FAST AVERAGE GOSSIPING UNDER ASYMMETRIC LINKS IN WSNS

César Asensio-Marco[†] *and Baltasar Beferull-Lozano*[†‡]

[†]Group of Information and Communication Systems, Universidad de Valencia, Spain
[‡]Department of Information and Communication Technologies, University of Agder, Norway
E-mail: {Cesar.Asensio, Baltasar.Beferull}@uv.es

ABSTRACT

Wireless Sensor Networks are a recent technology where the nodes cooperate to obtain, in a totally distributed way, certain function of the collected data. An important example of these distributed processes is the average gossip algorithm, which allows the nodes to obtain the global average by only using local data exchanges. This process is traditionally slow, but can be accelerated by introducing geographic information or by exploiting the broadcast nature of the wireless medium. However, when a gossip protocol utilizes long geographic routes or broadcast communications, its convergence is not easily guaranteed due to asymmetry in communications. Alternatively, we propose an asymmetric version of the gossip algorithm that exploits residual information involved in each asymmetric exchange. Our asymmetric gossip algorithm achieves convergence faster than existing studies in the related literature. Numerical results are presented to show clearly the validity and efficiency of our approach.

1. INTRODUCTION

The advent of Wireless Sensor Networks (WSNs) has recently attracted a great deal of research work providing a new scenario where sensor nodes can perform intelligent cooperative tasks under strong constraints in terms of communications and energy resources. In this scenario, it is crucial to devise new decentralized schemes of estimation using only in-network processing capabilities. In other words, decentralized algorithms for estimation, where each node only exchanges information with its immediate neighbors in each communication step. As an important example, the average consensus problem is an instance of a distributed problem which comes up in many applications such as coordination of autonomous agents [7], estimation of fields [10], etc. The goal of any average consensus algorithm is to obtain, in a distributed way, the average of the sensor data by processing the measurements collected by sensor nodes. It avoids the need of performing all the computations at one or more sink nodes, thus, reducing congestion around these nodes and incrementing the robustness of the network.

This work was supported by the Spanish MEC Grants TEC2010-19545-C04-04 “COSIMA”, CSD2008-00010 “COMONSENS”, the European STREP Project “HYDROBIONETS” Grant 287613 within the FP7 and by the Norwegian RFF Agder CIEMCoE Grant.

The simplest gossip algorithm solving the average consensus problem is the pairwise gossip [1], where each node randomly picks up a neighbor and iteratively computes a symmetric pairwise average with it. This approach introduces an interesting asynchronous mechanism to the original deterministic protocol for consensus [8] and the nodes are still able to converge to the average with certain accuracy. However, this approach suffers from the locality of the updates, taking lot of iterations to reach consensus. Since each iteration involves an associated communication cost, geographic routes and broadcast communications has been proposed to reduce the total number of performed iterations. However, ensuring symmetric long routes or efficient broadcast communications requires complex control mechanisms in practice.

Therefore, executing gossip algorithms and ensuring their convergence to the average under asymmetric communications is crucial in a real WSN. The majority of the existing related works assume symmetric exchanges of data to asymptotically ensure convergence to the average, with the notable exceptions of [4], [6] and [9]. In this work, we propose a novel gossip algorithm, based on the residual information generated when an asymmetric communication is performed. We exploit this information by making the residuals evolve appropriately to preserve the summation of the process and to accelerate it. Moreover, our proposal is useful in the cases of having both unicast and broadcast communications and we show that it presents faster convergence in both schemes, as compared with existing approaches. Numerical results are presented to show clearly the validity and the efficiency of our approach.

The remainder of this paper is structured as follows: The problem formulation and the motivation of this work is given in Section II. A detailed explanation of the proposed gossip protocol is presented in Section III. We then present, in Section IV, some numerical results to show the convergence performance of our approach. Finally, the conclusions are summarized in Section V.

2. BACKGROUND

In this section, we first revise the necessary background in graph theory to model a time-varying WSN. Then, we provide a brief revision of the classical randomized symmetric gossip algorithm scheme [1], from which we motivate our work.

2.1. Graph theory

Any sequence of random communications between nodes can be modeled as a time-varying graph $\mathbf{G}(k) = (\mathbf{V}, \mathbf{E}(k))$, consisting of a set $\mathbf{V}$ of N nodes and a set $\mathbf{E}(k) \subset \mathbf{E}$ of links, where k denotes the current iteration of the consensus process and $\mathbf{E}$ denotes the set of all possible links. A directed link from node i to node j is denoted as e_{ij}, which indicates that there exists a directed information flow from node i to node j. Given a time-varying graph $\mathbf{G}(k)$, we can assign an $N \times N$ adjacency matrix $\mathbf{A}(k)$ where an entry is equal to 1 if $e_{ij} \in \mathbf{E}(k)$ and 0 otherwise. The set of neighbors of a node i is defined as $\mathbf{\Omega}_i = \{j \in \mathbf{V} : e_{ij} \in \mathbf{E}\}$. This set contains all the nodes that receive a packet when node i broadcast it. The cardinality of this set defines the degree of node i, that is, $d_i = |\mathbf{\Omega}_i|$. Note that the subset $\mathbf{\Omega}_i(k) \subset \mathbf{\Omega}_i$ may contain zero, one or several nodes in a particular iteration k.

2.2. Symmetric gossip

Given a network of N nodes and assuming that each node i has some initial measurement x_i, the average gossip algorithm allows every node to estimate the global average within certain accuracy. In this process, no central entity is available so that the average has to be obtained by only using local information and local communications. Various strategies have been proposed in the literature to solve the problem of obtaining the global average in a totally distributed fashion. The symmetric gossip algorithm presented in Boyd et al. [1] focuses on a low complexity implementation. However, this protocol, in order to ensure convergence, requires the following: i) a strongly connected undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ in which all the instantaneous subgraphs are included $\mathbf{G}(k) \subset \mathbf{G}$, ii) a symmetric $N \times N$ matrix $\mathbf{W}(k)$, at each iteration k, such that $[\mathbf{W}(k)]_{ij} = 0$ if $[\mathbf{A}(k)]_{ij} = 0$ and iii) $\mathbf{W}(k)\mathbf{1} = \mathbf{1}$, $1^T\mathbf{W}(k) = \mathbf{1}^T$. In this scheme, the non-zero entries of $\mathbf{W}(k)$ are generated by randomly activating, at each time instant k, an undirected link, which determines the nodes that pairwise exchange and mix information at that iteration. The information is mixed according to a fix real number $\epsilon \in (0, 1)$, which is called the step size of the process. This scheme leads to the nodes i and j producing a new state according to the equations:

$$x_i(k+1) = (1-\epsilon)x_i(k) + \epsilon x_j(k) \quad (1)$$

$$x_j(k+1) = \epsilon x_i(k) + (1-\epsilon)x_j(k) \quad (2)$$

The state of the rest of nodes remains unaltered.

2.3. Motivation of our approach

The symmetric gossip algorithm, defined by (1) and (2) is simple and provides, by only using local communications, probabilistic average consensus:

$$\lim_{k\to\infty} \mathbf{x}(k) = \mathbf{x}_{\text{avg}} \quad (3)$$

where $\mathbf{x}_{\text{avg}}$ is a vector that contains N entries equal to the average μ of the initial data of nodes, that is, $\mathbf{x}_{\text{avg}} = \mu\mathbf{1}$.

However, this approach generally requires many iterations to converge. Several alternative schemes of communications have been proposed in order to accelerate consensus. Some important examples are the works of [2] and [3]. In particular, the work in [2] proposes the use of asymmetric broadcast communications, which provides convergence to some common value, but it may significantly differ from the average value. Alternatively, the work in [3] makes use of geographic information to accelerate the gossip process, which requires establishing long routes, making it difficult to ensure symmetric communications and thus average convergence. In practice, ensuring symmetric communications is difficult due to the existence of wireless interferences and other environmental factors. Therefore, an alternative asymmetric algorithm to ensure convergence to average in both schemes [2][3] is necessary.

The main problem of using asymmetric communications is that the summation of the initial data is not preserved along the iterations. When a packet is asymmetrically sent, the receiving nodes update its state producing a deviation from the target state (the average in our case). The intuition behind our approach is to exploit this deviation from the average, generated in each asymmetric data exchange, to preserve the summation of the system and accelerate consensus.

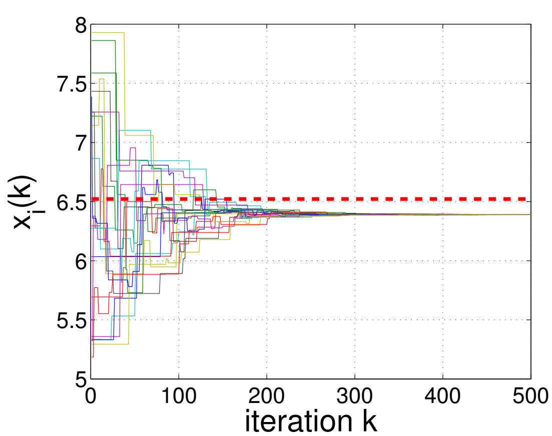


Fig. 1. Example of the convergence of $N = 20$ nodes with an asymmetric gossip algorithm. The process converges to a certain biased value that clearly differs from the initial average. This initial target value is represented by the red dashed line.

3. RESIDUAL GOSSIP

We propose an asymmetric gossip algorithm that presents fast convergence in both unicast and broadcast communication schemes. In this asymmetric scenario, all the instantaneous graphs $\mathbf{G}(k) \subset \mathbf{G}$ are directed. The weight matrix $\mathbf{W}(k)$, at each iteration k, is asymmetric and only the condition $\mathbf{W}(k)\mathbf{1} = \mathbf{1}$ is fulfilled. Finally, the data is mixed according to a fix real number $\epsilon \in (0, 1)$. Let us consider first the simplest asymmetric gossip scheme:

$$x_j(k+1) = \epsilon x_i(k) + (1-\epsilon)x_j(k),\ \forall j \in \mathbf{\Omega}_i(k) \quad (4)$$

The remaining nodes in the network, including the transmitter node i, do not update their state value, that is:

$$x_\ell(k+1) = x_\ell(k),\ \forall \ell \notin \mathbf{\Omega}_i(k) \quad (5)$$

This is the simplest asymmetric scheme for gossiping, which is studied in [5]. It ensures convergence to a common state value, but this state is not necessarily the average (see Fig. 1). In other words, there exists a random variable α such that $\forall i \in \mathbf{V}$ it is accomplished that $\lim_{k\to\infty} x_i(k) = \alpha$.

At each step of the asymmetric process, a residual component is ignored. Disregarding it is what causes the final deviation from the average. In particular, the following term is ignored:

$$r_j = x_j(k) - (\epsilon x_i(k) + (1-\epsilon)x_j(k)),\ \forall j \in \mathbf{\Omega}_i(k) \quad (6)$$

which coincides with the current variation of node j.

What we propose in our residual gossip algorithm is to include in the packet transmitted through each link e_{ij}, together with the current state of the transmitter node i, its accumulated residual r_i. As a consequence, the system evolves according to the following equation:

$$x_j(k+1) = \epsilon x_i(k) + (1-\epsilon)(x_j(k) + r_i(k)),\ \forall j \in \mathbf{\Omega}_i(k) \quad (7)$$

which implies that the residuals evolve as follows:

$$r_j(k+1) = r_j(k) + x_j(k) - x_j(k+1) + r_i(k),\ \forall j \in \mathbf{\Omega}_i(k) \quad (8)$$

and $r_i(k+1) = 0$.

This alternative asymmetric scheme defined by (7) and (8) preserves the summation of the initial data along iterations. Then, if this scheme is still able to converge to a common state, it is surely that this is the average of the initial data.

From expression (7), the weight matrix $\mathbf{W}(k)$ of the asymmetric gossip is:

$$[\mathbf{W}(k)]_{\ell j} = \begin{cases} 1-\epsilon & \text{if } j = \ell,\ \ell \in \mathbf{\Omega}_i(k) \\ \epsilon & \text{if } j \neq \ell,\ j = i,\ \ell \in \mathbf{\Omega}_i(k) \\ 1 & \text{if } j = \ell,\ \ell \notin \mathbf{\Omega}_i(k) \\ 0 & \text{otherwise} \end{cases}$$

where node i is chosen to send data in the current iteration k.

Equivalently, we define the residual vector $\mathbf{s}$ as follows:

$$[\mathbf{s}(k)]_j = \begin{cases} r_i(k) & \text{if } j \in \mathbf{\Omega}_i(k) \\ 0 & \text{otherwise} \end{cases}$$

which allows us to express (7) as follows:

$$\begin{aligned}
\mathbf{x}(1) &= \mathbf{W}(0)(\mathbf{x}(0) + \mathbf{s}(0)) \\
&= \mathbf{W}(0)\mathbf{x}(0) + \mathbf{W}(0)\mathbf{s}(0) \\
\mathbf{x}(2) &= \mathbf{W}(1)(\mathbf{x}(1) + \mathbf{s}(1)) \\
&= \mathbf{W}(1)\mathbf{W}(0)(\mathbf{x}(0) + \mathbf{s}(0)) + \mathbf{W}(1)\mathbf{s}(1) \\
&\vdots \\
\mathbf{x}(k+1) &= \mathbf{W}(k)(\mathbf{x}(k) + \mathbf{s}(k)) \\
&= \mathbf{W}(k)\mathbf{W}(k-1)\ldots\mathbf{W}(1)\mathbf{W}(0)\mathbf{x}(0) + \\
&\quad +\mathbf{W}(k)\mathbf{W}(k-1)\ldots\mathbf{W}(1)\mathbf{W}(0)\mathbf{s}(0) + \\
&\quad +\mathbf{W}(k)\mathbf{W}(k-1)\ldots\mathbf{W}(1)\mathbf{s}(1) + \\
&\quad \vdots \\
&\quad +\mathbf{W}(k)\mathbf{s}(k)
\end{aligned}$$

where $\mathbf{s}(0)$ is a vector with all of its entries equal to zero. This system evolution can be expressed as follows:

$$\mathbf{x}(k+1) = \prod_{i=0}^{k}(\mathbf{W}(i)\mathbf{x}(0)) + \sum_{j=0}^{k}\left(\prod_{\ell=j}^{k}(\mathbf{W}(\ell)\mathbf{s}(j))\right) \quad (9)$$

Proposition 1: The system defined by (9) reaches probabilistic average consensus $\lim_{k\to\infty} \mathbf{x}(k) = \mathbf{x}_{\text{avg}}$, provided that $\lim_{k\to\infty} r_i(k) = 0\ \forall i \in \mathbf{V}$.

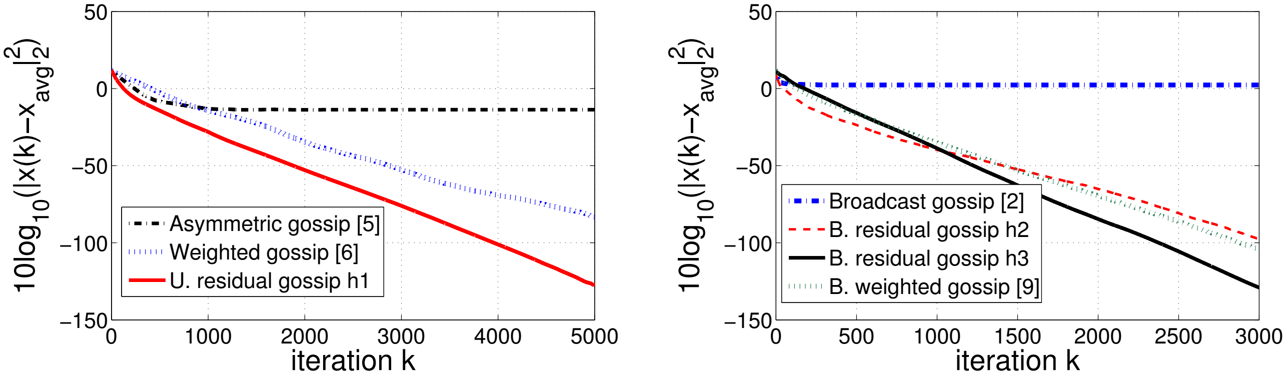
Proof. The first term of (9), given by $\prod_{i=0}^{k}(\mathbf{W}(i)\mathbf{x}(0))$, defines an asymmetric consensus as the one studied in [5]. Since every $\mathbf{W}(k)$ is row stochastic, the product of these matrices and the initial vector $\mathbf{x}(0)$ ensures probabilistic consensus to a scalar random variable α. The second term can be seen as the summation of several asymmetric consensus. Let us denote $k = C$ the first iteration in which $r_i(k) \leq \xi\ \forall i \in \mathbf{V}$, for an arbitrarily small value of ξ. When this consensus iteration $k = C$ is reached, we can continue the process Δ iterations more, that is, $\sum_{j=0}^{C+\Delta}\left(\prod_{\ell=j}^{C+\Delta}(\mathbf{W}(\ell)\mathbf{r}(j))\right)$. If Δ is large enough, all the products, started before the iteration $k = C$, are able to achieve probabilistic consensus to a scalar random variable α_i. Formally, our system provides:

$$\lim_{k\to\infty} \mathbf{x}(k) = \alpha\mathbf{1} + \sum_{i=0}^{C}\alpha_i\mathbf{1} + \sum_{i=C+1}^{C+\Delta}\alpha_i\mathbf{1}$$

where $\alpha_i \leq \xi$ if $i > C$.

Therefore, since we are ensuring that the summation is preserved along the iterations, our scheme provides probabilistic average consensus $\lim_{k\to\infty} x(k) = \alpha\mathbf{1} + \sum_{i=0}^{C}\alpha_i\mathbf{1} = \mathbf{x}_{\text{avg}}$.

□



(a) unicast communications

(b) broadcast communications

Fig. 2. Convergence of $N = 50$ nodes in several randomly deployed networks. (a) Comparison between our unicast residual gossip and the works of [5] and [6]. (b) Comparison between our broadcast residual gossip and the works of [2] and [9].

The process defined by (9) is suitable for both unicast and broadcast communications. The main difference between these two schemes is the number of nodes that receive the information. Our overall algorithm is described as follows: When a new iteration is started, a node i is randomly chosen. This node wakes up and sends a unicast or a broadcast packet with its own state value and the residual, which is divided by certain quantity $1 \leq \beta \leq d_i$ that depends on the number of receivers. The packet is received by the destination node or the whole neighborhood. Once it is received, the corresponding nodes update their information according to (7). Therefore at each iteration k, we have the following:

- Node i sends a packet containing $x_i(k)$ and $\frac{r_i(k)}{\beta}$ and sets $r_i(k) = 0$. How to choose β is explained later.
- The packet is successfully received by the destination node in the unicast scheme, or by every node $j \in \mathbf{\Omega}_i(k)$ in the broadcast scheme.
- The nodes receiving the packet from node i update its own state according to (7), which generates an update of its own residual by using (8). The state of the rest of the nodes remains unaltered.

This algorithm reaches probabilistic average consensus if $\lim_{k\to\infty} r_i(k) = 0 \ \forall i \in \mathbf{V}$, as shown in **Proposition 1**. Notice that intuitively both the nodes reaching consensus and the residuals approaching to zero, are closely related. In fact, if the residual sent from the transmitter to the receiver makes their difference among states to remain equal or to become smaller, the difference among states and the resulting residual are both decreased. This reduction is explained by the manner in which the state of the nodes and the residuals are updated. In the following subsections, we propose several heuristics based on the idea of the residuals that are able to accelerate convergence under both unicast and broadcast communications.

3.1. Unicast scheme of communications

When a unicast scheme of communications is considered, only one node receives the information from the origin node i. We want to minimize the value of the residual along the iterations. This intuitively leads to ensure $\lim_{k\to\infty} r_i(k) = 0 \ \forall i \in \mathbf{V}$. There are two ways to operate in this scenario: i) the origin node i randomly chooses a neighbor j from $\mathbf{\Omega}_i$, which updates its state if r_i and r_j are of opposite sign, or node j directly updates its residual with r_i otherwise or ii) the origin node decides to which node is best to send the information according to some local criteria. We focus on the second methodology, which leads us to the following heuristic:

Min-max approach (h1): This heuristic chooses the neighbor with minimum state value if the current residual of the origin node i is positive and it chooses the neighbor with maximum state value if the current residual of the origin node i is negative. This local information can be estimated during the process, updating the minimum and the maximum when the neighbors send their own information.

3.2. Broadcast scheme of communications

When a broadcast scheme of communications is considered, every node $j \in \mathbf{\Omega}_i(k)$ receives the information from the origin node i. The proposed heuristics in this scheme are:

Equitable approach (h2): This first method is based on equally distributing the residuals. For this purpose, the origin node i broadcastly sends $(x_i(k), \frac{r_i(k)}{d_i})$ to every node $j \in \mathbf{\Omega}_i(k)$, that computes its new state (7) and residual (8).

Large-small approach (h3): This heuristic updates the neighbors with smaller state value than node i if its current residual is positive or it updates the state of the neighbors with larger state value otherwise. The rest of neighboring nodes only update their residual.

3.3. Toy example

Let us define a network example composed by $N = 5$ nodes, where, after k iterations, the state of the nodes is $\mathbf{x}(k) = [7\ 3\ 2\ 8\ 5]^T$ and the residuals $\mathbf{r}(k) = [-3\ -1\ 2\ 2\ 0]^T$. Now imagine that node $i = 1$ has been chosen to broadcast its value to its neighbors (2, 3 and 4). In this case, if we use $\epsilon = 1/2$, then the entries of the matrix $\mathbf{W}$ are $W_{21} = W_{31} = W_{41} = 1/2$, $W_{11} = W_{55} = 1$, $W_{22} = W_{33} = W_{44} = 1/2$ and the rest are zero. Finally, if we consider the equitable approach (h3), the residual is spread as $r_i(k)/d_i = -3/3$, thus vector $\mathbf{s}$ is $\mathbf{s} = [0\ -1\ -1\ -1\ 0]^T$.

The resulting vectors are $\mathbf{x}(k+1) = [7\ \frac{9}{2}\ 4\ 7\ 5]^T$ and $\mathbf{r}(k+1) = [0\ -\frac{7}{2}\ -1\ 2\ 0]^T$. Note that the summation of the process $\sum_{i=1}^{5}(x_i(k) + r_i(k)) = 25$ is preserved.

4. NUMERICAL RESULTS

We model a WSN as a randomly deployed network of $N = 50$ nodes inside a $2D$ unit square area. The information is mixed as described in (7) where the instantaneous topology determines which data is mixed. We average our results over 100 different random topologies.

Fig. 2 shows that our residual gossip algorithm converges faster than existing methods, where we evaluate the convergence in terms of the deviation from the average of the initial data. In the unicast scheme of communications, the heuristic h1 is compared with [5] and [6]. In the broadcast scheme of communications, the two heuristics h2 and h3 are compared with [2] and [9]. Note that the broadcast scheme takes less iterations to converge with a similar error.

Fig. 3 shows the evolution of the residuals associated to our heuristics. The norm of the vector $\mathbf{r}(k)$ is evaluated along the iterations. Since the vector $\mathbf{r}(0)$ has all of its entries equal to zero, the norm of this vector grows along the first iterations, until a maximum is reached and from which it starts to rapidly decrease. The method that provides the fastest convergence is the one that produces the largest residuals at the beginning. As expected, all the residuals vanish as the iteration number k increases. To converge even faster, we could add or subtract some suitable value δ_i to each node state and initialize the residuals with it, i.e. $x_i(0) = x_i(0) - \delta_i$ and $r_i(0) = \delta_i$.

5. CONCLUSIONS AND FUTURE WORK

In this work, we propose a new protocol for performing asymmetric gossiping that lead to probabilistic average consensus with fast convergence speed. We compare our convergence results with several existing approaches, showing the superior performance of our protocol in both unicast and broadcast schemes of communications. For future work, the exact conditions for the residual to vanish with the iteration number, ensuring probabilistic average consensus have to be investigated.

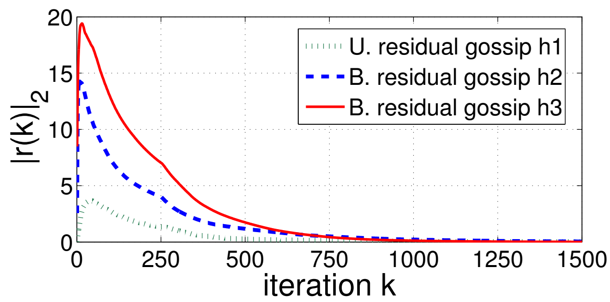


Fig. 3. Evolution of the residuals with the iteration number k.

REFERENCES

[1] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. "Randomized gossip algorithms." *IEEE Transactions on Information Theory*, 52(6):2508-2530, June 2006.

[2] T. Aysal, M. Yildiz, A. Sarwate, and A. Scaglione. "Broadcast gossip algorithms for consensus." *IEEE Transactions on Signal Processing*, 57(7):2748-2761, July 2009.

[3] Dimakis, A.D.G.; Sarwate, A.D.; Wainwright, M.J.; , "Geographic Gossip: Efficient Averaging for Sensor Networks," *IEEE Transactions on Signal Processing*, vol.56, no.3, pp.1205-1216, March 2008.

[4] Antunes, D.; Silvestre, D.; Silvestre, C.; , "Average consensus and gossip algorithms in networks with stochastic asymmetric communications," *IEEE Conference on Decision and Control and European Control Conference*. CDC 2011, , vol., no., pp.2088-2093, 12-15 December 2011.

[5] Fagnani, F.; Zampieri, S.; , "Randomized consensus algorithms over large scale networks," *IEEE Journal on Selected Areas in Communications*, vol.26, no.4, pp.634-649, May 2008.

[6] Bnzit, F.; Blondel, V.; Thiran, P.; Tsitsiklis, J.; Vetterli, M.; , "Weighted Gossip: Distributed Averaging using non-doubly stochastic matrices," *IEEE International Symposium on Information Theory Proceedings*. ISIT 2010, , vol., no., pp.1753-1757, 13-18 June 2010.

[7] Wei Ren; Beard, R.W.; Atkins, E.M., "A survey of consensus problems in multi-agent coordination," *Proceedings of the 2005 American Control Conference*, 2005, vol., no., pp.1859,1864 vol. 3, 8-10 June 2005.

[8] Olfati-Saber, R.; Fax, J.A.; Murray, R.M., "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE*, vol.95, no.1, pp.215-233, January 2007.

[9] Iutzeler, Franck; Ciblat, Philippe; Hachem, Walid; Jakubowicz, Jeremie; , "New broadcast based distributed averaging algorithm over wireless sensor networks," *IEEE International Conference on Acoustics, Speech and Signal Processing*. ICASSP 2012 , vol., no., pp.3117-3120, 25-30 March 2012.

[10] Zhang, H., Moura, J. M., and Krogh, B. 2009. "Dynamic field estimation using wireless sensor networks: tradeoffs between estimation error and communication cost." *IEEE Transactions on Signal Processing*, 57, 6 , 2383-2395, June 2009.