COALITIONAL GAME THEORETIC APPROACH TO DISTRIBUTED ADAPTIVE PARAMETER ESTIMATION

Nikola Bogdanović, Dimitris Ampeliotis, Kostas Berberidis

Department of Computer Engineering and Informatics, University of Patras
& C.T.I RU-8, 26500, Rio - Patra, Greece
E-mails: {bogdanovic, ampeliot, berberid}@ceid.upatras.gr

ABSTRACT

In this paper, the parameter estimation problem based on diffusion least squares strategies is studied from a coalitional game theoretical perspective. The problem has been modeled as a non-transferable coalitional game and two scenarios have been considered, one where the value function includes only a suitable estimation accuracy criterion and another one in which the cost of coalition formation is taken into account as well. In the former scenario, we first analyze the non-emptiness of the core of the games corresponding to traditional diffusion strategies, and then, we extend the analysis to a recently proposed node-specific parameter estimation setting where the nodes have overlapped but different estimation interests. In the latter scenario, we employ a distributed coalition formation algorithm, based on merge-and-split steps, which converges to a stable coalition structure.

Index Terms—Adaptive distributed networks, diffusion algorithm, cooperation, node-specific parameter estimation, coalitional game theory, NTU game.

1. INTRODUCTION

Several low-complexity distributed strategies for parameter estimation, based on least mean squares (LMS), have been studied in the literature, i.e., the consensus, the incremental and the diffusion strategy (see [1]-[2] and references therein). In most of the existing papers, it is assumed that the nodes have exactly the same interests. More recently, research efforts have focused on removing this restriction. In [3]-[7], a Node-Specific Parameter Estimation (NSPE) formulation has been introduced where the nodes have overlapped but different estimation interests. Also, in [8], the authors proposed an algorithm for a scenario where nodes have numerically similar estimation interests, while in [9], for the same scenario, the performance of the diffusion strategy has been studied. However, it is of interest to study the problem of parameter estimation over networks in the case where nodes or groups of nodes are allowed to be selfish using concepts from game theory.

In general, game theory can be defined as the study of mathematical models of conflict and cooperation between intelligent rational decision-makers. It has been applied to various disciplines such as economics, political sciences, philosophy and more recently, to engineering [10]. Unlike non-cooperative game theory where the modeling unit is a single player, coalitional game theory, that is the focus of this work, seeks for optimal coalition structure of players in order to optimize the worth of each coalition. According to [11], coalitional games can be classified into the following three categories, i.e., canonical coalitional games, coalition formation games and coalitional graph games.

The literature dealing with the analyses of the adaptive networks from the game theoretic perspective has been rather limited. Most studies focus mainly, although not exclusively, on the game-theoretical approaches based on non-cooperative game theory. The authors in [12]-[13] allow the nodes to be selfish while minimizing their own cost functions that combine both the estimation accuracy and the communication cost. In the pairwise one-shot successive game setting [12], after showing that the dominant strategy is not to cooperate, the authors proposed a reputation mechanism in order to motivate cooperation among selfish nodes. In the same setting, they provide an approximate cluster formation protocol in [13], where the nodes decide in a pairwise manner whether they should merge according to the estimation gain and communication cost associated with that action. A game-theoretic approach to node activation control in parameter estimation via diffusion LMS has been considered in [14]. The energy-aware activation control is modeled as a non-cooperative repeated game, where the aim for each node is to be activated only when its contribution outweighs the activation cost, and it is able to track the set of approximate correlated equilibria of the underlying activation control game. In [15], the authors formulated the distributed adaptive filtering problem as a graphical evolutionary game under the imitation strategy updating rule, and proved that the strategy of using information from nodes with good signal is always an evolutionarily stable strategy. In all these studies, the aim of the nodes was to estimate a set of global parameters that were identical to all nodes.

In this work, we analyze a distributed adaptive parameter estimation problem in the framework of coalitional game theory. We consider the coalitional game to be the Non-Transferable Utility (NTU) game for which the choice of coalitional actions defines each player’s payoff. The contribution of our work is two-fold: 1) we study the parameter estimation problem via diffusion strategy as a canonical game and extend the analysis to an NSPE setting, and 2) we propose coalition formation game for the specific NSPE setting when the coalition formation cost is considered.

2. CANONICAL GAMES FOR PARAMETER ESTIMATION

In this section, we analyze a distributed adaptive parameter estimation problem for a scenario where the value function includes only a
suitable estimation accuracy criterion. One could select various measures of the estimation performance, e.g., estimation error, speed of convergence, etc. In this work, we consider Mean Square Deviation (MSD) in the steady-state to be minimized at each node $k$. The nodes of the network represent the players in the game, while we want to study whether and under which conditions the players will form the grand coalition, i.e., the case where all the nodes of the network cooperate. To study the stability of the grand coalition, we will use the well-established solution concept of the core.

2.1. MSD game definition

Let us define an NTU coalitional game $(N, v)$, where $N = \{1, \ldots, N\}$ is the set of players (nodes) while $v(S) \subseteq \mathbb{R}^N$ is the set of vector values of a coalition $S \subseteq N$. The maximum estimation accuracy that a certain node $k \in S$ may achieve in a coalition $S$ is the payoff $x_k$, where $x_k$ is the $k^{th}$ element of payoff vector $x \in v(S)$, and is given by

$$x_k(S) = -\text{MSD}_k(S) = -E[\hat{w}_{k,S} - w^o]^2.$$  \hfill (1)

Note that in (1), $\hat{w}_{k,S}$ denotes a steady-state estimate of $w^o \in \mathbb{C}^{M \times 1}$ that a node $k$ may achieve while cooperating, through the diffusion LMS strategy (see Eq. 16 in [2]), with the nodes that also belong to the same network subset $S \subseteq N$. Naturally, each node is rationally selfish in the sense that it aims to maximize its own payoff.

2.2. Data assumptions

To convey the main ideas of this work, it is sufficient to assume that the data $\{d_k(i), u_k,i\}$ are realizations of zero-mean jointly wide-sense stationary random processes related through the model

$$d_k(i) = u_k,i w^o + v_k(i)$$  \hfill (2)

and satisfying the following assumptions:

A1) the regressors $u_{k,i} \in \mathbb{C}^{1 \times M}$ are temporally and spatially white with the autocovariance $R_{u_k} = R_u > 0$ equal for all nodes,

A2) $v_k(i) \in \mathbb{C}$ is temporally and spatially white noise, of variance $\sigma_{v_k}^2$, and which is independent of $u_{k,j}$ for all $k, i, l, j$.

A3) the step sizes $\mu_k = \mu$ in the LMS recursion of diffusion strategy are equal for all nodes, and are sufficiently small so that the higher-order terms of $\mu$ can be ignored.

2.3. MSD game analysis

To analyze the MSD game defined in 2.1, we first need to provide some definitions and concepts from the literature.

Firstly, in order to classify an NTU game as canonical, it is required to be [11]:

- in characteristic form, i.e., the value of a coalition $S$ is determined exclusively by the members of that coalition
- superadditive, i.e.,
  $$v(S_1) \cap v(S_2) \subseteq v(S_1 \cup S_2),$$  \hfill (3)
  $$\forall S_1 \subset N, S_2 \subset N, \text{ s.t. } S_1 \cap S_2 = \emptyset.$$

Due to the fact that, in a canonical game, cooperation is never harmful with respect to the non-cooperative case, it is important to study the properties of the grand coalition, i.e., the coalition of all nodes. To this end, we will employ the probably most studied solution concept, known as the core.

The core of a canonical game is the set of payoff allocations for which it holds that no coalition $S \subseteq N$, $S \neq \emptyset$ has an incentive to split off and leave the grand coalition. In order to solve a NTU game using the core, the following standard conditions should hold for a family of sets $v = (v(S))_{S \subseteq N}$ [16], [11]: (i) The value $v(S)$ must be a non-empty closed subset of $\mathbb{R}^N$, (ii) The value $v(S)$ must be comprehensive, i.e., if $x \in v(S)$ and $y \in \mathbb{R}^N$ are such that $y_k \leq x_k \forall k \in S$, then $y \in v(S)$, (iii) The set $\{x : x \in v(S) \text{ and } x_k \geq x_k, \forall k \in S\}$ with $x_k = \max\{y_k : y \in v(S)\} < \infty$, $\forall k \in N$ must be a bounded subset of $\mathbb{R}^N$. Finally, for a canonical NTU game, the core is defined as

$$C_{\text{NTU}} = \{x \in v(N) : \forall y \in v(S), \exists y \in \text{v(S), s.t. } x_k > x_k, \forall k \in S\}. \hfill (4)$$

However, the core of NTU canonical games are not always guaranteed to exist. Therefore, it is necessary to examine the non-emptiness of the core.

In general, proving the non-emptiness of the core is an $NP$-complete problem since the number of possible coalition structures grows exponentially with the number of players [17]. A way to prove that the core is nonempty is by showing that the game being analyzed is balanced [16]-[18], since the balanced condition is sufficient (yet not necessary) in case of an NTU game and is given by

$$\bigcap_{S \subseteq B} v(S) \subseteq v(N), \forall B$$  \hfill (5)

where $B$ is a balanced subsets’ family, i.e., a family of nonempty, proper subsets of $N$ where there exist positive weights $\delta_S$ for $S \subseteq B$ such that

$$\sum_{S \subseteq B, k \in S} \delta_S = 1, \forall k \in N.$$  \hfill (6)

Now, we analyze the MSD game defined in Subsec. 2.1. Note that, under the assumptions made in Subsec. 2.2, the MSD achieved at each node $k$ in a connected network of $N$ nodes, can be well-approximated by the following expressions [2, p.163]:

1) in case of doubly stochastic combination weights

$$\text{MSD}_{k}^{\text{doubly}}(N) \approx \frac{\mu M}{2} \cdot \frac{1}{N} \left( \frac{1}{N} \sum_{k=1}^{N} \sigma_{v_k}^2 \right), \hfill (7)$$

2) in case of combination weights using the Hastings rule

$$\text{MSD}_{k}^{\text{Hasting}}(N) \approx \frac{\mu M}{2} \cdot \left( \sum_{k=1}^{N} \sigma_{v_k}^{-2} \right)^{-1}. \hfill (8)$$

Theorem 1. Under the assumptions A1-A3, the MSD game with doubly stochastic combination weights is, in general, not canonical and has empty core.

\textbf{Proof}. The proof is based on assuming the case of arbitrarily different noise variances at different nodes which indicates that the game is not superadditive in general.

Theorem 2. The MSD game with Hastings rule combination weights, under the assumptions A1-A3, is a canonical one and has non-empty core.

\textbf{Proof}. The proof is based on the game superadditivity due to (8), and proving the core non-emptiness by showing that the game is balanced which is a sufficient condition.
2.4. Extension to NSPE-MSD game

In this subsection, the coalitional game-theoretical discussion is extended to a more general, Node-Specific Parameter Estimation (NSPE) setting. Let us assume that there are two vectors of parameters affecting an area of \( N \) nodes, i.e., \( w^a_k \) and \( w^b_k \), where not all the nodes are influenced by both vectors of parameters. The subset of nodes affected by \( w^a_k \) is denoted by \( \mathcal{P}_a \) while \( \mathcal{P}_b \) stands for the subset of nodes interested in estimating \( w^b_k \), see Fig. 1. In this setting, which can be seen as a special case of the NSPE formulation introduced in [3]-[7], the data of each node \( k \) is related through the model

\[
d_k(i) = \begin{cases} \sum_{l \in P_a \cap P_b} \mu_{k,l} w^a_k + \sum_{l \in P_a \setminus P_b} \mu_{k,l} w^b_k + v_k(i) & \text{if } k \in \mathcal{P}_a \cap \mathcal{P}_b, \\ \sum_{l \in P_a \setminus \mathcal{P}_b} \mu_{k,l} w^a_k + v_k(i) & \text{if } k \in \mathcal{P}_a \setminus \mathcal{P}_b, \\ \mu_{k,P_b} w^b_k + v_k(i) & \text{otherwise.} \end{cases} \tag{9} \]

To simplify the analysis for this NSPE scenario, besides the assumptions made in Subsec. 2.2, we further assume that

A4) the regressors \( u_{k,l}^a \in \mathbb{C}^{1 \times M_a} \) and \( u_{k,l}^b \in \mathbb{C}^{1 \times M_b} \) are independent.

Even under assumptions A1-A2 and A4, regardless of the cooperation strategy, there is a coupling between the estimation processes related to \( w^a_k \) and \( w^b_k \) in the general case. This coupling is due to the influence of higher order data moments that are multiplied by \( \mu^2 \) in the expressions for mean-square performance, (see discussion for standalone LMS in [19]), or for cooperative NSPE settings in [6]-[7]. However, for sufficiently small step size (A3), this effect can be ignored. Therefore, under assumptions A1-A4, and for the observation model given in (9), the MSD achieved at each node \( k \) in the considered diffusion NSPE setting can be well-approximated as follows

\[
\text{MSD}_k(\mathcal{N}) \approx \text{MSD}_{k,u^a} (\mathcal{P}_a) + \text{MSD}_{k,u^b} (\mathcal{P}_b) \tag{10}
\]

where \( \text{MSD}_{k,u^a} (\mathcal{P}_j) \) \( \forall j \in \{a, b\} \), and, for instance, for Hasting combination rule, it is given by

\[
\text{MSD}_{k,u^a} (\mathcal{P}_j) = \begin{cases} \frac{\sigma_{\text{MSD}}}{2} \cdot \left( \sum_{k \in \mathcal{P}_j} \sigma_{v_k}^2 \right)^{-1} & \text{if } k \in \mathcal{P}_j, \\ 0 & \text{otherwise,} \end{cases} \tag{11}
\]

Now, let us define the NSPE-MSD game by extending the MSD game from Subsec. 2.1. In particular, let us define a player as a node per each estimation task that is within its interest. For instance, the node 7 in Fig. 1 is representing two players, one for each estimation task. Finally, we generalize Thm 2 for the described NSPE scenario as follows:

**Theorem 3.** The NSPE-MSD game with Hasting rule combination weights, under the assumptions A1-A4, is a canonical one and has non-empty core.

**Proof.** The proof is a generalization of the proof of Thm 2. \( \square \)

3. COALITION FORMATION GAME FOR PARAMETER ESTIMATION

So far, we have analyzed the (NSPE-)MSD games and the formation of the grand coalition in case where the nodes cooperate without any cost. Now, the goal is to find the coalition structure that enables the nodes to maximize their utilities while taking into account the cost of coalition formation as well.

3.1. NSPE-MSD-COMM game definition

We define an NTU coalitional game \((\mathcal{N}^n, v)\) where we model now each player as a node per its estimation task with \( |\mathcal{N}^n| = |\mathcal{P}_a| + |\mathcal{P}_b| \). Furthermore, the payoff \( x_k^{(j)} \) achieved by each node \( k \) in a coalition \( S^j \) per each estimation task \( j \in \{a, b\} \) where \( k \in \mathcal{P}_j \), as:

\[
x_k^{(j)}(S^j) = gain_k^{(j)} - c_k^{(j)} \tag{12}
\]

where

- \( gain_k^{(j)} = \text{MSD}_{k,w^j} (k) - \text{MSD}_{k,w^j} (S^j) \),

- \( c_k^{(j)} \) stands for the total communication cost that is incurred by a node \( k \) in order to establish a coalition \( S^j \).

3.2. NSPE-MSD-COMM game analysis

Let us now analyze the grand coalition formation of the NSPE-MSD-COMM game defined above.

**Theorem 4.** The NSPE-MSD-COMM game with either doubly stochastic combination weights or with Hasting rule combination weights has, in general, empty core.

**Proof.** The proof is based on assuming the case of two arbitrary disjoint coalitions \( S^j_1 \), \( S^j_2 \), where \( k \in S^j_1 \), that are far enough (high communication cost) so that \( x_k(S^j_1 \cup S^j_2) < x_k(S^j_1) \), implying also that the game is not superadditive in general. \( \square \)

Due to the fact that the grand coalition will not be formed in general, we aim to apply a coalition formation algorithm and analyze the properties of the coalitional structure being formed. In particular, we focus on the so-called Merge-Split approach, initially presented in [20], and applied in the context of several applications [11],[21]-[24].

To this purpose, let us make the following definitions. A collection of coalitions \( \mathcal{S} \) is the set of mutually disjoint coalitions, i.e., \( S = \{S_1, \ldots, S_l\} \), where \( S_n \subset \mathcal{N} \) for \( n = 1, \ldots, l \). If a collection \( \mathcal{S} \) comprises all the nodes of \( \mathcal{N} \), then the collection \( \mathcal{S} \) is a partition of \( \mathcal{N} \). The Pareto order operator \( \triangleright \) for comparing two collections \( \mathcal{R} = \{R_1, \ldots, R_m\} \) and \( \mathcal{S} = \{S_1, \ldots, S_l\} \), that are partitions of the same subset of nodes \( \mathcal{A} \subset \mathcal{N} \), is defined as follows

\[
\mathcal{R} \triangleright \mathcal{S} \iff x_k(\mathcal{R}) \geq x_k(\mathcal{S}), \quad \forall k \in \mathcal{R}, \mathcal{S} \tag{13}
\]

with at least one node satisfying the strict inequality \( \succ \).

Now, we can define a distributed algorithm based on two simple steps, i.e., merge and split [20],[21], as follows,

- \( \forall j \in \{a, b\} \), merge any set of coalitions \( \{S^j_1, \ldots, S^j_l\} \) if it holds \( \bigcup_{n=1}^l S^j_n \triangleright \{S^j_1, \ldots, S^j_l\} \),

- \( \forall j \in \{a, b\} \), split any coalition \( \bigcup_{n=1}^l S^j_n \triangleright \{S^j_1, \ldots, S^j_l\} \), if it holds \( \bigcup_{n=1}^l S^j_n \triangleright \{S^j_1, \ldots, S^j_l\} \).
In other words, $|P_j|$-player game is formed using merge-split operations for each estimation task $j$. Note that coalitions will merge (or split) only if there is at least one node that is improving its payoff while there is no decrease in the payoffs of the other nodes of the coalitions involved in the merge (or split) operation. The algorithm terminates for any initialization, and the coalition structure always converges to a $D_{hp}$-stable coalition structure [20], i.e., no player has incentive to leave its coalition. In order to account for time-varying noise variances and/or node mobility, the coalition formation process may be repeated periodically during the network operation.  

4. SIMULATIONS

Firstly, in order to verify Thm 1-3, obtained based on the approximate MSD expressions (7)-(8), we simulate a) the setting where all the nodes estimate a common parameter vector and b) the NSPE setting. We consider a network of $N = 10$ nodes with the measurements following the observation model (2) with $M = 8$ and $\mu = 5 \cdot 10^{-3}$. The regressors $u_{k,t}$ are zero mean with autocovariance $R_u = I$. The noise variances for nodes $4$ and $7$ are $\sigma^2_{o,4} = \sigma^2_{o,7} = 8 \cdot 10^{-3}$, while for all other nodes they are chosen from the interval $(0.04, 0.2)$. We compare the LMS-based non-cooperative strategy and the diffusion LMS, where nodes cooperate all together (grand coalition) and where there are two disjoint coalitions of nodes, i.e., $S_1 = \{4, 7\}$ and $S_2 = N \setminus S_1$. In Fig. 2, it can be seen that the grand coalition does not bring improvement to all nodes in the case of doubly-stochastic combination weights. On the contrary, in the Hasting case, the grand coalition makes all nodes better off. For the NSPE scenario with two estimation tasks, i.e., $w^a_o$ and $w^b_o$, we set $P_a = \{1, \ldots, 7\}$, $P_b = \{5, \ldots, 10\}$ and $P_a \cap P_b = \{5, 6, 7\}$, as in Fig. 1. The filter lengths are chosen to be the same, i.e., $M_a = M_b = 8$. Again, it is shown that the grand coalition for the Hasting case is stable (Fig. 3).

To illustrate the effectiveness of the proposed NSPE merge-split protocol we provide some indicative simulation results. The protocol is practically implemented as follows. For each estimation task $j \in \{a, b\}$, any coalition $T_j^k$ from an initial network partition $T$ starts the merging process by performing pairwise negotiations with other coalitions. If a merge occurs, the newly formed coalition continues the search for merging until it is possible. Then, the merging process is repeated for all other coalitions from $T$ that have not been merged yet. Afterwards, the resulting coalition(s) are performing split operations, if any is possible. We model communication cost between the nodes $k$ and $\ell$ with a simple exponential model, i.e., $c_{k,\ell} = T \cdot e^{d_{k,\ell}/d_0}$, where $d_0$ is a reference distance, $d_{k,\ell}$ denotes the distance between the nodes $k$ and $\ell$, while $T$ is a normalization coefficient. Assuming the broadcast nature of the communication, the total communication cost for a node $k$ in order to establish a coalition $S^j$ is given by $C_{k,S^j} = \max_{\ell \in S^j} c_{k,\ell}$. We consider a network of $N = 10$ nodes where each node $k$ has a noise variance $\sigma^2_{o,k}$ between $0.1$ and $0.6$. There are two vectors of parameters to be estimated, i.e., $w^a_o$ and $w^b_o$, with $P_a = \{1, \ldots, 8\}$ and $P_b = \{6, \ldots, 10\}$. We set $\mu = 0.001$ and $M_a = M_b = 10$. We evaluate the MSD expressions for the Hasting case. In Fig. 4, we plot the sizes of coalitions resulting from the NSPE merge-split protocol as a function of the communication cost (normalization coefficient). For each estimation task, we plot the maximum and the average coalition size. The results have been averaged over 100 experiments where we randomly set $d_{k,\ell}$ between 0 and 5. The figure shows that for zero (or relatively small) communication costs, the coalition structures for estimating $w^a_o$ and $w^b_o$ correspond to $P_a = 8$ and $P_b = 5$, respectively, i.e., the grand coalition. As the communication costs increase, the coalitions get split, and finally, they reduce to the non-cooperative nodes.

5. CONCLUSIONS

We have modeled the distributed adaptive parameter estimation problem as a non-transferable coalitional game. Initially, we have studied the parameter estimation problem via diffusion strategy as a canonical game and we have extended the analysis to NSPE setting. Then, we have proposed a coalition formation game for NSPE setting when the coalition formation cost is taken into account. The proposed algorithm may adapt to a dynamic environment as the coalitions can periodically merge or split depending on the nodes’ distance-based communication cost and/or time-varying noise variances. Our future work may consider extending the aforementioned games to a coalitional graph games setting.
6. REFERENCES


