Engineering Applications of Artificial Intelligence 56 (2016) 157–174

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

# A hierarchical consensus method for the approximation of the consensus state, based on clustering and spectral graph theory

Rita Morisi, Giorgio Gnecco*, Alberto Bemporad

*IMT School for Advanced Studies, Piazza S. Francesco, 19, 55100 Lucca, Italy*

ARTICLE INFO

*Article history:*
Received 15 April 2016
Received in revised form
21 June 2016
Accepted 28 August 2016

*Keywords:*
Consensus problem
Approximation
Hierarchical consensus
Clustering
Spectral graph theory

ABSTRACT

A hierarchical method for the approximate computation of the consensus state of a network of agents is investigated. The method is motivated theoretically by spectral graph theory arguments. In a first phase, the graph is divided into a number of subgraphs with good spectral properties, i.e., a fast convergence toward the local consensus state of each subgraph. To find the subgraphs, suitable clustering methods are used. Then, an auxiliary graph is considered, to determine the final approximation of the consensus state in the original network. A theoretical investigation is performed of cases for which the hierarchical consensus method has a better performance guarantee than the non-hierarchical one (i.e., it requires a smaller number of iterations to guarantee a desired accuracy in the approximation of the consensus state of the original network). Moreover, numerical results demonstrate the effectiveness of the hierarchical consensus method for several case studies modeling real-world networks.


## 1. Introduction

The theory of complex systems deals with the study of the behavior of systems made of several agents (or units) that interact among each other; typical examples are social (Del Vicario et al., 2016) and economic (Battiston et al., 2016) networks, physical systems made of interacting particles (Castellano et al., 2009), and biological (Pastor-Satorras et al., 2015) and ecological (Vivaldo et al., 2016) systems. In all these cases, one has often to deal with a large number of units, which have no global knowledge about the structure of the whole system, as their interactions are limited to their neighbors in the network. Control problems on such systems are strongly influenced by structural properties of their graph of interconnections, described, e.g., in terms of a weighted/unweighted adjacency or graph-Laplacian matrix (Mesbahi and Egerstedt, 2010; Liu et al., 2011). In particular, several studies (see, e.g., Lovisari and Zampieri, 2012 for a tutorial) deal with the analysis of the conditions under which a complex system has all its agents reach asymptotically a common state, called consensus state (i.e., they agree asymptotically with the same opinion) and, in case of a positive answer, with investigating the rate of convergence to the consensus state. It is well-known (see, e.g., Boyd et al., 2004; Lovisari and Zampieri, 2012) that such a convergence rate is related to the spectral properties of the graph of interconnections (e.g., the ones of a transition probability matrix one can associate to it). The work (Boyd et al., 2004) optimizes such properties by solving a suitable convex optimization problem, called *Fastest Mixing Markov-Chain* (FMMC) problem. In our previous work (Gnecco et al., 2015), we optimized a suitable trade-off between the rate of convergence to the consensus state and the sparsity of the graph of interconnections, which is a way to insert in the model a possible cost of communication associated with each link used. In more details, the optimization problem considered in Gnecco et al. (2015) (which is a substantial extension of the conference paper, Gnecco et al., 2014) is an $l_1$-norm (convex) regularization of the FMMC problem, called FMMC-$l_1(\eta)$ problem, where $\eta > 0$ is a regularization parameter. Its main contributions are some theoretical results about the choice of $\eta$ to avoid triviality of the resulting optimal solution, and an interpretation of the FMMC-$l_1(\eta)$ problem as a robust version of the FMMC problem, in which one is allowed to select only nominal weights associated with the edges of the graph, as such weights enter the model together with an intrinsic relative uncertainty, which cannot be removed unless the nominal values are chosen to be equal to 0. A (nonconvex) $l_0$-pseudo-norm regularized version of the FMMC problem is also analyzed in Gnecco et al. (2015). Some ways to restrict the search for its optimal solution to suitable feasible solutions are also investigated therein. Finally, numerical results demonstrate the effectiveness of both regularized approaches (with computational advantages for the convex case) in achieving – as desired – a "good" trade-off between sparsity of the network and its rate of convergence to the consensus state.

* Corresponding author.
*E-mail addresses:* rita.morisi@imtlucca.it (R. Morisi), giorgio.gnecco@imtlucca.it (G. Gnecco), alberto.bemporad@imtlucca.it (A. Bemporad).

http://dx.doi.org/10.1016/j.engappai.2016.08.018
0952-1976/

The approach followed in this paper is substantially different from Gnecco et al. (2015), although the goal is similar. In more details, the main idea of the present work is the following: for a fixed network topology, we aim at speeding up consensus using a "hierarchical" approach, whose theoretical motivation relies on spectral properties of the agents' network. Our approach is based on dividing the original connected graph into many connected subgraphs, which are expected (due to spectral graph theory arguments, Chung, 1997) to have "good" spectral properties. In this case, the rate of convergence to the "local" consensus state (i.e., the consensus state of each subgraph) is faster than the one to the "global" consensus state of the original graph. In a second phase, the resulting approximations of the local consensus states of the subgraphs are mixed to get (up to a certain tolerance) global consensus on an auxiliary graph, whose nodes are selected nodes of the subgraphs (one for each subgraph), and for which "good" spectral properties are still expected (again, due to spectral graph theory arguments). To generate the subgraphs, we apply both a technique known as spectral clustering (von Luxburg, 2007), and a second ad hoc technique that we call *nearest supernode approach*, which are both expected to extract sufficiently "dense" subgraphs (e.g., made of a single cluster of nodes, with each node connected directly to several other nodes of the same cluster). For such subgraphs, the rate of convergence to the local consensus state is relatively fast (since the second-largest eigenvalue modulus of the transition probability matrix associated with each such subgraph is relatively small). In this way, in the hierarchical approach, one fixes the sparsity of the graph, then speeds up the approximation of its consensus state possibly even more than through the resolution of the FMMC problem, since the latter does not allow for a hierarchical solution. It is worth noting that, in case the original graph is not sparse, one can still apply the hierarchical consensus method described in the paper after a preliminary step of edge sparsification (this could be achieved, e.g., applying the algorithms detailed in Batson et al., 2013), to construct another graph with a very similar spectral behavior, but with a (typically much) smaller number of edges. Then, the hierarchical consensus method could be applied directly to this sparsified graph. It has to be remarked that approaches similar to the one presented in this paper have been proposed also in Epstein et al. (2008) and Li and Bai (2012). In such works, the multi-agent system is also decomposed into a hierarchical structure. Nevertheless, neither Epstein et al. (2008) nor Li and Bai (2012) consider techniques that exploit spectral graph theory arguments for the generation of the subgraphs. Hence, compared with Epstein et al. (2008) and Li and Bai (2012), the main original contribution of the present work lies on the techniques we adopt to determine the different connected subgraphs, and on the theoretical motivations we provide for such techniques, based on spectral graph theory arguments. In addition to this, we perform an extensive numerical evaluation of the hierarchical consensus method on several case studies modeling real-world networks, achieving in most cases better performance with respect to a non-hierarchical consensus method.

The paper is structured as follows. Section 2 presents an introduction to the consensus problem, and provides an overview of the hierarchical consensus method. Section 3 provides some theoretical arguments supporting the method, based on spectral graph theory. Section 4 describes clustering techniques used by the method, whereas Section 6 provides a study of its approximation of the global consensus state. In Section 7, numerical examples are presented. Section 8 provides a refinement of the basic setting, based on the results of the numerical examples. Finally, Section 9 offers conclusions.

## 2. An overview of the hierarchical consensus method

Let $G = (V, E)$ be a connected undirected graph with $N = |V|$ nodes and $|E|$ edges. In the context of the paper, the nodes represent agents (or units), which locally interact among each other. Such an interaction is governed by non-negative weights associated with the edges, which have to be chosen in a suitable way. Assuming a linear time-invariant model and describing each agent as a 1-dimensional dynamical system, the consensus problem refers to the investigation of the convergence to the consensus state (see the next formula (2)), for the following linear dynamical system:

$$\underline{x}(t+1) = P\underline{x}(t), \tag{1}$$

where the column vector $\underline{x}(t) \in \mathbb{R}^N$ contains the states (opinions) of the $N$ agents at a generic discrete time instant $t$, while $P \in \mathbb{R}^{N\times N}$ is a symmetric doubly stochastic matrix (i.e., $P_{ij} \geq 0$ for all $i, j = 1, \ldots, N$, $P\underline{1} = \underline{1}$, and $P = P^T$, where $\underline{1}$ is the $N$-dimensional vector whose components are all equal to 1). Moreover, $P_{ij} = 0$ when the two nodes $i$ and $j$ are different and are not linked by an edge. Due to the stated assumptions, $P$ can be interpreted as the matrix of transition probabilities associated with a finite-states Markov chain, possibly containing self-loops, since $P_{ii} \geq 0$ for all $i \in 1, \ldots, N$. If all the diagonal entries of $P$ are positive and the weighted graph associated with $P$ is connected, then it is well-known (see, e.g., Lovisari and Zampieri, 2012) that, for the $i$th component $x^{(i)}(t)$ of $\underline{x}(t)$, one has

$$x^{(i)}(t) \xrightarrow{t\to\infty} \frac{1}{N}\underline{1}^T \underline{x}(0), \quad \forall\, i = 1, \ldots, N, \tag{2}$$

with $\underline{x}(0)$ being the vector of the initial opinions of the agents. The expression $\Sigma = \frac{1}{N}\underline{1}^T \underline{x}(0)$, which is the average of the initial opinions of the agents, is the consensus state of the system.[1]

It is well-known (see, e.g., Como et al., 2012; Fagnani, 2014) that, at any discrete time instant $t$, the distance from the consensus state can be bounded from above as a function of the second-largest eigenvalue modulus $\mu(P)$ of the matrix $P$, in the following way:

$$\left\| \underline{x}(t) - \frac{1}{N}\underline{1}\,\underline{1}^T \underline{x}(0) \right\|_2^2 \leq \mu^{2t}(P) \|\underline{x}(0)\|_2^2, \tag{3}$$

where $\|\cdot\|_2$ denotes the $l_2$ norm.[2] For a given $P$, this rate of convergence cannot be improved, since there exist choices of the initial state $\underline{x}(0)$ for which a better rate cannot be obtained. Using (3), the rate of convergence to the consensus state was optimized in Boyd et al. (2004) by solving a suitable convex optimization problem, whose optimization variables are the entries of the matrix $P$. Differently from that approach, in the paper we intend to speed up consensus by considering local consensus subproblems formulated on different subgraphs $G_m = (V_m, E_m)$ of the original

[1] Since in the paper we are dealing with undirected graphs, hence with symmetric transition probability matrices, the consensus state is the average of the initial opinions of the agents. Without this assumption, the consensus state belongs only to the convex hull of the set of such opinions. To distinguish between these two situations, the consensus problem considered in this paper is sometimes called "average" consensus problem (Lovisari and Zampieri, 2012).

[2] The proof of (3) is as follows (see also Como et al., 2012). The matrix $P$ has the eigendecomposition $P = \frac{1}{N}\underline{1}\,\underline{1}^T + \sum_{j=1}^{N-1} \lambda_j \underline{v}_j \underline{v}_j^T$, where the eigenvalues are 1 and, for $j = 1, \ldots, N-1$, $\lambda_j$ (with $|\lambda_j| \leq \mu(P)$). The corresponding unit-norm and orthogonal eigenvectors are $\frac{1}{\sqrt{N}}\underline{1}$ and, for $j = 1, \ldots, N-1$, $\underline{v}_j$. Then, using also (1), one gets

$$\left\| \underline{x}(t) - \frac{1}{N}\underline{1}\,\underline{1}^T \underline{x}(0) \right\|_2^2 = \left\| \sum_{j=1}^{N-1} \lambda_j^t \underline{v}_j \underline{v}_j^T \underline{x}(0) \right\|_2^2 = \sum_{j=1}^{N-1} \left\| \lambda_j^t \underline{v}_j \underline{v}_j^T \underline{x}(0) \right\|_2^2 \leq \mu^{2t}\, \|\underline{x}(0)\|_2^2.$$