum_{\ell=1}^{n_u} y_\ell \in \mathcal{N}_k^{(n_u, n_k)}\left(y_{n_u+1}, \ldots, y_n, \theta'_{n_u+1}, \ldots, \theta'_n\right)$$

$$\Leftrightarrow \widehat{\theta}_{i,\text{MAP}} = k. \quad (10)$$

IV. OPTIMAL OPPONENT SELECTION FOR PAIA

Assume that Agent i has performed n PCTs. It aims at determining the LoA its $n+1$-th opponent should have so as to minimize its probability of mis-classification once it has used the outcome of the $n+1$-th PCT.

Assume that the n first PCTs are with outcomes $\mathbf{y}^{(n)} = (y_1, \ldots, y_n)$, and are done without knowing the opponents' LoA. In the case $\Theta_i = \theta_i$, the probability of LoA estimation error is

$$p_{e,i}^{(n)} = \Pr\left(\sum_{\ell=1}^{n} Y_\ell \notin \mathcal{N}_{\theta_i}^{(n)} | \theta_i, \mathbf{y}^{(n)}\right)$$

$$= \begin{cases} 0 & \text{if } \sum_{\ell=1}^{n} y_\ell \in \mathcal{N}_{\theta_i}^{(n)} \\ 1 & \text{else,} \end{cases} \quad (11)$$

since, once $\sum_{\ell=1}^{n} y_\ell$ is known, the outcome of the MAP estimator is deterministic.

Assume now that the $n+1$-th opponent is chosen such that its LoA is $\Theta'_{n+1} = \theta'_{n+1}$. The expected probability of error $p_{e,i}^{(n+)}$ before performing the $n+1$-th PCT, when $\mathbf{y}^{(n)}$ as well as the LoA of the $n+1$ opponent are known but not y_{n+1} is

$$p_{e,\theta_i}^{(n+)}\left(\theta'_{n+1}\right)$$

$$= \Pr\left(\sum_{\ell=1}^{n} Y_\ell \notin \mathcal{N}_{\theta_i}^{(n,1)}\left(Y_{n+1}, \theta'_{n+1}\right) | \theta_i, \theta'_{n+1}, \mathbf{y}^{(n)}\right)$$

$$= \sum_{y_{n+1} \in \{0,1\}} \Pr\left(\sum_{\ell=1}^{n} Y_\ell \notin \mathcal{N}_{\theta_i}^{(n,1)}\left(y_{n+1}, \theta'_{n+1}\right), y_{n+1} | \theta_i, \theta'_{n+1}, \mathbf{y}^{(n)}\right)$$

$$= \sum_{y_{n+1} \in \{0,1\}} \Pr\left(\sum_{\ell=1}^{n} Y_\ell \notin \mathcal{N}_{\theta_i}^{(n,1)}\left(y_{n+1}, \theta'_{n+1}\right) | \theta_i, \theta'_{n+1}, \mathbf{y}^{(n)}, y_{n+1}\right)$$

$$\Pr\left(y_{n+1} | \theta_i, \theta'_{n+1}, \mathbf{y}^{(n)}\right)$$

Since $\Pr\left(y_{n+1} | \theta_i, \theta'_{n+1}, \mathbf{y}^{(n)}\right)$ does not depend on $\mathbf{y}^{(n)}$, once θ_i and θ'_{n+1} are known, one has

$$\Pr\left(y_{n+1} | \theta_i, \theta'_{n+1}\right) = f\left(y_{n+1}, \theta_i, \theta'_{n+1}\right).$$

The first term in the sum is evaluated as (11). One observes that $p_{e,\theta_i}^{(n+)}\left(\theta'_{n+1}\right)$ for a given value of θ_i and θ'_{n+1} only depends on $\sum_{\ell=1}^{n} y_\ell = m$.

To get the average (over all possible values of Θ_i) expected probability of error before performing the $n+1$-th PCT, once the $n+1$-th opponent has been chosen, one has to evaluate

$$p_e^{(n+)}\left(\theta'_{n+1}\right) =$$

$$\Pr\left(\sum_{\ell=1}^{n} Y_\ell \notin \mathcal{N}_{\theta_i}^{(n,1)}\left(Y_{n+1}, \theta'_{n+1}\right) | \theta'_{n+1}, \mathbf{y}^{(n)}\right)$$

$$= \sum_{\theta_i=1}^{K} \Pr\left(\sum_{\ell=1}^{n} Y_\ell \notin \mathcal{N}_{\theta_i}^{(n,1)}\left(Y_{n+1}, \theta'_{n+1}\right), \theta_i | \theta'_{n+1}, \mathbf{y}^{(n)}\right)$$

$$= \sum_{\theta_i=1}^{K} \Pr\left(\sum_{\ell=1}^{n} Y_\ell \notin \mathcal{N}_{\theta_i}^{(n,1)}\left(Y_{n+1}, \theta'_{n+1}\right) | \theta_i, \theta'_{n+1}, \mathbf{y}^{(n)}\right)$$

$$\Pr\left(\theta_i | \theta'_{n+1}, \mathbf{y}^{(n)}\right). \quad (12)$$

One has to determine $\Pr\left(\theta_i | \theta'_{n+1}, \mathbf{y}^{(n)}\right)$. Knowing the LoA of the $n+1$-th opponent, without knowing the outcome of the $n+1$-th PCT does not bring more information on θ_i than the outcomes of the last n PCTs. Thus $\Pr\left(\theta_i | \theta'_{n+1}, \mathbf{y}^{(n)}\right) = \Pr\left(\theta_i | \mathbf{y}^{(n)}\right)$ may be evaluated using Proposition 1.

As a consequence, to minimize the probability of LoA estimation error after $n+1$ PCTs, an agent has to select the $n+1$-th opponent which LoA θ'_{n+1} minimizes

$$p_e^{(n+)}\left(\theta'_{n+1}\right) = \sum_{\theta_i=1}^{K} p_{e,\theta_i}^{(n+)}\left(\theta'_{n+1}\right) \Pr\left(\theta_i | \mathbf{y}^{(n)}\right). \quad (13)$$

In (13), the impact of $\mathbf{y}^{(n)}$ is only through the sum of its components $m = \sum_{\ell=1}^{n} y_\ell$. The decision on θ'_{n+1} will thus only depend on m.

These derivations are easily extended to the selection of the $n+1$-th opponent that minimizes the probability of LoA estimation error assuming that $n = n_u + n_k$ PCTs have been performed against n_u agents with unknown LoA and against n_k agents selected depending on their LoA.

V. SIMULATION RESULTS

One considers $K = 4$ different LoAs and the following preference matrix

$$Q = \begin{bmatrix} 0.5 & 0.855 & 0.958 & 0.988 \\ 0.145 & 0.5 & 0.855 & 0.958 \\ 0.042 & 0.145 & 0.5 & 0.855 \\ 0.012 & 0.042 & 0.145 & 0.5 \end{bmatrix}$$

The *a priori* information on the proportions of agents is $\mathbf{p} = (0.5, 0.25, 0.125, 0.125)^{\mathsf{T}}$.

A. Evolution of the decision regions $\mathcal{N}_k^{(n)}$

Using (7), one is able to plot the various decision regions $\mathcal{N}_k^{(n)}$ as a function of the number n of PCTs performed against agents with unknown LoA. Figure 1 (top) shows that as soon as n is large enough, the decision regions do not overlap.
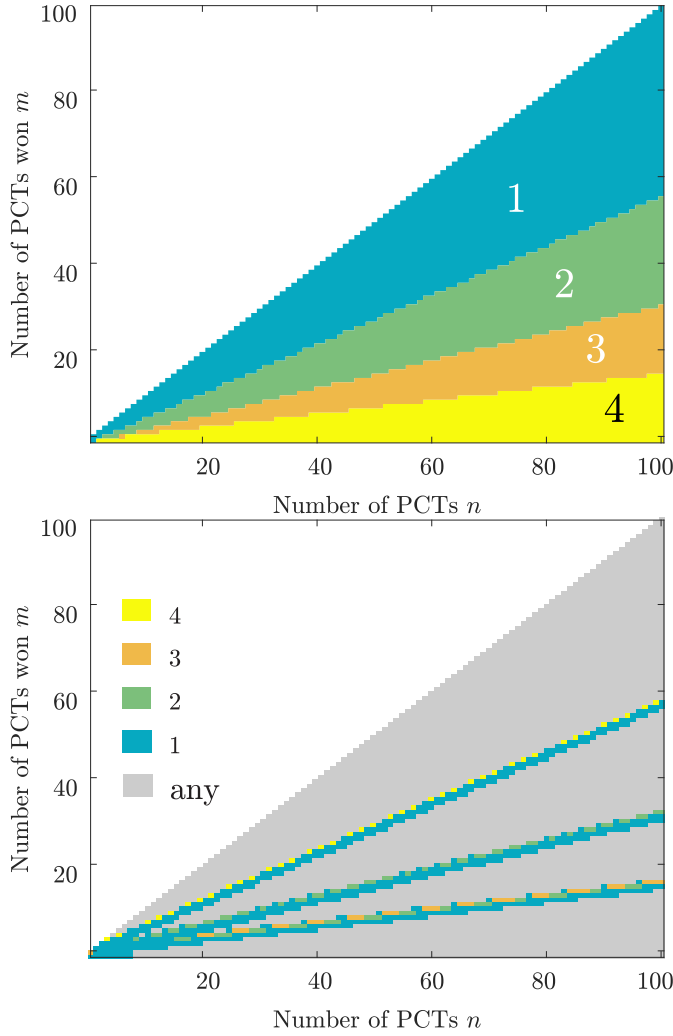
Fig. 1. Decision regions $\mathcal{N}_k^{(n)}$ (top) and optimal LoA of the next opponent (bottom) as a function of the number of PCTs and of the number of PCTs won

B. Optimal LoA of next opponent

Using (13), one is able to evaluate the error probability $p_{\mathrm{e}}^{(n+)}\left(\theta_{n+1}'\right)$ as a function of m, see Table I in the case $n = 10$, *i.e.*, when an agent has to select its most informative opponent after having performed $n = 10$ PCTs as a function of m, the number of PCTs won.

One observes that for several values of m, the next opponent may be chosen randomly, without any effect on the LoA estimation error. In what follows those values of m are indicated to belong to Class R (random agent). Such values of m belong to the middle of the decision regions $\mathcal{N}_k^{(n)}$, where the next PCT will not affect the LoA estimate, see also Figure 1 (bottom).

In several cases, the most informative opponent should be with $\theta_{n+1}' = 1$. The corresponding values of m are indicated to belong to Class B (best agent). Finally, here, when $m = 4$, the most informative opponent is such that $\theta_{n+1}' = 2$. Values of m for which the most informative opponent is not such

that $\theta_{n+1}' = 1$ and cannot be chosen randomly are indicated to belong to Class O (other agents). Values of m belonging to Class B and O are close to the boundaries of the decision regions.

Table II shows the proportions of values of m belonging to each class. One observes that when n gets large, the next opponent may be chosen at random without significantly affecting the performance of the LoA estimation algorithm. One may alternatively always select an agent such that $\theta_{n+1}' = 1$. Except for small values of n, this choice minimizes the LoA estimation error in 95% of the cases when $n = 40$ and in more than 97% of the cases when $n = 80$. This confirms the intuition found in [16] that performing PCT always against opponents with LoA $\theta' = 1$ is a reasonable choice.

To evaluate the benefits of selecting the most informative opponents in terms of probability of LoA estimation error, one has considered $N_{\mathrm{a}} = 1000$ agents randomly moving in a unit square, with LoA randomly chosen according to the *a priori* probability vector **p**. Figure 2 represents the evolution of the proportion of agents having erroneously estimated their LoA as a function of the number of PCTs performed. Results are averaged over 100 simulations. Three cases are considered: in a first case, agents always perform a PCT when they come in vincinity to an other agent. In the second case, after n_{u} PCTs against random agents, PCTs are only performed when the encountered agent estimates its LoA is $\theta' = 1$. The last case is similar to the second case, but now, the opponent with estimated LoA minimizing the probability of LoA estimation error is considered.

The estimation performance is much better in the last two cases. Selecting the agent with an estimated LoA $\theta' = 1$ provides slightly better results than selecting an opponent with LoA minimizing the probability of LoA estimation error. This may be due to the increased robustness of the second approach when the LoA is still not perfectly estimated.

VI. CONCLUSION

This paper presents two MAP estimation algorithms to determine the LoA of agents at doing some task in a crowsensing application . These estimators take as input the outcomes of PCTs performed with other agents. When the LoA of opponents is not known, the proportion of PCTs won against randomly chosen opponents is a sufficient statistic to estimate the LoA of the considered agent. When the LoA of some opponents is known, the expression of the MAP estimator is less straightforward.

This paper introduces a criterion to choose the LoA of the next opponent of a given agent to minimize its LoA estimation error. Using this criterion, one observes that in many cases, performing PCTs with opponents with the best LoA provides very good results, as already experienced in [16]. Performing PCTs with opponents at random leads to an increased probability of error. One observes that the selection of an opponent with a specific LoA is important when the current LoA estimate remains uncertain.

TABLE I
LoA ESTIMATION ERROR PROBABILITY AND OPTIMAL OPPONENT WHEN $n = 10$; $*$ INDICATES THAT ALL LoAs ARE OPTIMAL

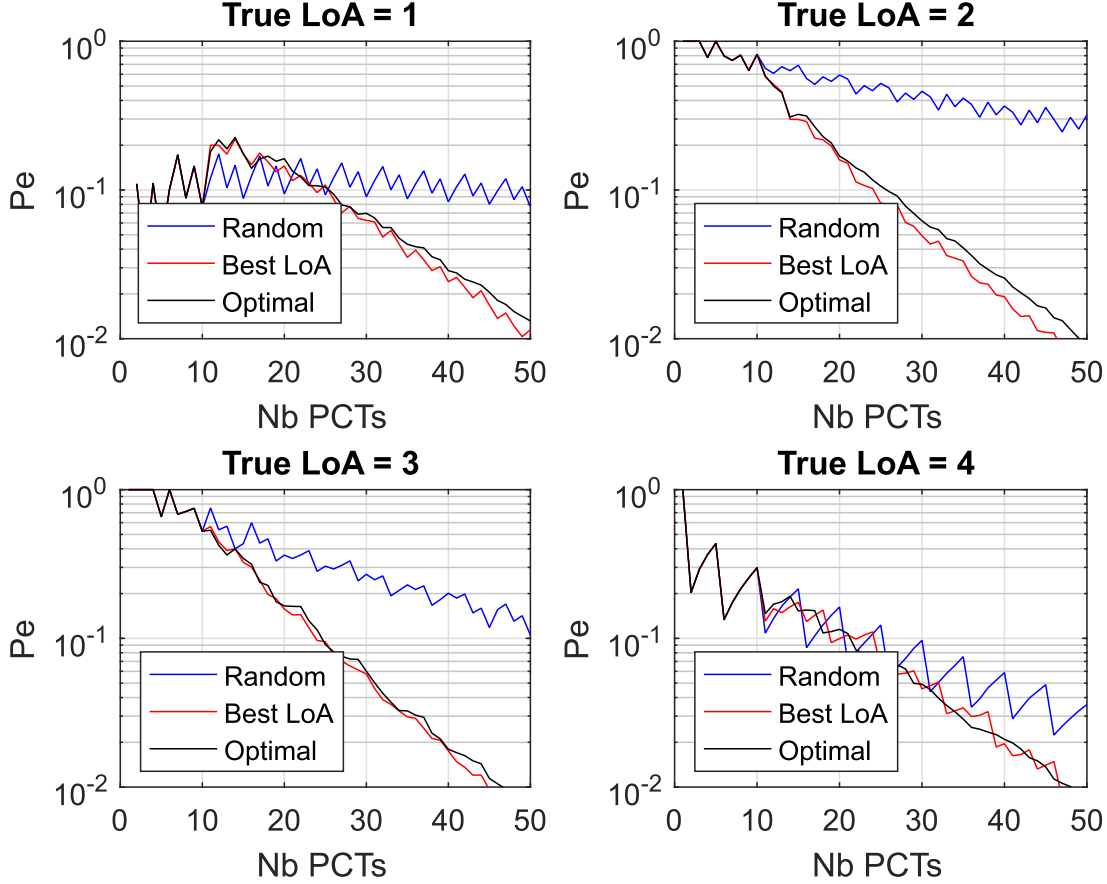| opponent's LoA	Number m of PCTs won										
	0	1	2	3	4	5	6	7	8	9	10
	Percentage of LoA estimation error										
$\theta'_{n+1} = 1$	1.1	3.59	**2.7**	**1.7**	3.2	5.4	**5.0**	**1.2**	0.4	0.06	0.004
$\theta'_{n+1} = 2$	1.1	3.6	3.1	3.5	**3.1**	5.4	8.5	1.2	0.4	0.06	0.004
$\theta'_{n+1} = 3$	1.1	3.6	4.4	6.5	7.9	5.4	9.5	1.3	0.4	0.06	0.004
$\theta'_{n+1} = 4$	1.1	3.6	5.7	8.0	8.0	5.4	9.9	1.3	0.4	0.06	0.004
optimal LoA	$*$	$*$	1	1	2	$*$	1	1	$*$	$*$	$*$

Fig. 2. Proportion of erroneous LoA estimation as a function of the number of PCTs, when PCTs are against (*i*) random opponents, (*ii*) opponents estimating their LoA as the best (except for the first 10 PCTs), and (*iii*) opponents with estimated LoA minimizing the probability of LoA estimation error.

TABLE II
PROPORTIONS OF VALUES OF m BELONGING TO EACH CLASS AS A FUNCTION OF n

	number of PCTs			
	$n = 20$	$n = 40$	$n = 80$	$n = 100$
Class R (%)	61.90	80.49	90.12	93.70
Class B (%)	28.57	14.63	7.41	4.95
Class O (%)	9.52	4.88	2.47	1.98

Significant work remains to be done to address situations where only the preference matrix is provided but no knowledge about the initial proportions of agents, or vice-versa. One has also to evaluate the impact of the selection of the LoA of the two or more next opponents on the LoA estimation error.

APPENDIX

One has

$$\Pr\left(\Theta_i = \theta \mid \mathbf{y}^{(n)}\right) \propto \Pr\left(\mathbf{y}^{(n)} \mid \Theta_i = \theta\right) \Pr\left(\Theta_i = \theta\right) \quad (14)$$

Since the PCT outcomes have been assumed as independent,

$$\Pr\left(\Theta_i = \theta \mid \mathbf{y}^{(n)}\right) \propto \prod_{\ell=1}^{n} \Pr\left(y_\ell \mid \Theta_i = \theta\right) p_\theta. \quad (15)$$

Let Θ'_ℓ be the random variable representing the LoA of the ℓ-th opponent of Agent i. One has

$$\Pr\left(y_\ell \mid \Theta_i = \theta\right) = \sum_{k=1}^{K} \Pr\left(y_l, \Theta'_\ell = k \mid \Theta_i = \theta\right)$$

$$= \sum_{k=1}^{K} \Pr\left(y_\ell \mid \Theta_i = \theta, \Theta'_\ell = k\right) \Pr\left(\Theta'_\ell = k \mid \Theta_i = \theta\right).$$

Since the ℓ-th opponent is randomly chosen, one has

$$\Pr\left(\Theta'_\ell = k \mid \Theta_i = \theta\right) = \Pr\left(\Theta'_\ell = k\right) = p_k$$

and

$$\Pr\left(y_\ell \mid \Theta_i = \theta\right) = \sum_{k=1}^{K} \left[y_\ell q\left(\theta, k\right) + \left(1 - y_\ell\right)\left(1 - q\left(\theta, k\right)\right)\right] p_k$$

$$= y_\ell s_\theta + \left(1 - y_\ell\right)\left(1 - s_\theta\right) \tag{16}$$

for any $\ell = 1, \dots, n$, where one uses (6). Finally, since $y_\ell \in \{0, 1\}$, introducing (6) in (15), one gets

$$\Pr\left(\Theta_i = \theta \mid y_1, \dots, y_n\right) \propto s_\theta^m \left(1 - s_\theta\right)^{n-m} p_\theta. \tag{17}$$

The *a posteriori* probability in (8) may be rewritten as

$$\Pr\left(\Theta_i = \theta \mid \mathbf{y}^{(n)}, \theta'_{n_u+1}, \dots, \theta'_n\right)$$
$$\propto \Pr\left(\mathbf{y}^{(n)} \mid \Theta_i = \theta, \theta'_{n_u+1} \dots, \theta'_n\right) \Pr\left(\Theta_i = \theta \mid \theta'_{n_u+1} \dots, \theta'_n\right) \tag{18}$$

The knowledge of the LoAs of the n_k last opponents does not provide any information about the LoA of Agent i, thus

$$\Pr\left(\Theta_i = \theta \mid \theta'_{n_u+1}, \dots, \theta'_n\right) = \Pr\left(\Theta_i = \theta\right).$$

Moreover, since the PCT outcomes have been assumed as independent, one has

$$\Pr\left(\Theta_i = \theta \mid \mathbf{y}^{(n)}, \theta'_{n_u+1}, \dots, \theta'_n\right)$$
$$\propto \left(\prod_{\ell=1}^{n_u} \Pr\left(y_\ell \mid \Theta_i = \theta\right)\right) \left(\prod_{\ell=n_u+1}^{n} \Pr\left(y_\ell \mid \theta, \theta'_\ell\right)\right) p_\theta. \tag{19}$$

The first products are obtained from Proposition 1. A given term of the second product may be expressed as

$$\Pr\left(y_\ell \mid \theta, \theta'_\ell\right) = y_\ell q\left(\theta, \theta'_\ell\right) + \left(1 - y_\ell\right)\left(1 - q\left(\theta, \theta'_\ell\right)\right). \tag{20}$$

Finally, (9) is obtained combining (17), (19), and (20).

References

[1] *Pairwise Ranking Aggregation in a Crowdsourced Setting*. ACM, February 2013.

[2] Hock Hee Ang, Vivekanand Gopalkrishnan, Wee Keong Ng, and Steven Hoi. Communication-efficient classification in P2P networks. In *Proc. Joint Eur. Conf. on Mach. Learn. and Knowl. Disc. in Databases*, pages 83–98, Bled, Slovenia, 2009.

[3] C. Borcea, M. Talasila, and R. Curtmola. *Mobile Crowdsensing*. Chapman and Hall/CRC, 2016.

[4] R. A. Bradley and M. E. Terry. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

[5] Róbert Busa-Fekete and Eyke Hüllermeier. *A Survey of Preference-Based Online Learning with Bandit Algorithms*, pages 18–39. Springer International Publishing, Cham, 2014.

[6] Alessandro Chiuso, Fabio Fagnani, Luca Schenato, and Sandro Zampieri. Gossip algorithms for simultaneous distributed estimation and classification in sensor networks. *IEEE Jnl Sel. Top. Sig. Proc.*, 5(4):691–706, 2011.

[7] Fabio Fagnani, Sophie M Fosson, and Chiara Ravazzi. A distributed classification/estimation algorithm for sensor networks. *SIAM Jnl on Cont. and Optim.*, 52(1):189–218, 2014.

[8] J. Furnkranz and E. Hullermeier. *Preference Learning*. Springer-Verlag, 2011.

[9] L. Galluccio, B. Lorenzo, and S. Glisic. Sociality-aided new adaptive infection recovery schemes for multicast DTNs. *IEEE Trans. Veh. Techn.*, 65(5):3360–3375, 2016.

[10] R. Gentz, S. X. Wu, H.-T. Wai, A. Scaglione, and A. Leshem. Data injection attacks in randomized gossiping. *IEEE Transactions on Signal and Information Processing over Networks*, 2(4):523–538, 2016.

[11] Thore Graepel. Score-based bayesian skill learning. January 2012.

[12] B. Guo, Z. Wang, Z. Yu, Y. Wang, N.Y. Yen, R. Huang, and X Zhou. Mobile crowd sensing and computing: the review of an emerginghuman-powered sensing paradigm. *ACM Computing Surveys*, 48(1), September 2015.

[13] B. Kailkhura, S. Brahma, and P. K. Varshney. Data falsification attacks on consensus-based detection systems. *IEEE Transactions on Signal and Information Processing over Networks*, 3(1):145–158, 2017.

[14] B. Kantarci and H. T. Mouftah. Reputation-based sensing-as-a-service for crowd management over the cloud. In *Proc. IEEE ICC*, pages 3614–3619, Sydney, Australia, June 2014.

[15] W. Li, L. Galluccio, F. Bassi, and M. Kieffer. Distributed faulty node detection in delay tolerant networks: Design and analysis. *IEEE Transactions on Mobile Computing*, PP(99):1–1, 2017.

[16] Wenjie Li, Francesca Bassi, Laura Galluccio, and Michel Kieffer. Peer-assisted individual assessment in a multi-agent system. *Automatica*, 83(Supplement C):351 – 360, 2017.

[17] R. D. Luce. *Individual choice behaviour: A theoretical analysis*. Wiley, New-York, 1959.

[18] Ping Luo, Hui Xiong, Kevin Lü, and Zhongzhi Shi. Distributed classification in peer-to-peer networks. In *Proc. Int. Conf. on Knowl. Discov. and Data Mining*, pages 968–976. ACM, 2007.

[19] J. Ren, Y. Zhang, K. Zhang, and X. Shen. SACRM: Social Aware Crowdsourcing with Reputation Management in mobile sensing. *Comp. Commun.*, 65, July 2015.

[20] Nihar B Shah, Sivaraman Balakrishnan, and Martin J Wainwright. A permutation-based model for crowd labeling: Optimal estimation and robustness. *arXiv preprint arXiv:1606.09632*, 2016.

[21] G. Suarez-Tangil, J. E. Tapiador, P. Peris-Lopez, and A. Ribagorda. Evolution, detection and analysis of malware for smart devices. *IEEE Communications Surveys Tutorials*, 16(2):961–987, Second 2014.

[22] L. Thurstone. A law of comparative judgement. *Psychological Review*, 34:273–286, 1927.

[23] H. T. Waiy, A. E. Ozdaglarz, and A. Scaglione. Identifying susceptible agents in time varying opinion dynamics through compressive measurements. In *Proc. IEEE ICASSP*, pages 4114–4118, 2018.

[24] Z. Yu and M. Van der Schaar. Reputation-based incentive protocols in crowdsourcing applications. In *Proc. IEEE INFOCOM*, pages 2140–2148, Orlando, FL, 2012.

[25] W. Zamora, C. T. Calafate, J.-C. Cano, and P. Manzoni. A survey on smartphone-based crowdsensing solutions. *Mobile Information Systems*, 2016:26, 2016.

[26] Hongzi Zhu, Luoyi Fu, Guangtao Xue, Yanmin Zhu, Minglu Li, and L.M. Ni. Recognizing exponential inter-contact time in vanets. In *Proc. IEEE INFOCOM*, pages 1–5, San Diego, CA, March 2010.