Greedy Gossip Algorithm with Synchronous Communication for Wireless Sensor Networks

Zhang, Jie; Hendriks, R. C.; Heusdens, R.

Publication date
2016

Document Version
Accepted author manuscript

Published in
Proceedings of the 37th WIC Symposium on Information Theory in the Benelux and The 6th Joint WIC/IEEE Symposium on Information Theory and Signal Processing in the Benelux

Greedy Gossip Algorithm with Synchronous Communication for Wireless Sensor Networks

Jie Zhang, Richard C. Hendriks and Richard Heusdens
Signal and Information Processing Lab.,
Dept. of Microelectronics, Delft University of Technology,
2628 CD Delft, The Netherlands
{j.zhang-7, r.c.hendriks, r.heusdens}@tudelft.nl

Abstract

Randomized gossip (RG) based distributed averaging has been popular for wireless sensor networks (WSNs) in multiple areas. With RG, randomly two adjacent nodes are selected to communicate and exchange information iteratively until consensus is reached. One way to improve the convergence speed of RG is to use greedy gossip with eavesdropping (GGE). Instead of randomly selecting two nodes, GGE selects the two nodes based on the maximum difference between nodes in each iteration. To further increase the convergence speed in terms of transmissions, we present in this paper a synchronous version of the GGE algorithm, called greedy gossip with synchronous communication (GGwSC). The presented algorithm allows multiple node pairs to exchange their values synchronously. Because of the selection criterion of the maximum difference between the values at the nodes, there is at least one node pair with different information, such that the relative error must be reduced after each iteration. The convergence rate in terms of the number of transmissions is demonstrated to be improved compared to GGE. Experimental results validate that the proposed GGwSC is quite effective for the random geometric graph (RGG) as well as for several other special network topologies.

1 Introduction

Distributed signal processing in wireless sensor networks (WSNs) has many operational advantages. For instance, there is no need to have a fusion centre (or host) for facilitating computations, communication and time-synchronization. Positions of the network nodes are not necessarily known *a priori*, and the network topology might change as nodes join or disappear. For the design of fault-tolerant computation and information exchange algorithms over such WSNs, decentralized randomized gossip (RG) based averaging consensus is attractive, because it does not require any special routing, there is no bottleneck or single point of failure, and it is robust to unreliable and changing wireless network conditions. Moreover, the decentralized RG puts no constraints on the network topology and requires no information about the actual topology.

Since the original RG algorithm was proposed in [1], many derivatives were proposed to improve its convergence rate, and it has been employed into various applications (see e.g., [2] and references therein). Dimakies introduced a geographic gossip [3], which enables information exchange over multiple hops with the assumption that nodes have knowledge of their geographic locations, such that it is a good alternative for the grid network topology. In [4], a synchronous communication process was considered and improvements were made to the synchronous RG of [1] in a speech enhancement context. They allowed multiple node pairs to exchange their current values per iteration synchronously. Other improvements to increase the convergence speed are to use clique-based RG (CbRG) and cluster-based RG (see e.g., [5] and [6]), where cliques or clusters

are used to compress the original graph. Deniz *et al* presented a greedy gossip with eavesdropping (GGE) to accelerate the convergence [7]. Instead of randomly choosing two nodes, they chose the two nodes to communicate that have the maximum difference between values per iteration. Another more competitive broadcasting based algorithm was proposed in [8], although it cannot guarantee to reach the actual consensus surely.

To further increase the convergence speed in terms of transmissions, we present in this paper a synchronous version of the GGE algorithm, called greedy gossip with synchronous communication (GGwSC). Each time slot is divided into two time scales, one is the time used for node pairs selection, and the other is for the *gossip exchange* between every node pair. The simultaneous communicating node pairs are chosen recursively. Each time, one node selects the node from its neighbors that has the maximum difference. Then, the additional communicating node pairs are chosen recursively by excluding the node pairs that are already formed. Finally, the chosen node pairs communicate synchronously. Thus, unlike the synchronous gossip in [1] or [4], which performs updates completely at random, the GGwSC, like GGE, makes use of the greedy neighbor selection procedure. Whereas unlike GGE, we also permit multiple node pairs to communicate so as to accelerate the convergence rate. Experiments have demonstrated the effectiveness of the proposed method. The convergence rate in terms of the number of transmissions for random geographic graphs (RGGs) is accelerated compared to the GGE algorithm. Additionally, we also test the improvement on the convergence rate of the proposed method under different conditions in this paper, e.g., different initializations for the nodes and different network topologies.

2 Fundamentals of GGE

To guide the reader, we first give a brief overview of the GGE algorithm presented in [7]. We consider a network of N nodes and represent network connectivity as a graph, $G = (V, E)$, with vertices $V = \{1, 2, ..., N\}$ and edge set $E \subset V \times V$ such that $(i, j) \in E$ if and only if nodes i and j directly communicate. We assume that communication relationships are symmetric and that the graph is connected. Let $\mathcal{N}_i = \{j : (i, j) \in E\}$ denote the set of neighbors of node i (excluding i). Each node in the network has an initial value y_i, and the goal is to use only local information exchanges to arrive at a state where every node knows the average $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$. Each node is initialized with $x_i(0) = y_i$.

At the kth iteration of GGE [7], an activated node s_k is chosen uniformly at random. This can be accomplished using the asynchronous time model, where each node "ticks" according to a Poisson clock with rate 1. Then, s_k identifies a neighboring node t_k satisfying

$$t_k \in \arg\max_{t \in \mathcal{N}_j} \Big\{ \frac{1}{2}(x_{s_k}(k-1) - x_t(k-1))^2 \Big\}, \tag{1}$$

in other words, s_k identifies a neighbor that currently has the most different value from itself. This choice is possible because each node i maintains not only its own local variable, $x_i(k-1)$, but also a copy of the current values at its direct neighbors, $x_j(k-1)$, for $j \in \mathcal{N}_i$, because of eavesdropping with wireless communications. When s_k has multiple neighbors whose values are equally (and maximally) different from s_k's, it chooses one of these neighbors at random. Then the update is performed by enforcing the average $\frac{1}{2}(x_{s_k}(k-1) - x_{t_k}(k-1))$ to s_k and t_k, while all other nodes $i \notin \{s_k, t_k\}$ hold their values at $x_i(k) = x_i(k-1)$. Finally, the two nodes, s_k and t_k, broadcast these new values so that their neighbors have up-to-date information. If the values x_i on all sensors are stacked as a vector, i.e., $\boldsymbol{x}(k) = [x_1(k), x_2(k), ..., x_N(k)]^T$, we can formulate the above update as

$$\boldsymbol{x}(k) = \boldsymbol{U}_{GGE}(k)\boldsymbol{x}(k-1), \tag{2}$$

where $\boldsymbol{U}_{GGE}(k)$ is an $n \times n$ dimensional update matrix, which is dependent across time. For two communicating nodes x_{s_k} and x_{t_k} at iteration k, the update matrix is

$$\boldsymbol{U}_{GGE}(k) = \boldsymbol{I} - \frac{1}{2}(\boldsymbol{e}_{s_k} - \boldsymbol{e}_{t_k})(\boldsymbol{e}_{s_k} - \boldsymbol{e}_{t_k})^T, \tag{3}$$

where $\boldsymbol{e}_i = [0, ..., 1, 0, ..., 0]^T$ is an $N-$ dimensional vector with the ith entry equal to 1. Note that similar to the standard RG, the update matrix is doubly stochastic, which implies $\boldsymbol{U}_{GGE}\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^T\boldsymbol{U}_{GGE} = \mathbf{1}^T$ with $\mathbf{1}$ denoting a vector of all ones.

Given the initial vector of a network $\boldsymbol{x}(0) = [x_1(0), x_2(0), ..., x_N(0)]^T$, the theoretical consensus will be $\tilde{x}_{ave} = \mathbf{1}\tilde{\boldsymbol{x}}(\mathbf{0})/N$. To measure the convergence rate, we use the relative convergence error defined as

$$RE = \frac{\|\tilde{\boldsymbol{x}}(k) - \tilde{x}_{ave}\mathbf{1}\|}{\|\boldsymbol{x}(0) - \tilde{x}_{ave}\mathbf{1}\|}, \tag{4}$$

such that the iteration can be quitted when $RE \leq \varepsilon$ (or after a fixed amount of iterations).

3 GGwSC

In this section, we will present the proposed GGwSC algorithm based on GGE. As mentioned above, in GGE, a node selects a neighboring node whose state value is most different from its own value. This strategy can indeed accelerate the convergence at the cost of additional communication bandwidth compared to the original gossip algorithm [1], because it has to send (broadbast) the new values (eavesdrop) to all its neighbors. In spite of this, it still has a relatively slow convergence because only two nodes are allowed to exchange their state values at each iteration. In [4], a synchronous randomized gossip (SRG) was proposed for distributed delay and sum beamforming (DDSB) based speech enhancement in WSNs, where each node is permitted to communicate with one of its neighbors randomly at each iteration, such that the state values of multiple nodes are updated after each iteration. Given sufficient communication bandwidth, we combine the idea of GGE and SRG to further accelerate the convergence. Hence for the GGwSC, multiple node pairs can communicate at each iteration. These active node pairs are constrained to be disjoint, and the communicating node pairs are chosen according to **arg max** distance vectors.

This newly proposed GGWSC algorithm can generally be described as in **Algorithm 1**. For the practical realization, there are several points worthy to be noted:

- ◇ Given N (even) nodes, the desired case is that $N/2$ node pairs are chosen synchronously by the *SelectNodePair* function at each iteration. This would be most efficient. However, this will not always happen. For example, at kth iteration, when the node s_k is randomly activated, but all of its neighbors are selected already (i.e., $\mathcal{N}_{s_k} = \mathcal{O}$), s_k has a bye (i.e., $x_{s_k} = x_{s_{k-1}}$) and needs to wait for the next iteration $k+1$.

- ◇ For the kth iteration, the update matrix $\boldsymbol{U}_{GGwSC}(k)$ is a manifold stochastic process approximately, that is, $\boldsymbol{U}_{GGwSC}(k) = \prod_{\{s_k,t_k\}\in V} \boldsymbol{U}_{GGE}^{s_k,t_k}(k)$.

- ◇ Note that for a communicating node pair, two transmissions are required during an iteration, e.g., s_k computes the average, such that s_k broadcasts it to its neighbors, and t_k also needs to broadcast the received average from s_k to its neighbors.

Algorithm 1: *GGwSC*

Input: $\boldsymbol{x}(0) = [x_1(0), x_2(0), ..., x_N(0)]^T$, $G = (V, E)$

while $RE > \varepsilon$ **do**

  function $SelectNodePair(G)$

    $s_k = N \times rand$;

    $t_k \in \arg\max_{t \in \mathcal{N}_{s_k}} \left\{ \frac{1}{2}(x_{s_k}(k-1) - x_t(k-1))^2 \right\}$;

    updating topology by excluding $\{s_k, t_k\}$ to $G' = (V', E')$;

    if $(V' \neq \mathcal{O})$ $SelectNodePair(G')$;

    else break;

  end function

  $\boldsymbol{U}_{GGE}^{s_k,t_k}(k) = \boldsymbol{I} - (e_{s_k} - e_{t_k})(e_{s_k} - e_{t_k})^T/2$;

  $\boldsymbol{U}_{GGwSC}(k) = \prod_{\{s_k,t_k\} \in V} \boldsymbol{U}_{GGE}^{s_k,t_k}(k)$;

  $\boldsymbol{x}(k) = \boldsymbol{U}_{GGwSC}(k)\boldsymbol{x}(k-1)$;

end

return $\boldsymbol{x}(k)$

3.1 Convergence Rate: GGwSC versus GGE

In the following, we investigate the convergence rate in terms of the underlying communication topology. The convergence rate for gossip algorithms [1] is typically defined in terms of the $\varepsilon-$avergaing time

$$T_{ave}(\varepsilon) = \sup_{\boldsymbol{x}(0) \neq \boldsymbol{0}} \inf \Big\{ k : \Pr\Big(\frac{\|\tilde{\boldsymbol{x}}(k) - \tilde{x}_{ave}\boldsymbol{1}\|}{\|\boldsymbol{x}(0) - \tilde{x}_{ave}\boldsymbol{1}\|} > \varepsilon \Big) \leq \varepsilon \Big\}. \tag{5}$$

The averaging time $T_{ave}(\varepsilon, Pr)$ is bounded by the second largest eigenvalue of the expected value of the update matrix $E[\boldsymbol{U}_{GGwSC}]$, that is [1]

$$\frac{0.5 \log \varepsilon^{-1}}{\log \lambda_2(E[\boldsymbol{U}_{GGwSC}])^{-1}} \leq T_{ave}(\varepsilon, Pr) \leq \frac{3 \log \varepsilon^{-1}}{\log \lambda_2(E[\boldsymbol{U}_{GGwSC}])^{-1}}. \tag{6}$$

Although this bound is suitable for the GGwSC as well, it is hard to relate it as a homogeneous Markov chain, and $T_{ave}(\varepsilon, Pr)$ is difficult to calculate as a function of $\lambda_2(E[\boldsymbol{U}_{GGwSC}])$, because $E[\boldsymbol{U}_{GGwSC}]$ depends on the network topology. Therefore, we use here an alternative bound to investigate the convergence rate, which is based on results from [7]. Given a graph $G = (V, E)$, we will have

$$E[\|\tilde{\boldsymbol{x}}(k) - \tilde{X}_{ave}\boldsymbol{1}\|^2] \leq A(G)^k \|\boldsymbol{x}(0) - \tilde{X}_{ave}\boldsymbol{1}\|^2, \tag{7}$$

where $A(G)$ is the graph-dependent constant defined as

$$A(G) = \max_{\boldsymbol{x} \neq \tilde{x}_{ave}\boldsymbol{1}} \frac{1}{N} \sum_{s=1}^{N} \Big(1 - \frac{\|\boldsymbol{g}_s(k)\|^2}{4\|\tilde{\boldsymbol{x}} - \tilde{x}_{ave}\boldsymbol{1}\|^2} \Big), \tag{8}$$

where $\boldsymbol{g}(k)$ is the subgradient function defined in [7]. Indeed, $A(G)$ is equivalent to $\lambda_2(E[\boldsymbol{U}_{GGWSC}])$ functionally. Obviously, the smaller of $A(G)$, the faster of the convergence rate. For the kth iteration of GGwSC, there is at least one node pair (s_k, t_k) communicating synchronously, such that $\boldsymbol{g}(k)$ has more than two elements unequal

to 0. Yet for the GGE algorithm, one node pair (s_k, t_k) is allowed to communicate per iteration, such that there are only two elements of the subgradient function unequal to 0. Therefore, we have the relationship between the subgradient functions, as $\|\boldsymbol{g}_{GGwSC}(k)\|^2 \geq \|\boldsymbol{g}_{GGE}(k)\|^2$, which leads to

$$A_{GGwSC}(G)^k \leq A_{GGE}(G)^k, \tag{9}$$

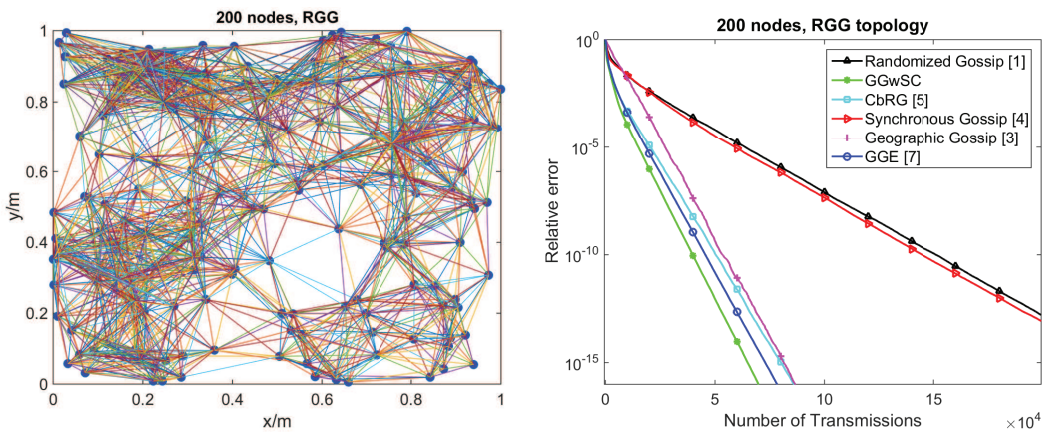
with equality if and only if only one node pair gossips per iteration. Consequently, we have demonstrated theoretically that GGwSC converges faster than GGE.

4 Performance Analysis

In this section, we present simulations to compare the GGwSC with several state-of-the-art methods, including Boyd's original RG [1], GGE [7], synchronous gossip [4], CbRG [5] and geographic gossip [3], by observing the convergence rate in terms of transmissions. We also investigate how this is effected by the network topology.

4.1 Random Geometric Graph (RGG)

Firstly, in order to observe the general performance of convergence, we place 200 nodes randomly in a (1×1) m enclosure. A Gaussian distribution $\mathcal{N}(0, 1)$, is used to initialize the values of $\boldsymbol{x}(0)$ on each sensor. The maximum number of transmissions is fixed to 20000, and the results are averaged over 100 realizations for the RGG. The transmission radius is set to be $\sqrt{\log N/N}$, which determines the RGG topology.



(a) RGG (b) Convergence

Figure 1: Convergence of relative error of the state-of-the-art methods for the RGG topology with 200 nodes.

Fig. 1(a) shows a typical RGG with 200 nodes, and Fig. 1(b) shows the corresponding convergence behaviours. We can see that our method achieves the fastest convergence rate, and randomized gossip and synchronous gossip are slowest.

4.2 Initialization

Secondly, we examine performance for four different initial conditions, $\boldsymbol{x}(0)$, which are consistent to those in [7], in order to explore the impact of the initial values on the

convergence behaviour. The first two of these cases are a Gaussian bumps field, and a linearly-varying field. For these two cases, the initial value $\boldsymbol{x}(0)$ is determined by sampling these fields at the locations of the nodes. The remaining two initializations consist of the "spike" signal, constructed by setting the value of one random node to 1 and all other node values to 0, and a random initialization where each value is i.i.d. drawn from a Gaussian distribution $\mathcal{N}(0,1)$ of zero mean and unit variance. The first three of these signals were also used to examine the performance of geographic gossip in [3].

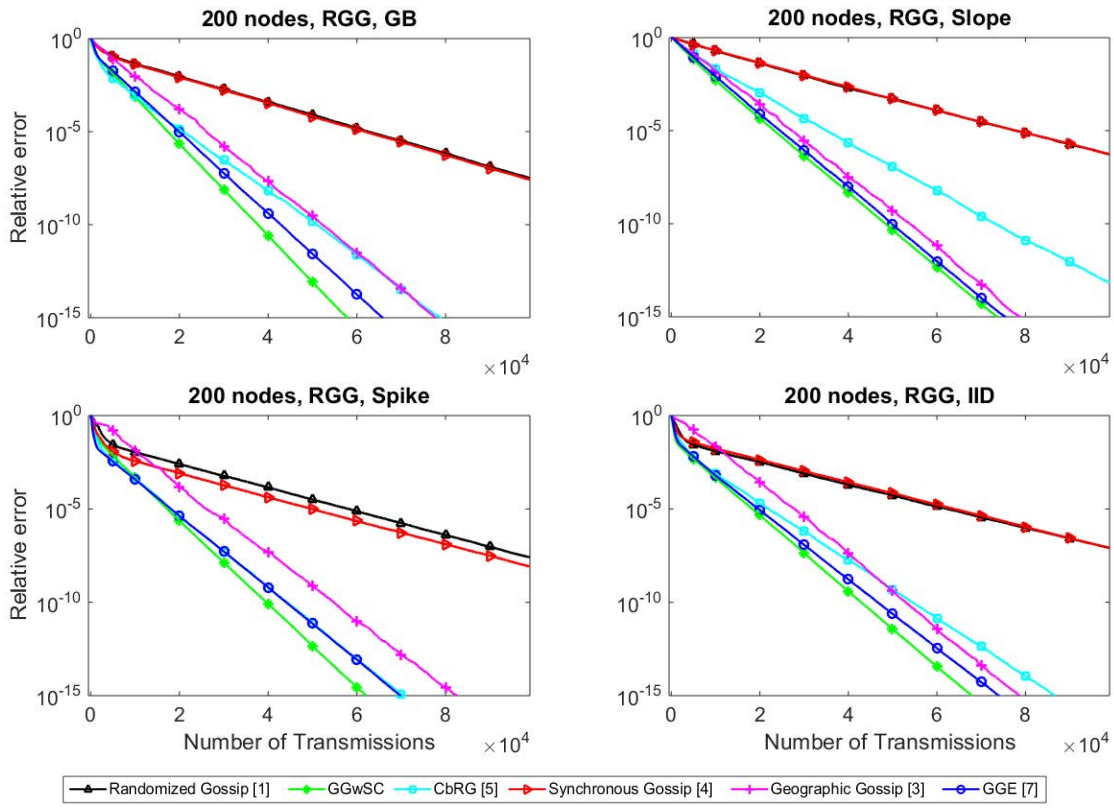


Figure 2: Comparison of the performance of the state-of-the-art methods with four different initializations of $\boldsymbol{x}(0)$.

Fig. 2 shows that GGwSC converges to the average at a faster rate asymptotically than the other state-of-the-art methods for all initial conditions. Out of these candidate initializations, the linearly-varying field is the worst case, because it improves the convergence rate least compared to GGE. This is not surprising since the convergence analysis in Section 3.1 suggests that constant differences between neighbors cause both GGwSC and GGE to provide minimal gain.

4.3 Special topologies

Finally, we investigate the influence of the network topologies on the convergence rates. We test three special kinds of topologies, including complete connected, grid, and a star topology. Note that for the grid network topology, the number of nodes must be a square. Some results are shown in Fig. 3 versus the number of transmissions. To this end, we can conclude:

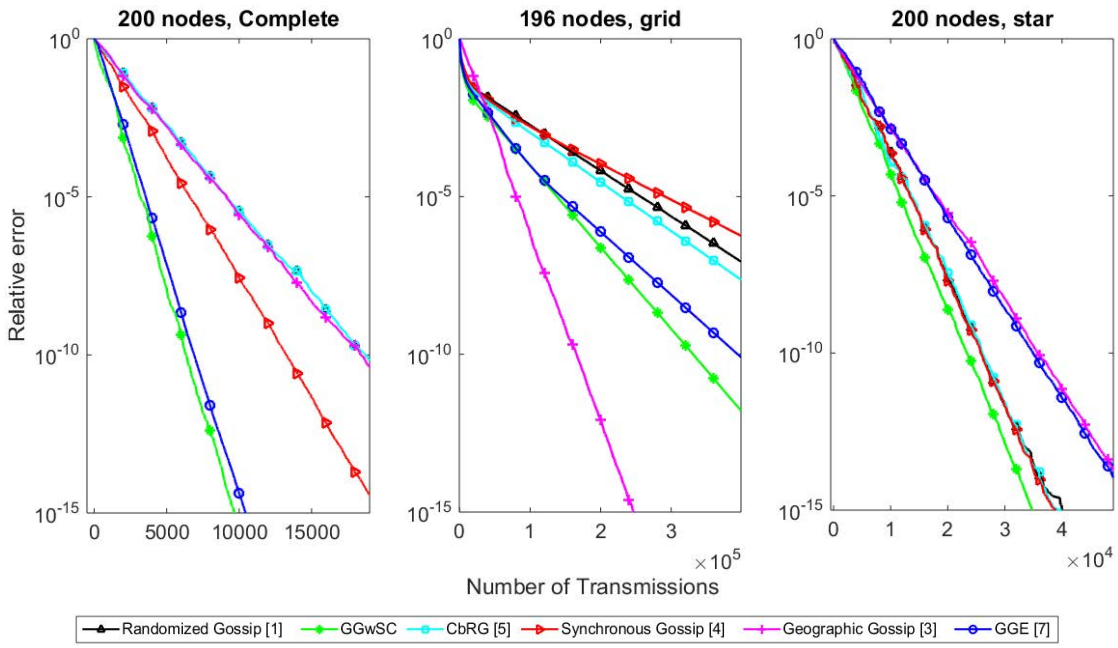


Figure 3: Comparisons of the performance of the state-of-the-art methods for three special network topologies (left: grid; middle: grid, where the number of nodes must be a square, e.g., 196; right: star).

- ◇ GGwSC is the most effective gossiping strategy, and it has the fastest convergence rate generally, except for the grid topology. For these grid-structured networks, geographic gossip has the best performance, because it is specified to thes kind of networks.
- ◇ Although both GGwSC and GGE perform gossiping according to the difference between neighboring nodes, through the synchronous communication strategy the former guarantees that at least one node pair has a value difference per iteration except in the case when the average is reached. That is why GGwSC is faster than GGE in terms of transmissions.

Accordingly, in general the proposed GGwSC algorithm obtains the fastest rate of convergence.

5 Conclusions

In this paper, we proposed a greedy gossip with synchronous communication (GGwSC) as an extension of the GGE algorithm [7] for averaging consensus. The convergence rate of GGwSC was analyzed theoretically as being faster than GGE. The experimental results demonstrated the effectiveness of the proposed method. Additionally, we also tested the performance on the convergence rate of our method under several conditions, e.g., different initializations for the nodes and different network topologies. In general, the proposed GGwSC algorithm obtained the fastest rate of convergence.

References

[1] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[2] A. G. Dimakis, S. Kar, J. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[3] A. G. Dimakis, A. D. Sarwate, and M. J. Wainwright, "Geographic gossip: efficient aggregation for sensor networks," in *ACM Int. Conf. Inform. Process. in Sensor Networks (IPSN)*, 2006, pp. 69–76.

[4] Y. Zeng and R. C. Hendriks, "Distributed delay and sum beamformer for speech enhancement via randomized gossip," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 22, no. 1, pp. 260–273, 2014.

[5] Y. Zeng, R. C. Hendriks, and R. Heusdens, "Clique-based distributed beamforming for speech enhancement in wireless sensor networks," in *EURASIP Europ. Signal Process. Conf. (EUSIPCO)*, 2013, pp. 1–5.

[6] W. Li and H. Dai, "Cluster-based distributed consensus," *IEEE Transactions on Wire. Communicat.*, vol. 8, no. 1, pp. 28–31, 2009.

[7] D. Üstebay, B. N. Oreshkin, M. J. Coates, and M. G. Rabbat, "Greedy gossip with eavesdropping," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3765–3776, 2010.

[8] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2748–2761, 2009.