2015 Sixth Argentine Symposium and Conference on Embedded Systems (CASE)
ISBN: 978-987-45523-4-1 E-Book: 978-987-45523-5-8 Printed IEEE CATALOG: CFP15C02-PRT IEEE XPLORE: CFP15C02-ART

MESA: A Formal Approach to Compute Consensus in WSNs

Zacchigna, Federico G.
and Lutenberg, Ariel
Laboratorio de Sistemas Embebidos
Facultad de Ingeniería, Universidad de Buenos Aires
Paseo Colón 850, Ciudad Autónoma de Buenos Aires
federico.zacchigna@gmail.com, lse@fi.uba.ar

Vargas, Fabian
SiSC Group
Electrical Engineering Department
Catholic University (PUCRS)
Porto Alegre, Brasil
vargas@computer.org

***Abstract*—Consensus algorithms are implemented over Wireless Sensor Networks (WSN) for computing a distributed variable measurement or performing Data Fusion (DF). In a previous work the main idea of the MESA algorithm presented, but with no mathematical formalization. The MESA approach is a novel consensus algorithm that implements the transmission censoring technique aiming to reduce the energy used for reaching consensus. Moreover, it makes use of neighborhood data, which allows it to perform estimations over unknown distributed variables. This condition renders the proposed approach more efficient than the existing censoring techniques found in the literature.**

In this article a formalization of the MESA algorithm over cluster based WSN is presented. It is addressed from the statistical point of view of the algorithm, which gives us a better understanding of the consequences of transmission censoring and leads us to predict the results of its utilization.

I. Introduction

This work is focused on Wireless Sensor Networks (WSN) with a Cluster Head (CH) and N nodes around it. Each node performs a measurement and a Consensus Value (CV) is achieved after that:

$$x_c = f(\mathbf{X}) \tag{1}$$

where $\mathbf{X} = [X_1 X_2 \dots X_N]^T$, X_i is the i's node measurement and x_c is the CV. This is called Data Fusion (DF).

In this work it will be presented the statistical model that describes this process when the censoring technique is used. Section I-A will explain in more detail the measurement process and the differences between performing an estimation or a detection, and its difficulties. In section II the theoretical derivation of the algorithm is performed. The simulations, their results and the comparison with the theoretic derivation is shown in section III. Finally, conclusions are stated in section IV.

A. The measurement

The consensus process starts with the measurement of a set of initial values named initial vector or initial state. Though it is common to talk about a measurements, this is not necessarily precise. The initial vector can be obtained by measuring an environment variable with an actual sensor, for example in cognitive radios, where cooperative spectrum sensing is performed to detect the presence of a carrier by measuring its power [3] [2], but these initial values can also be obtained, for example as a node internal variable that can represent the remaining energy stored in its battery [1]. This allows the node to discover if its battery level is above or below the average value of the other nodes in the network. Such information if of prime importance, if one desires to extend the network lifetime as a whole, as long as possible.

We will consider each of the measurements a Gaussian distributed Random Variable (RV) with mean μ and variance σ^2. The Probability Density Function (PDF) of them may have the following characteristic:

$$X_i \overset{\mathcal{H}_h}{\sim} \mathcal{N}\left(\mu_h, \sigma^2\right), \text{ with } h \in \{0, \dots, h_{max}\} \tag{2}$$

or

$$X_i \sim \mathcal{N}\left(\mu, \sigma^2\right), \text{ with } \mu \in \mathbb{R} \tag{3}$$

for a detection or a estimation scheme respectively.

B. Different cases

Different schemes are shown in Table I. Cases I and IV are the standard consensus algorithms for clustered and distributed WSN respectively. Case II is the one studied in Rago's work [4]. In this article the analysis will be focused on cases I to III.

Case I provides no innovation at all, but it is derived to be used as a benchmark in this work. Case II is also not innovative, but in this work it is addressed in a different way, not from the likelihood ratio test as done by Rago [4], but from its statistics. Case III is the one that provides a new way of facing the consensus problem by using neighborhood data.

II. Theoretical Analysis

A. Case I

Consider a WSN with a CH and N nodes, that performs an initial measurement and $\mathbf{X} = [X_1, X_2, \dots, X_N]^T$ is obtained, with $X_i \sim \mathcal{N}(\mu, \sigma^2)$ $\forall i \in \{1, \dots, N\}$ and X_i, X_j are Independent Identically Distributed (IID) RVs $\forall i \neq j$.

Every node transmits its value to the CH, where the DF is performed as

$$X_{CH_I} = \sum_{i=1}^{N} w_i X_i = \mathbf{w}^T \mathbf{X} \tag{4}$$

ISBN: 978-987-45523-4-1 E-Book: 978-987-45523-5-8 Printed IEEE CATALOG: CFP15C02-PRT IEEE XPLORE: CFP15C02-ART

TABLE I
CASES CLASSIFICATION ACCORDING THE CENSORING TECHNIQUE. *knw*, *unkwn*, *ngbhd* AND *N/A*, ARE USED FOR *known*, *unknown*, *neighborhood* AND *Not Applicable* RESPECTIVELY.

Case	WNS type	Consensus strategy	Distribution type	Data usage	Estimation Detection
I	Cluster	standard	kwn/unkwn	N/A	Both
II		censoring	kwn	local	Detection
III		censoring	kwn/unkwn	ngbhd	Both
IV	Distributed	standard	kwn/unkwn	N/A	Both
V		censoring	kwn	local	Detection
VI		censoring	kwn/unkwn	ngbhd	Both

X_{CH_I} is the new random variable in the CH and is obtained as a Linear Combination (LC) of the $\mathbf{X}$'s elements. In this case the the average consensus is considered, which means that $w_i = 1/N \; \forall i$ and it is easy to demonstrate that

$$X_{CH_I} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{N}\right) \tag{5}$$

This consensus scheme is useful for performing estimations or detections.

B. Case II

As well as in case I, in case II the initial values are measured and $\mathbf{X}$ is obtained, again with X_i, $X_j \in$ IID RV $\forall\, i \neq j$ and X_i distributed as in (2) with $h \in \{h_0, h_1\}$. $\hat{X}_{II_i}$ is defined as

$$\hat{X}_{II_i} = \begin{cases} X_i & \text{if } x_i \in R_i \\ X_{II_i}^{est_{\mathcal{H}}} & \text{if } x_i \in \bar{R}_i \end{cases} \tag{6}$$

where R_i and $\bar{R}_i$ are two disjoint regions, in which the data is sent and censored respectively. Both are define by th_L and th_H, the two limits between them

$$R_i = \{x \,|\, th_L \leq x < th_H\} \tag{7}$$

$$\bar{R}_i = \{x \,|\, x < th_L \vee th_H \leq x\} \tag{8}$$

In (6) $X_{IIi}^{est_{\mathcal{H}}}$ is the estimated value, when no transmission is received and it differs according to the hypothesis $\mathcal{H}$ taken. This estimation can be performed in several ways. As follows two examples are shown:

1) $X_{IIi}^{est_{\mathcal{H}_0}} = \mathbb{E}[X_i \,|\, x_i \in \bar{R}_i, \mathcal{H}_0]$: This is an unbiased estimator for $\mathcal{H}_0$, but biased for $\mathcal{H}_1$. By exchanging $\mathcal{H}_0$ y $\mathcal{H}_1$ and adjusting th_L and th_H accordingly, the same scheme is reached.
2) $X_{IIi}^{est_{\mathcal{H}_0}} = \mathbb{E}[X_i \,|\, x_i \in \bar{R}_i, \mathcal{H}_0]\,\mathbf{P}(\mathcal{H}_0) + \mathbb{E}[X_i \,|\, x_i \in \bar{R}_i, \mathcal{H}_1]\,\mathbf{P}(\mathcal{H}_1)$: This estimator is biased in both cases (but less biased than in 1). Again $\mathcal{H}_0$ and $\mathcal{H}_1$ may be exchanged for getting the symmetric case.

Fig. 1 shows a a simple example of the distribution X_i y $\hat{X}_{II_i}$.

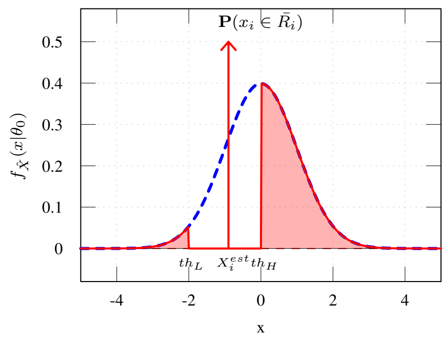

Fig. 1. Example of the $\hat{X}_{II_i}$ PDF. In this case $th_L = -2$, $th_H = 0$, $\mu = 0$ y $\sigma = 1$.

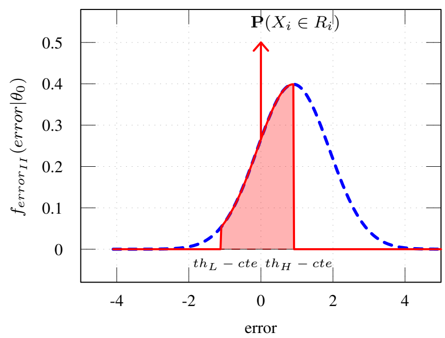

Fig. 2. Example of the $error_{II_i}$ PDF. In this case $th_L = -2 - cte$, $th_H = 0 - cte$, $\mu = 0$, $\sigma = 1$ and $cte = \mathbb{E}(x|th_L < x < th_H)$.

1) Considering the $\mathcal{H}_0$ hypothesis: By taking the first estimator previously explained ($X_{II_i}^{est_{\mathcal{H}_0}} = \mathbb{E}[X_i \,|\, x_i \in \bar{R}_i, \mathcal{H}_0]$) we may say

$$\mu_{\hat{X}_{II_i}} = \mu \tag{9}$$

The CH receives the i's $\hat{X}_{II_i}$ RV and performs de DF as

$$\hat{X}_{CH_{II}} = \sum_{i=1}^{N} \frac{\hat{X}_{IIi}}{N} = \frac{1}{N}\, \mathbf{1}^T\, \hat{\mathbf{X}}_{\mathbf{II}} \tag{10}$$

The $\hat{X}_{IIi}$ RV may be expressed as $\hat{X}_{II_i} = X_i + error_{II_i}$, which lead to

$$\hat{X}_{CH_{II}} = \sum_{i=1}^{N} \frac{X_i}{N} + \sum_{i=1}^{N} \frac{error_{II_i}}{N} = X_{CH_I} + error_{II} \tag{11}$$

where X_{CH_I} is the RV obtained in case I and $error_{II} = \sum_{i=1}^{N} \frac{error_{II_i}}{N}$. The $error_{II_i}$ RV PDF is shown in Fig. 2.

By applying the Central Limit Theorem (CLT), we may say that

$$\hat{X}_{CH_{II}} \overset{aprox.}{\sim} \mathcal{N}\left(\mu_{\hat{X}_{CH_{II}}}, \sigma^2_{\hat{X}_{CH_{II}}}\right) \tag{12}$$

or that

$$error_{II} \overset{aprox.}{\sim} \mathcal{N}\left(\mu_{error_{II}}, \sigma^2_{error_{II}}\right) \tag{13}$$

ISBN: 978-987-45523-4-1 E-Book: 978-987-45523-5-8 Printed IEEE CATALOG: CFP15C02-PRT IEEE XPLORE: CFP15C02-ART

We may now say that

$$error_{II_i} = \hat{X}_{II_i} - X_i \tag{14}$$

and

$$error_{II} = \hat{X}_{CH_{II}} - X_{CH_I} \tag{15}$$

both RV have zero mean because $\mu_{\hat{X}_{CH_{II}}} = \mu_{X_{CH_I}} = \mu$, and by using CLT and considering that N is big enough:

$$\sigma^2_{\hat{X}_{CH_{II}}} \approx \frac{\sigma^2_{\hat{X}_{II_i}}}{N} \tag{16}$$

and

$$\sigma^2_{error} \approx \frac{\sigma^2_{error_{II_i}}}{N} \tag{17}$$

The variance of the RVs are related by their correlation $\rho_{\hat{X}_{II_i}X_i}$ and $\rho_{X_i error_{II_i}}$ respectively.

$$\sigma^2_{error_{II}} = \frac{\sigma^2_{\hat{X}_{II_i}} + \sigma^2_{X_i} - 2\,\rho_{\hat{X}_{II_i}X_i}\,\sigma_{\hat{X}_{II_i}}\,\sigma_{X_i}}{N} \tag{18}$$

$$\sigma^2_{\hat{X}_{CH_{II}}} = \frac{\sigma^2_{X_i} + \sigma^2_{error_{II_i}} + 2\,\rho_{X_i error_{II_i}}\,\sigma_{X_i}\,\sigma_{error_{II_i}}}{N} \tag{19}$$

the correlation can be calculated if it is desired and it is easy to see that $\rho_{\hat{X}_{II_i}X_i} \geq 0$ and $\rho_{\hat{X}_i error_{II_i}} \leq 0$.

2) Considering the $\mathcal{H}_1$ hypothesis: Again taking the first estimator named before ($X^{est_{\mathcal{H}_1}}_{II_i} = \mathbb{E}[X_i \,|\, x_i \in \bar{R}_i, \mathcal{H}_1]$) the error is

$$error_{II_i} = \hat{X}_{II_i} - X_i \tag{20}$$

and

$$error_{II} = \hat{X}_{CH_{II}} - X_{CH_I} \tag{21}$$

which now does not have a zero mean, because $\mu_{\hat{X}_{II_i}} = \mu + bias_{\mathcal{H}_1}$ and $bias_{\mathcal{H}_1} \neq 0$, so the RV means are

$$\mu_{\hat{X}_{CH_{II}}} = \mu + bias_{\mathcal{H}_1} \tag{22}$$

$$\mu_{error_{II}} = bias_{\mathcal{H}_1} \tag{23}$$

with variances

$$\sigma^2_{\hat{X}_{CH_{II}}} \approx \frac{\sigma^2_{\hat{X}_{II_i}}}{N} \tag{24}$$

$$\sigma^2_{error_{II}} \approx \frac{\sigma^2_{\hat{X}_{error_{II_i}}}}{N} \tag{25}$$

respectively.

It worth to remark that the PDFs for $\hat{X}_{CH_{II}}$ and $error_{II}$ differ according to the $\mathcal{H}$ considered and of course this leads to different $\sigma^2_{\hat{X}_{CH_{II}}}$ and $\sigma^2_{error_{II}}$.

3) Conclusions for the case II: The efficiency of the algorithm can be measured by means of the metrics presented in II-E and hangs on the estimator used when the transmissions are censored:

- Unbiased for $\mathcal{H}_0$ and biased for $\mathcal{H}_1$.
- Unbiased for $\mathcal{H}_1$ and biased for $\mathcal{H}_0$.
- Both biased but less than in the two previous cases.

The bias is a function of $\mathbf{P}(\mathcal{H}_0)$ and $\mathbf{P}(\mathcal{H}_1)$ and $\bar{R}_i$ region should be on the side of the most probable hypothesis, so that the bias obtained will be smaller with the same amount of censored transmissions.

In [4] it is proved that for minimizing the *miss* or the *false detection* probabilities in the detection, th_L should be equal to $-\infty$, but should be different if we want to minimize another metric.

4) Problems to be solved in case II: In this section some useful ways of using the censoring technique are enumerated now:

- Given:
  - a transmission constraint, that is equivalent to an area limitation:
    - $th_c \geq \mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_0)$, or
    - $th_c \geq \mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_1)$, or
    - $th_c \geq \mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_1) + \mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_0)$.

    where th_c is the threshold value of the constraint.
  - μ and σ for $\mathcal{H}_0$ and $\mathcal{H}_1$.

  find the th_L y th_H, so that one of the metrics named in section II-E is minimized. Equivalent to find the optimum DR.
- Given:
  - a constraint in one of the metrics in section II-E.
  - μ and σ for $\mathcal{H}_0$ and $\mathcal{H}_1$.

  find the DR (th_L and th_H) so that the number of censored transmissions is maximized:
  - $\mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_0)$, or
  - $\mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_1)$, or
  - $\mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_0) + \mathbf{P}(x_i \in \bar{R}_i | \mathcal{H}_1)$

C. Case III

As well as in case I and II the initial values are measured. The X_is RV may be distributed as in (2) or (3). The initial vector $\mathbf{X}$ is obtained, where X_i, X_j are IID RV $\forall i \neq j$.

In this work the following assumptions and restrictions are considered for limiting the scope and the length of this article.

- The nodes are synchronized and they transmit in order, one node per Time Slot (TS) as in [5]. In each realization:
  - The first node is randomly selected.
  - In the next TS the corresponding node decides whether to transmit or to censor the data according to the DR.
  - The realization ends when the N nodes had their TS, including the first one.
- The nodes are able to hear the data transmitted only from the previous node.

ISBN: 978-987-45523-4-1 E-Book: 978-987-45523-5-8 Printed IEEE CATALOG: CFP15C02-PRT IEEE XPLORE: CFP15C02-ART

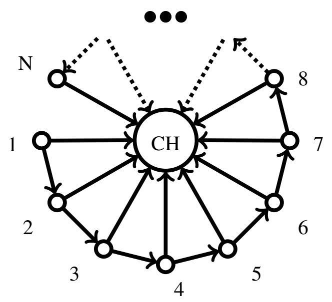

Fig. 3. This figures shows the CH and the N nodes around it and how they communicate. The i is able to send data to the CH and to hear data from the $i-1$ node.

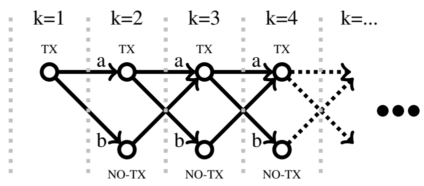

Fig. 4. Graph representing the algorithm and the transmission probability in each TS.

TABLE II
PROBABILITY IN EACH TS

Turn (k)	1	2	3	4	5	...
$\mathbf{P}(\text{TX}_k)$	1	a	a^2+b	$a^3+2\,ab$	$a^4+3\,a^2+b^2$	...
$\mathbf{P}(\text{NO-TX}_k)$	0	b	ab	a^2b+b^2	$a^3b+2\,ab^2$	...

Fig. 3 shows the corresponding communication diagram used in this section. In section II-D other interesting variations to be considered are named.

The algorithm may be resumed in the following way:

- Node in the first TS sends its data.
- The k node sends it if the $k-1$ has not send its data.
- The k node sends it if the $k-1$ has send its data and the DR in k is fulfilled.
- The k node does not sends its data if $k-1$ sent it and the DR in k is not fulfilled.

by saying 'k node', it means the 'the corresponding node in the k TS'. A good way of representing the algorithm is shown in Fig. 4, where the horizontal axis represents the TSs. In each TS the corresponding node has a probability $\mathbf{P}(\text{TX}_k)$ of performing the transmission and a probability $\mathbf{P}(\text{NO-TX}_k)$ of censoring it, except in the first where the transmission is mandatory.

In the first TS, it is mandatory for the node to transmit, after that, in the second TS, the probability of transmitting is a and the probability of censoring is b. Each time the previous node has transmitted, there exist a probability a for a transmission and a probability b for a censorship. If the previous node has censored the data, then is mandatory to perform the transmission, this is that the transmission probability is 1. The probabilities for each of the firsts 5 TSs is shown in Table II.

The DR used is defined as:

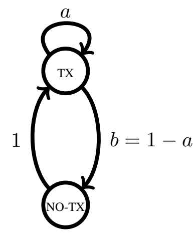

Fig. 5. The Markov chain representing the communication scheme in Fig. 4. The initial state of the chain shall always be the 'TX' state.

- If $(X_i - X_{i-1}) > th$ then the transmission is performed.
- If $(X_i - X_{i-1}) \leq th$ then the transmission is censored.

it uses th as parameter and $a = f(\sigma, th) = \mathbf{P}(X_i - X_{i-1} > th)$, $b = 1-a$ and $\mathbf{P}(\text{NO-TX}_k) = 1-\mathbf{P}(\text{TX}_k)$. The X_{i-1} PDF is considered Gaussian in both situations (TX and NO-TX). This is a valid approximation used in this work for small values of th, thus a and b may be considered constant for every k There exists different ways of finding the probabilities in each TSs. As follows two of them are shown:

1) Through recurrence equations. We know:

$$\mathbf{P}(\text{TX}_k) = a\,\mathbf{P}(\text{TX}_{k-1}) + \mathbf{P}(\text{NO-TX}_{k-1}) \quad (26)$$

$$\mathbf{P}(\text{NO-TX}_k) = b\,\mathbf{P}(\text{TX}_{k-1}) \quad (27)$$

combining (26) and (27) leads us to the recurrence equation (28) with the initial conditions in (29).

$$\mathbf{P}(\text{TX}_k) = a\,\mathbf{P}(\text{TX}_{k-1}) + b\,\mathbf{P}(\text{TX}_{k-2}) \quad (28)$$

$$\mathbf{P}(\text{TX}_1) = 1,\ \mathbf{P}(\text{NO-TX}_1) = 0 \quad (29)$$

This particularly equation is homogeneous and with constant coefficients and thus has a very simple solution:

$$\mathbf{P}(\text{TX}_k) = A\,x_0^k + B\,x_1^k$$

x_0 and x_1 are the characteristic polynomial roots of (28) $(t^2 - a\,t - (1-a) = 0)$ A and B are found through (29).

2) Through Markov chains. The process in Fig. 4 may be considered as the Markov chain showed in figure 5. The Markov operator may be defined as:

$$\mathbf{K} = \begin{bmatrix} a & 1 \\ b & 0 \end{bmatrix} = \begin{bmatrix} a & 1 \\ 1-a & 0 \end{bmatrix} \quad (30)$$

and the probability in the k TS can be calculated as:

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{P}(\text{TX}_k) \\ \mathbf{P}(\text{NO-TX}_k) \end{bmatrix} = K^{k-1}\,\mathbf{P}_0\ ,\ \mathbf{P}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Both solutions leads as to the same result showed in Fig. 6, that shows $\mathbf{P}(\text{TX}_k)$ and $\mathbf{P}(\text{NO-TX}_k)$, as a function of k. It can be seen that process converges and the convergence rate is $\propto a$.

These probabilities are needed for computing the RV obtained in the CH after the DF. In each TS, the CH incorporates a new $\hat{X}_{III_i}$ to the set used to perform the DF:

$$\hat{X}_{III_i} = \begin{cases} X_i & \text{if } x_i \in R_i \\ X_{i-1} & \text{if } x_i \in \bar{R}_i \end{cases} \quad (31)$$

ISBN: 978-987-45523-4-1 E-Book: 978-987-45523-5-8 Printed IEEE CATALOG: CFP15C02-PRT IEEE XPLORE: CFP15C02-ART

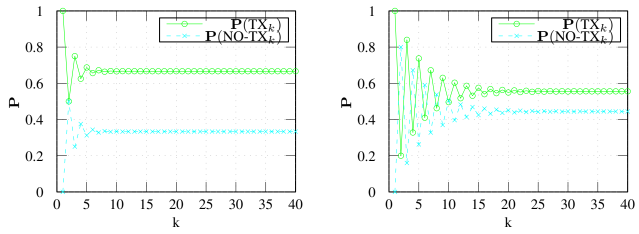

Fig. 6. The evolution of the process of Fig. 4 for $a = 0.5$ and $a = 0.2$ and $N = 40$.

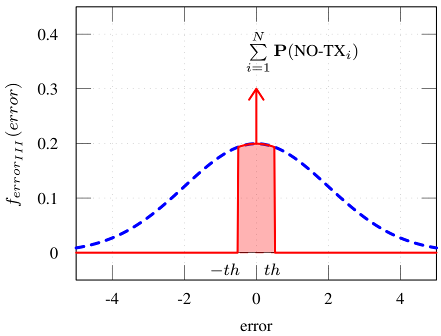

Fig. 7. The error PDF for the case III.

Calculating the variance of $\hat{X}_{CH_{III}} = \sum_{i=1}^{N} \hat{X}_{III_i}$ is not easy because of the complex correlation that exists between $\hat{X}_{III_i}$ and $\hat{X}_{III_j}$. For dealing with this problem we rewrite

$$\hat{X}_{III_i} = X_i + error_{III_i} \tag{32}$$

with

$$error_{III_i} = \begin{cases} 0 & \text{if } x_i \in R_i \\ X_i - X_{i-1} & \text{if } x_i \in \bar{R}_i \end{cases} \tag{33}$$

for small values of th its PDF can be considered as a delta in 0 with area $\mathbf{P}(\text{TX}_i)$ and the rest of the area is a trimmed Gaussian RV in $\pm th$ with area $\mathbf{P}(\text{NO-TX}_i)$. The original Gaussian PDF has a zero mean and variance equal to $2\,\sigma^2$. The RVs $error_{III_i}$ and $error_{III_j}$ are not independent, but because of nature of the algorithm it is possible to demonstrate that $\rho_{error_{III_i} error_{III_j}} = 0$. The $error_{III}$ PDF is shown in Fig. 7.

By doing this, we may calculate the $\hat{X}_{CH_{III}}$ value in the CH as:

$$\hat{X}_{CH_{III}} = \left(\frac{1}{N}\sum_{i-1}^{N} X_i\right) + \left(\frac{1}{N}\sum_{i-1}^{N} error_{III_i}\right) = \\ = X_{CH_I} + error_{III} \tag{34}$$

where X_{CH_I} is the one from case I, the 'ideal case'. With these approach we may say, that in the CH we compute the mean of every X_i plus an error. This error ($error_{CH_{III}}$) is known and we are able to present its statistics:

$$error_{III} = \hat{X}_{CH_{III}} - X_{CH_I} \tag{35}$$

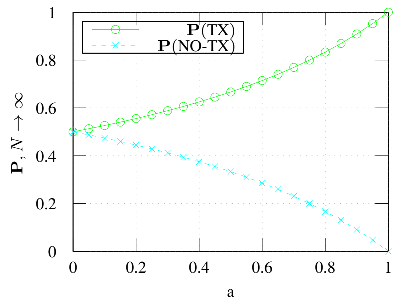

Fig. 8. The convergent values of $\mathbf{P}(\text{TX})$ and $\mathbf{P}(\text{NO-TX})$ as a function of a probability, when $N \to \infty$.

which has zero mean because $\mu_{\hat{X}_i} = \mu$, and by using CLT and considering that N is big enough:

$$\sigma^2_{error_{III}} \approx \frac{\sigma^2_{error_{III_i}}}{N} \tag{36}$$

The variance of the RVs are related by their correlation $\rho_{\hat{X}_{III_i} X_i}$

$$\sigma^2_{error_{III}} = \frac{\sigma^2_{\hat{X}_{III_i}} + \sigma^2_{X_i} - 2\,\rho_{\hat{X}_{III_i} X_i}\,\sigma_{\hat{X}_{III_i}}\,\sigma_{X_i}}{N} \tag{37}$$

It is easy to see that $\rho_{\hat{X}_{III_i} X_i} \geq 0$.

To conclude this section, Fig. 8 shows the convergence values for $\mathbf{P}(\text{TX}) = \sum_{i=1}^{N} \mathbf{P}(\text{TX}_i)$ and $\mathbf{P}(\text{NO-TX}) = \sum_{i=1}^{N} \mathbf{P}(\text{TX}_i)$ as a function of a probability, when $N \to \infty$. It can be interpreted as the superior limit of the Censored Transmission Rate (CTR) as a function of a. For $N < \infty$, the CTR decreases, the first mandatory and the subsequent ones do not influence too much in the convergence.

D. Case III variations

Possible variations of case III are: *a)* The asynchronous case III. *b)* The case III, when two or more previous nodes can be heard. *c)* The case III, when the amount of nodes than can be heard depends on i. *d)* The case III, when there exists a hearing probability of others node's transmissions (depending on the distance, energy, etc.), becoming random graphs. *e)* Any combination of the previous cases.

They will not be be analyzed in this work.

E. Metrics

The following metrics are suggested for evaluating the algorithm performance:

- Statistic distance between distributions $d\left(f_{X_{CH}}, f_{\hat{X}_{CH}}\right)$.
- The 'miss' and 'false detection' probabilities.
- Mean Square Error (MSE) as in [6].
- CTR.
- Time or iterations needed for reaching consensus.

ISBN: 978-987-45523-4-1 E-Book: 978-987-45523-5-8 Printed IEEE CATALOG: CFP15C02-PRT IEEE XPLORE: CFP15C02-ART

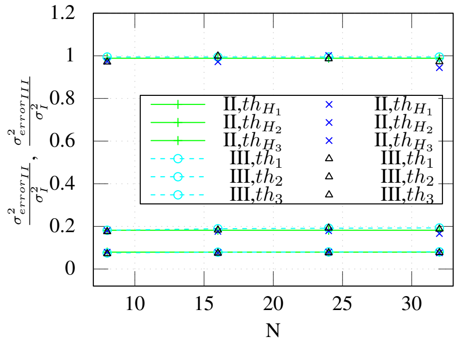

Fig. 9. The relative variance of the error with respect to the variance of the case I as function of N. Theoretic results plotted with lines and markers, simulated ones only with markers.

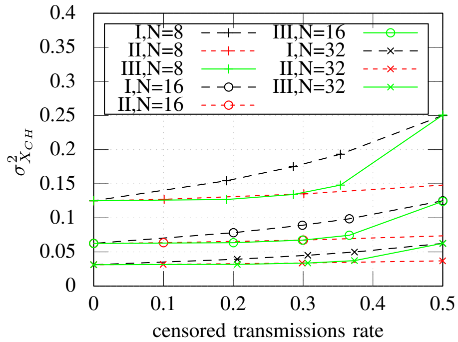

Fig. 10. The X_{CH} variance as a function of the CTR. Case I works as benchmark, in it no transmission is censored and this rate shall be interpreted as having less sensor data to perform the DF. For example, when CTR $r = 0.25$ and $N = 8$, the result for case I is the same as having 6 sensors.

III. Simulations and Results

This section deals with the validation of the proposed mathematical analysis. The considered network topology is the one presented in section II.

The first results to be shown are the ones that validates the theory, this is the error variance in Fig. 9, where it can be seen that the predicted results are in accordance with the simulated ones. This figures shows the relative variance of the error wrt. the variance in case I ($\frac{\sigma^2_{error}}{\sigma^2_{CH_I}}$).

The most interesting results from the point of view of the algorithm performance are show in Fig. 10. The aim of performing DF is to obtain a better measurement, this is the measurement with the smallest possible variance. In the standard case $\sigma^2_{CH} \propto 1/N$. In this figure, the following convention is adopted: In case I, no transmission is censored, so the variance should not depend on CTR, it should be interpreted as having less quantity of sensors. For example, given N nodes and CTR $= r$, is equivalent to have $(r\,N)$ nodes, which leads to a $\sigma^2_{CH} = \frac{\sigma^2}{r\,N}$. The only objective of doing this, is to have a benchmark value to compare the rest of the results.

From Fig. 10 the following results may be extracted:

- $\sigma^2_{CH_{III}} = \sigma^2_{CH_I}$ for $r \in \{0, 0.5\}$: The algorithm becomes dummy for those values of r, no DR is evaluated. For $r = 0$ every node performs the transmission and for $r = 0.5$ every odd node performs the transmission, being the same as having $\lceil \frac{N}{2} \rceil$ nodes.
- $\sigma^2_{CH_{III}} < \sigma^2_{CH_I}$ for $0 < r < 0.5$. In this region is where the algorithm achieves better results, between 0.10 to 0.35 the $\sigma^2_{CH_{III}}$ remains almost the same, while the censored transmissions increase. Moreover, the performance is similar to the one in case II.
- For CTR above 0.35, the algorithm derates and case II outperforms it. This is clearly because case II has the advantage of knowing the distribution a priori, but at the same time case II is not able of performing estimations.

IV. Conclusions

MESA is a novelty consensus algorithm over WSN for performing distributed detection or estimation, while using a censoring technique. The preliminary theoretic model of the MESA algorithm, has been successfully derived.

In particular the scheme presented in this work, where only one previous transmission can be heard by the nodes, the maximum CTR can rise up to $\lfloor \frac{N}{2} \rfloor \frac{1}{N}$. The algorithm demonstrated to have the better performance when the CRT is not close to zero or close to its upper limit, but around rates between 0.10 and 0.35. This rates may increase for other topologies.

The results of the work encourage us to develop the theory for the variations presented in II-D, which are left for future works.

References

[1] Dong Wentao; Li Wenfeng; Zhang Fan, "Consensus algorithm for energy consumption of wireless sensor networks," Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on , vol., no., pp.2620,2625, 9-12 Oct. 2011 doi: 10.1109/ICSMC.2011.6083992

[2] Chunhua Sun; Wei Zhang; Ben, K., "Cluster-Based Cooperative Spectrum Sensing in Cognitive Radio Systems," Communications, 2007. ICC '07. IEEE International Conference on , vol., no., pp.2511,2515, 24-28 June 2007 doi: 10.1109/ICC.2007.415

[3] Ian F. Akyildiz, Brandon F. Lo, and Ravikumar Balakrishnan. 2011. Cooperative spectrum sensing in cognitive radio networks: A survey. Phys. Commun. 4, 1 (March 2011), 40-62. DOI=10.1016/j.phycom.2010.12.003 http://dx.doi.org/10.1016/j.phycom.2010.12.003

[4] Rago, Constantino; Willett, P.; Bar-Shalom, Y., "Censoring sensors: a low-communication-rate scheme for distributed detection," Aerospace and Electronic Systems, IEEE Transactions on, vol.32, no.2, pp.554,568, April 1996 doi: 10.1109/7.489500

[5] Zacchigna, Federico G.; Lutenberg, Ariel, "A novel consensus algorithm proposal: Measurement estimation by silent agreement (MESA)," Embedded Systems (SASE/CASE), 2014 Fifth Argentine Symposium and Conference on , vol., no., pp.7,12, 13-15 Aug. 2014 doi: 10.1109/SASE-CASE.2014.6914462

[6] Pereira, S.S.; Pages-Zamora, A., "Consensus in Correlated Random Wireless Sensor Networks," Signal Processing, IEEE Transactions on , vol.59, no.12, pp.6279,6284, Dec. 2011 doi: 10.1109/TSP.2011.2166552