Distributed estimation of the maximum value over a Wireless Sensor Network

Franck Iutzeler*, Jérémie Jakubowicz*, Walid Hachem†and Philippe Ciblat*
* Institut Télécom ; Télécom ParisTech ; CNRS LTCI
Email: {iutzeler,jakubowi,ciblat }@telecom-paristech.fr
† CNRS LTCI ; Télécom ParisTech
Email: hachem@telecom-paristech.fr

***Abstract*—This paper focuses on estimating the maximum of the initial measures in a Wireless Sensor Network. Two different algorithms are studied : the RANDOM GOSSIP , relying on pairwise exchanges between the nodes, and the BROADCAST in which each sensor sends its value to all its neighbors; both are asynchronous and distributed. We prove the convergence of these algorithms and provide tight bounds for their convergence speed.**

I. INTRODUCTION

Distributed estimation algorithms over Wireless Sensors Networks (WSN) have been widely studied since the seminal work of Tsitsiklis [1]. The goal of these algorithms is to make the network reach a consensus over the value of interest by means of local communications between the sensors; in particular, a lot of results have been shown in the past few years concerning the problem of averaging (that is to say make consensus over the mean of the initial values) [2], [3], [4], [5], [6].

However, various applications such as stock management or distributed computing could need the maximum value of the network (obviously, the estimation of the maximum value is equivalent to the estimation of any extrema over any function of the network measures such as the minimum value or the closest to a constant, ...). For example, if all the sensors have i) a task for which they want the collaboration of the network and ii) a priority coefficient, by making consensus over the maximal coefficient and the associated task or sensor ID, the network can process the most urgent task in priority.

Beyond going further, a clarification on the word *broadcast* is needed. In the rest of the paper, *broadcast* will refer to the fact of sending information to all reachable sensors as in [5] and not to the goal of information propagation like in the computer scientific literature [7]. An important point here is that the sensors do not know if they have the maximal value or not, because if it was the case a flooding algorithm would be more suitable (see [7], [8] for some results about the speed of convergence of flooding algorithms).

To share the maximum value over the entire network, the standard Random Pairwise Gossip approach (analyzed in [2], [3] in the case of averaging) can be used, but it does not take benefit of the broadcast nature of the wireless channel. Actually, broadcasting causes a major issue for averaging algorithms because the sum of measures is not conserved so the broadcast-based algorithms do not generally converge the statistical average [5]. This is obviously not a problem while estimating the maximum value. Hence, broadcasting information can be a good way to speed up convergence for the distributed estimation of the maximum in a WSN.

The paper is organized as follows : in Section II we will present our model and explicit the studied algorithms. Section III will be dedicated to the study of the convergence and in Section IV we will give speed convergence bounds for the consensus over the maximum value which are the main contributions of the paper. Finally, Section V will illustrate our results with some simulations and Section VI will be devoted to concluding remarks.

II. PROPOSED ALGORITHMS

A. Model

We consider a network of N sensors modeled by an unweighted undirected graph $\mathcal{G} = (V, E)$ where V is the set of vertices/sensors ($|V| = N$) and E is the set of edges/perfect links between the sensors. Each sensor i can exchange data with its neighborhood $\mathcal{N}_i = \{j \in V | (i, j) \in E\}$ and we define $d_i = |\mathcal{N}_i|$, the degree of node i and $d_{max} = \max_{i\in V} \ d_i$, the maximal degree across the network. We will denote by $\mathbf{A}$ the adjacency matrix of the graph such that $(\mathbf{A})_{ij} = 1$ if and only if $i \in \mathcal{N}_j$ and 0 elsewhere; as the considered graph is undirected, the adjacency matrix is obviously symmetric. We also introduce $\mathbf{D}$ the diagonal matrix of the degrees and $\mathbf{L} = \mathbf{D} - \mathbf{A}$ the Laplacian matrix of graph [9].

For practical reasons, we will suppose that $\mathcal{G}$ is connected and that each sensor is equipped with an independent Poisson clock of common parameter λ for its activation. This setup is equivalent to a global clock of parameter $N\lambda$ and uniform selection of the awaking sensor. We will note t the instant of the t-th ring of the global clock. We will also suppose that the communication duration is small with respect to the time between two clock ticks so there are no collisions between messages.

Each sensor i has an initial value $x_i(0)$; we introduce $x_{max} = \max_{i\in V} \ x_i(0)$ and $\mathbf{x}(t) = [x_1(t) \ldots x_N(t)]^T$ where $x_i(t)$ is the estimate of the i-th sensor at global time t. It is therefore obvious that we wish $\mathbf{x}(t)$ to converge to $x_{max}\mathbf{1}$ (with $\mathbf{1}$ being the size N vector of ones) as t goes to infinity.

Hence, the goal of the studied algorithms will be to achieve **max-consensus** in **finite time** τ, *i.e.* $\forall \mathbf{x}(0) \in \mathbb{R}^N, \exists \tau < \infty$:

$$\forall t > \tau,\ \mathbf{x}(t) = x_{max}\mathbf{1} \tag{1}$$

B. RANDOM GOSSIP *and* BROADCAST *algorithms*

Let us now present our two algorithms of interest:

- The RANDOM GOSSIP comes from the classical algorithm for average estimation over Sensor Networks [3]. When a sensor's clock ticks, it chooses uniformly another sensor among its neighbors and they update their estimate by taking the max of both their received and former value.

Algorithm 1 RANDOM GOSSIP (RG)

When the sensor i wakes up (at global time t) :
- The sensor i chooses uniformly a neighbor $j \in \mathcal{N}_i$
- i and j exchange their value
- i and j update as follows
  $x_i(t+1) = x_j(t+1) = \max(x_i(t), x_j(t))$

- The BROADCAST algorithm uses the broadcast nature of the wireless channel: when a sensor wakes up, it broadcasts its estimate to all its neighbors. The sensors that receive a communication update their value accordingly.

Algorithm 2 BROADCAST (BC)

When the sensor i wakes up (at global time t) :
- The sensor i broadcasts its value to all its neighbors
- The sensors of the neighborhood $\mathcal{N}_i$ update as follows
  $\forall j \in \mathcal{N}_i,\ x_j(t+1) = \max(x_i(t), x_j(t))$

III. PROOFS OF CONVERGENCE

The goal of both algorithms is to reach *max-consensus* in finite time. Let $H_t = \{i \in V | x_i(t) = x_{max}\}$ be the set of vertices that have the maximum value at time t and $\bar{H}_t$ its complementary set. We can remark that once a sensor has the maximal value, its value does not change anymore ($\forall x \in \mathbf{x}(t), \max(x, x_{max}) = x_{max}$), hence the number of sensors informed with x_{max} (that is $|H_t|$ at time t) is non-decreasing in t. Let $\tau_N = \arg\min_t\{|H_t| = N\}$ be the first time the *max-consensus* is reached, we can easily notice that as soon as the consensus is reached, it is stable : $\forall t > \tau_N,\ |H_t| = N$. So, proving the convergence of an algorithm with high probability is equivalent to proving that $\mathbb{E}[\tau_N] < \infty$.

Theorem 1. *The* RANDOM GOSSIP *algorithm reaches max-consensus in finite time* τ_{RG} *with high probability.*

Proof: While max-consensus is not reached, there is at least one vertex in H_t that is connected to a vertex in $\bar{H}_t$ as the graph is supposed to be connected. The probability of choosing one of these two vertices at time t is $2/N$ and the probability that they exchange with each other is at least $1/d_{max}$. Hence,

$$\mathbb{P}[|H_{t+1}| = |H_t| + 1] \geq \frac{2}{Nd_{max}}$$

Considering the Bernouilli variable $B_t = |H_{t+1}| - |H_t|$ of parameter $p \geq 2/(Nd_{max})$, we can conclude that the time for a new sensor to be informed is a random variable U_t that follows a geometrical distribution of parameter p. Finally, the mean time for all N sensors to be informed is upper bounded by N times the expectation of U_t :

$$\mathbb{E}[\tau_{RG}] \leq d_{max}\frac{N^2}{2}.$$

■

Corollary 1. *The* BROADCAST *algorithm reaches max-consensus in finite time* τ_{BC} *with high probability.*

Proof: Similarly, the probability of incrementing the number of max-informed sensors is the probability of choosing an informed node connected to a uninformed one which is greater than $1/N$ as long as the consensus is not reached. Considering the worst case where only one sensor is informed when an informed node at the inner border of H_t broadcasts, we have

$$\mathbb{E}[\tau_{BC}] \leq N^2.$$

■

IV. CONVERGENCE SPEED BOUNDS

Simple convergence speed bounds have already been obtained in the previous section but they do not depend much neither on the graph nor on the algorithm; yet, simulations show that the max-consensus time changes significantly according to the underlying graph and the algorithm (see section V for an illustration). Therefore, we propose here graph-dependent bounds with a different approach for each algorithm.

A. Bound for the RANDOM GOSSIP

Result 1. $\mathbb{E}[\tau_{RG}] \leq \frac{Nd_{max}}{\lambda_2^L}\sum_{k=1}^{N-1}\frac{1}{k} \sim \frac{Nd_{max}\log(N-1)}{\lambda_2^L}$ *where* λ_2^L *is the second smallest eigenvalue of the Laplacian of the graph.*

Proof: It is straightforward to see that incrementing the number of sensors informed with the maximum value is the same as exchanging along an edge that goes from the set of informed sensors, H_t , to the set of uninformed sensors, $\bar{H}_t$.

Let S be a subset of V and $\partial S = \{e = (i,j) \in E | i \in S, j \in \bar{S}\}$ the edge frontier of S. We know from Cheeger [10], [11] that

$$|\partial S| \geq \lambda_2^L |S| \left(1 - \frac{|S|}{|V|}\right)$$

So, as the probability of choosing a particular edge is greater than $2/(Nd_{max})$ and as there are $|\partial H_t|$ *increasing edges*, we have

$$\begin{aligned}\mathbb{P}[|H_{t+1}| = |H_t| + 1] &\geq 2\frac{|\partial H_t|}{Nd_{max}} \\ &\geq 2\lambda_2^L |H_t| \frac{|V| - |H_t|}{d_{max}|V|^2}\end{aligned}$$

As in the proof of Theorem 1, this inequality gives us an upper bound for the parameter of the geometrical law representing the time to inform one more node knowing that $|H_t| = i$ sensors have already been informed. Then,

$$\mathbb{E}\left[\tau_{i+1}|i \text{ nodes informed }\right] \leq \frac{N d_{max}}{2\lambda_2^L}\frac{N}{i(N-i)}$$

Finally, $\mathbb{E}[\tau_N]$ is obtained by summing the above inequality over i from 1 to $N-1$. ■

B. *Bound for the* BROADCAST

Result 2. $\mathbb{E}[\tau_{BC}] \leq \bar{\epsilon}N + (\bar{\epsilon}-1)\, N \log\left(\frac{N-2}{\bar{\epsilon}-1}\right)$ *where $\bar{\epsilon}$ is the mean eccentricity of the graph.*

Proof: First, like in the works of Feige, Frieze and Grimmett [12], [7], let us consider the spanning tree subgraph $\mathcal{G}'$ of $\mathcal{G}$ rooted on a sensor r with maximal value at time 0. It is evident that the BROADCAST algorithm will reach max-consensus less quickly on $\mathcal{G}'$ than it does on $\mathcal{G}$.

Let us denote by $\mathcal{L}^{(i)}$ the set of nodes (or layer) at distance i from the root node. As the spanning tree transformation keeps the smallest distance between the nodes and the root node, the number of layers is the eccentricity of the root node $\epsilon_r = \max_{i\in V\setminus\{r\}} d(r,i)$ where $d(r,.)$ stands for the distance between r and another node in the graph.

Let us denote by $\tau^{(j+1)}$ the time for all the nodes of the layer $\mathcal{L}^{(j+1)}$ to be informed knowing that the previous layer $\mathcal{L}^{(j)}$ is informed. This time is the time needed to pick $M = \left|\mathcal{L}^{(j)}\right|$ elements among N by uniform picking. The expectation of this time is the sum over k going from 1 to M of the expectation of the time to pick one element among the k remaining that is

$$\begin{aligned}\mathbb{E}[\tau^{(j+1)}] &= \sum_{k=1}^{M}\sum_{t=0}^{\infty} t.\frac{k}{N}.\left(1-\frac{k}{N}\right)^{t-1} \\ &= N\sum_{k=1}^{M}\frac{1}{k}\end{aligned}$$

where the last equality comes from the fact that $\forall x : |1-x| < 1,\ \sum_{k=1}^{\infty} k\,(1-x)^{k-1} = 1/x^2$.

Now, we upper bound the time to inform all the graph $\mathcal{G}'$ by the time to inform it layer by layer. Also, in order to have a tighter inequality we separate the time to inform the first layer $\mathcal{L}^1$ which is equal to N.

$$\begin{aligned}\mathbb{E}[\tau_{BC}] &\leq N + \sum_{j=1}^{\epsilon_r-1}\mathbb{E}[\tau^{(j+1)}|\mathcal{L}^{(j)} \text{ informed}] \\ &= N + N\sum_{j=1}^{\epsilon_r-1}\sum_{k=1}^{|\mathcal{L}^{(j)}|}\frac{1}{k} \\ &\leq N + N\sum_{j=1}^{\epsilon_r-1}\left(\log\left(|\mathcal{L}^{(j)}|\right)+1\right)\end{aligned}$$

using the fact that $\forall n \geq 1, \sum_{k=1}^{n} \leq \log(n)+1$. Now, by applying the arithmetico-geometric inequality and the fact that the sum of the elements of layers 2 to $N-1$ is less than $N-2$ (the layer 1 is treated apart and there is no need to hit the last layer, the two layer containing both at least one node), we have

$$\begin{aligned}\mathbb{E}[\tau_{BC}] &\leq \epsilon_r N + N\log\left(\prod_{j=1}^{\epsilon_r-1}|\mathcal{L}^{(j)}|\right) \\ &\leq \epsilon_r N + (\epsilon_r-1)N\log\left(\frac{\sum_{j=1}^{\epsilon_r-1}|\mathcal{L}^{(j)}|}{\epsilon_r}\right) \\ &\leq \epsilon_r N + (\epsilon_r-1)N\log\left(\frac{N-2}{\epsilon_r-1}\right)\end{aligned}$$

Remarking that $\mathbb{E}[\epsilon_r] = \bar{\epsilon}$ since the root node can be anywhere in the graph along with Jensen's inequality concludes the proof. ■

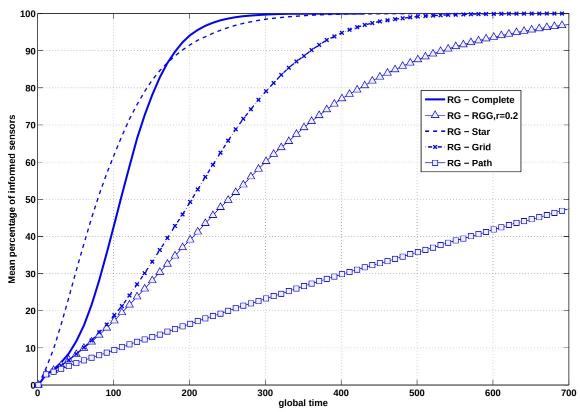
V. SIMULATIONS AND REMARKS ON THE TIGHTNESS

First, let's introduce Random Geometric Graphs (RGGs) [13]. These graphs are pretty well suited for Wireless Sensor Networks because they consist in uniformly disposing sensors on the unit square and creating a (perfect) link between two sensors if and only if their euclidean distance is lower than a fixed radius r_0. The simulations below are done on connected Random Geometric Graphs of radius $r_0 = \sqrt{4\log(N)/N}$.

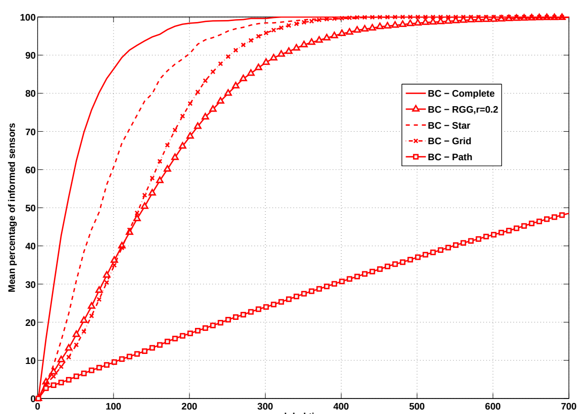
Figure 1 illustrates the fact that the convergence speed of the proposed algorithms can be very different according to the graph type (which is advocated in section IV) by plotting the mean percentage of informed sensors with respect to the global time. The mean percentage of informed sensors is evaluated by doing $1,000$ trials of the algorithm, the graph being unchanged. The graphs considered here are : the *complete* graph, the *star* graph where one node is connected to all others, the *path* graph, a *Random graph* (see below for a proper definition) and the *2D grid*, all with 49 sensors. Along with the fact that the convergence speed is strongly graph-dependent, this figure illustrates the fact that RANDOM GOSSIP and BROADCAST have different behaviors according to the encountered graph. On the *star* and *path* graphs, their convergence speed is quite the same whereas there are huge differences on the *grid* or on *random graphs*. Roughly speaking, this is due to the fact that the RANDOM GOSSIP works along edges whereas the BROADCAST algorithm works on neighborhoods.

In Figure 2, we plot the mean percentage of informed sensors versus the global time for a RGG sensors network with 75 nodes. As expected, the BROADCAST scheme informs a lot faster the unaware sensors, especially at the beginning. The curve of RANDOM GOSSIP suggest a 3-phase progression : *starting*, *permanent regime* and *ending* as stated in [8] and the proofs therein.

In Figure 3, we plot the *communication cost* for the RANDOM GOSSIP and the BROADCAST algorithm and the proposed bounds given in Section IV versus the number of sensors. In the Wireless Sensor Networks, the power constraints are often heavy and as the power consumption is rather associated to the number of communications than to the number of algorithms iterations, we advocate the use of the *communication cost* as a measure of an algorithm performance. This cost can be easily deduced from τ_N (the *iteration cost*) by multiplying it by the number of communications per iteration, namely, 2 for the RANDOM GOSSIP and 1 for the BROADCAST (we do not differentiate pairwise and broadcast communications here as their power consumption and speed are very close). We can see that the proposed bounds are good for both algorithms. It is also worth noticing these bounds are tight for the complete graph and the bound for the BROADCAST is also tight for the path graph. In our simulation, the proposed bounds are less tight when the number of sensors grows since we have remarked that the simulated RGG graphs are less connected, i.e., their smallest non-null eigenvalue of the Laplacian matrix decreases.



(a) RANDOM GOSSIP algorithm



(b) BROADCAST algorithm

Fig. 1. Mean percentage of informed sensors versus global time for different graph types.

VI. CONCLUSION AND FUTURE WORK

We proved that the RANDOM GOSSIP and the BROADCAST algorithms dealing with the distributed maximum value estimation over a WSN converge to the max-consensus and we gave tight bounds for their convergence speed.

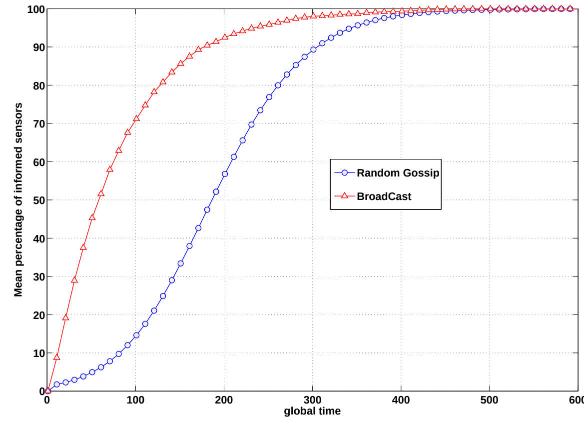


Fig. 2. Mean percentage of informed sensors versus global time for the RANDOM GOSSIP and the BROADCAST algorithms on RGGs.

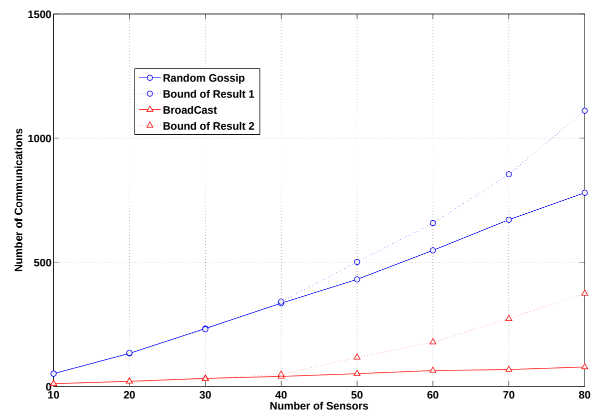


Fig. 3. Communication costs and proposed bounds for RANDOM GOSSIP and BROADCAST algorithms on RGGs.

We are currently working on concentration inequalities around the time of max-consensus in order to add a bound on the dispersion of consensus times. Also, it is easy to see that choosing uniformly the sensors might not always be the best thing to do (for example, as long as a sensor does not receive any update on its value, its action is useless). To cover up this defect, we are working on distributed clock management to better select the awaking sensors which would speed up the convergence.

ACKNOWLEDGMENT

This work was partially funded by the French Defence Agency (DGA) and the Télécom/Eurécom Carnot Institute.

REFERENCES

[1] J. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, M. I. T., Dept. of Electrical Engineering and Computer Science, 1984.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Analysis and optimization of randomized gossip algorithms," in *Proc. CDC Decision and Control 43rd IEEE Conf*, vol. 5, 2004, pp. 5310–5315.

[3] ——, “Randomized gossip algorithms,” *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[4] A. D. G. Dimakis, A. D. Sarwate, and M. J. Wainwright, “Geographic Gossip: Efficient Averaging for Sensor Networks,” *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1205–1216, 2008.

[5] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, “Broadcast Gossip Algorithms for Consensus,” *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2748–2761, 2009.

[6] F. Benezit, A. G. Dimakis, P. Thiran, and M. Vetterli, “Order-Optimal Consensus Through Randomized Path Averaging,” *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 5150–5167, 2010.

[7] U. Feige, D. Peleg, P. Raghavan, and E. Upfal, “Randomized broadcast in networks,” *Random Structures & Algorithms*, vol. 1, no. 4, pp. 447–460, 1990.

[8] N. Fountoulakis, A. Huber, and K. Panagiotou, “Reliable Broadcasting in Random Networks and the Effect of Density,” in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.

[9] N. Biggs, *Algebraic graph theory*. Cambridge Univ Pr, 1993, vol. 67.

[10] J. Cheeger, “Spectral geometry of singular Riemannian spaces,” *Journal of differential geometry*, vol. 18, no. 4, pp. 575–657, 1983.

[11] D. Spielman, “Spectral Graph Theory,” Lecture Notes, Yale University, 2009. [Online]. Available: http://www.cs.yale.edu/homes/spielman/561/

[12] A. Frieze and G. Grimmett, “The shortest-path problem for graphs with random arc-lengths,” *Discrete Applied Mathematics*, vol. 10, no. 1, pp. 57–77, 1985.

[13] M. Penrose, *Random geometric graphs*, O. S. in Probability, Ed. Oxford University Press, USA, 2003.