A Simple Information-Theoretic Analysis of Clustered Sensor Networks with Decentralized Detection

Marco Martalò, *Member, IEEE*, and Gianluigi Ferrari, *Member, IEEE*

***Abstract*—In this letter, we present a simple information-theoretic framework to analyze clustered sensor networks with hierarchical multi-level majority-like fusion and decentralized detection. The sensor nodes observe a binary phenomenon and transmit their own data to an access point (AP), possibly through intermediate fusion centers (FCs). We investigate the impact of uniform and non-uniform clustering on the system performance, evaluated in terms of mutual information between the true phenomenon status and its estimate at the AP. Being the overall system binary-input binary-output (BIBO), it will be shown that the probability of decision error (P_e) is a specific function of the input-output mutual information (I). In other words, the network operational point lies over a specific $P_e - I$ curve and depends on the network characteristics (e.g., topology, observation and communication noise levels, etc.).**

***Index Terms*—Clustered sensor networks, decentralized detection, noisy communication links, information-theoretic framework.**

I. Introduction and Motivation

DISTRIBUTED detection has been an active research field for a long time [1]. The increasing interest for sensor networks has spurred a significant activity on the design of efficient distributed detection techniques [2]. Information-theoretic approaches have also been proposed for the study of sensor networks with decentralized detection. In [3], the authors propose a framework to characterize a sensor network in terms of its entropy and false alarm/missed detection probabilities. In [4], the mutual information is evaluated in a scenario with censoring sensors which transmit their local likelihood ratios, by maximizing the probability of correct decision.

In this letter, we consider a sensor networking scenario where the information collected by a sensor can be transferred to the access point (AP) through *multiple hops*, i.e., by exploiting intermediate nodes as *relays*. Besides the need to support multiple communications, in several scenarios the information received by a relay from sensors placed in a specific region might be redundant. In this case, the relay does not need to forward the information received by all sensors, but can extract a concise "picture" of the status of the monitored scenario.

In order to carry out the analysis outlined in the previous paragraph, we consider network scenarios where sensors, which observe a binary phenomenon, are grouped into *clusters* and are directly connected with local fusion centers (FCs) (one per cluster), denoted as first-level FCs. We assume that the observed phenomenon is spatially constant. This is meaningful, for example, when it is of interest to detect if the phenomenon under observation (e.g., temperature, humidity, pressure) overcomes a critical threshold. Each first-level FC makes a local decision based on the data collected from its associated sensors and then transmits its decision to the AP, possibly through other intermediate FCs.

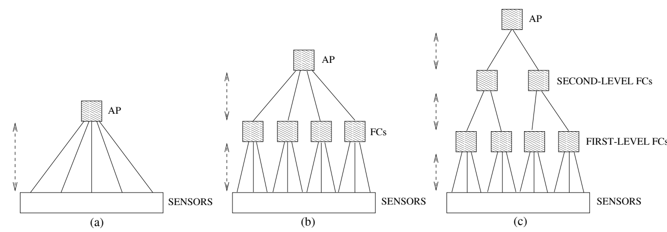

Fig. 1. Basic structures for sensor networks with decentralized detection. Three cases are shown: (a) absence of clustering, (b) uniform clustering with two levels of information fusion, and (c) uniform clustering with three levels of information fusion.

While a communication-theoretic framework is presented in [5], the goal of the current letter is to derive a simple, yet insightful, information-theoretic perspective on clustered sensor networks with decentralized detection. This is carried out by modelling the entire network as a binary-input binary-output (BIBO) system, where the binary input is the observed phenomenon and the binary output is its estimate at the AP. The key metric in this analysis is the mutual information between the true status of the phenomenon under observation and its estimate at the AP. As it will be shown, its value determines the probability of decision error at the AP. While clustered architectures are inherently scalable and allow to manage a very large number of sensors, our information-theoretic framework is expedient to concisely investigate (i) the impact of the specific clustered topologies (either uniform or non-uniform) and (ii) the cost, in terms of observation accuracy at the sensors and robustness against communication noise, of multi-level information fusion algorithms.

II. Preliminaries

The reference scenario is shown in Fig. 1 for three possible networking schemes: (a) no clustering, (b) two decision level uniform clustering,[1] and (c) three decision level uniform clustering. Note that topologies (b) and (c) can be easily extended to scenarios with non-uniform clustering, i.e., scenarios where the cluster size may vary from cluster to cluster. In the following, the clustered configurations will be denoted by

[1]By uniform clustering, we mean that the number of nodes per cluster is the same among all clusters.

Manuscript received November 23, 2009. The associate editor coordinating the review of this letter and approving it for publication was H.-H. Chen.

The authors are with the Wireless Ad hoc and Sensor Networks (WASN) Laboratory, Department of Information Engineering, University of Parma, Italy (e-mail: {marco.martalo, gianluigi.ferrari}@unipr.it).

Digital Object Identifier 10.1109/LCOMM.2010.06.092286

using the number of sensors in each cluster. For instance, if a cluster contains 14 sensors and other two clusters contain 1 sensor each, the configuration will be denoted as 14-1-1.

We consider a network scenario where N sensors observe, in a noisy manner, a *common binary phenomenon* whose status is defined as follows:

$$H = \begin{cases} H_0 & \text{with probability } p_0 \\ H_1 & \text{with probability } 1 - p_0 \end{cases}$$

where $p_0 \triangleq P(H = H_0)$. The sensors are clustered into $n_c < N$ groups, and each sensor can communicate only with its local first-level FC. The first-level FCs collect data from the sensors in their corresponding clusters and make local decisions on the status of the binary phenomenon. In a scenario with two levels of information fusion, each local FC transmits to the AP, which makes the final decision, denoted as $\widehat{H}$.

According to the approach followed in [5], a common signal-to-noise ratio (SNR) at the sensors, denoted as $\mathrm{SNR}_{\mathrm{sensor}}$, can be defined. The i-th sensor ($i = 1, \ldots, N$) makes a decision comparing its observation r_i with a threshold value τ_i and computes a local decision $u_i = U(r_i - \tau_i)$, where $U(\cdot)$ is the unit step function. In the following, we assume that all sensors use the same decision threshold τ and its value will be optimized in all considered scenarios, by minimizing the probability of decision error at the AP. In particular, the optimized value of the common threshold is around $\sqrt{\mathrm{SNR}_{\mathrm{sensor}}}/2$ [5].

In a scenario with noisy communication links, modeled as binary symmetric channels (BSCs), the decision u_i sent by the i-th sensor can be *flipped* with a probability corresponding to the cross-over probability of the BSC model and denoted as p [5]. The received bit at the fusion point (either an FC for clustered networks or directly the AP in the absence of clustering), referred to as $u_i^{(\mathrm{r})}$, can be expressed as

$$u_i^{(\mathrm{r})} = \begin{cases} u_i & \text{with probability } 1 - p \\ 1 - u_i & \text{with probability } p. \end{cases}$$

The probability of decision error at the AP is defined as

$$P_{\mathrm{e}} \triangleq p_0 P\left(\widehat{H} = H_1 | H_0\right) + (1 - p_0) P\left(\widehat{H} = H_0 | H_1\right). \quad (1)$$

In [5], it is shown how to numerically evaluate (1) and the behavior of P_{e} is investigated for many clustering configurations.

III. Joint Communication/Information-Theoretic Characterization

The considered sensor network schemes can be modeled as "black boxes" with a binary input (the phenomenon H) and a binary output (the decision $\widehat{H}$ at the AP). Using the model introduced in Section II, the final decision $\widehat{H}$ can be described as a binary random variable characterized by the parameter[2] $\hat{p}_0 \triangleq P(\widehat{H} = H_0)$, which can be rewritten as

$$\hat{p}_0 = p_0 P\left(\widehat{H} = H_0 | H_0\right) + (1 - p_0) P\left(\widehat{H} = H_0 | H_1\right). \quad (2)$$

[2]Note that $\hat{p}_0 \triangleq P(\widehat{H} = H_0)$ (relative to the decision $\widehat{H}$) is different from the *a priori* probability of the phenomenon $p_0 \triangleq P(H = H_0)$ given in Section II.

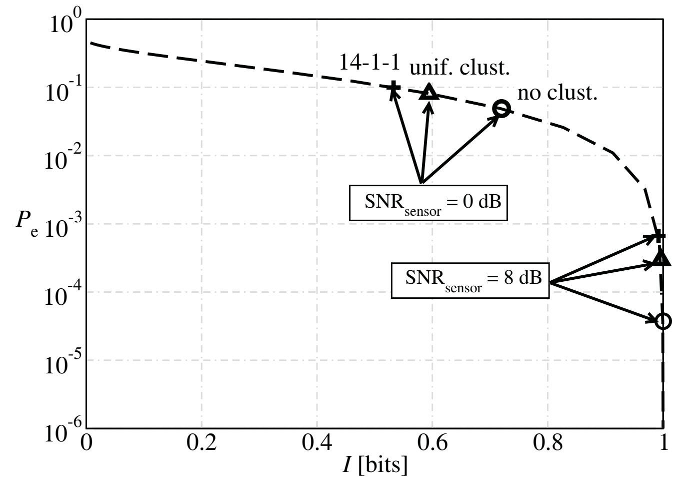

Fig. 2. Probability of decision error, as a function of the mutual information. The operational points for various clustering configurations and two sensor SNRs are shown.

We remark that equation (2) may look identical to (1). In (2), however, the first term at the right-hand side contains $P(\widehat{H} = H_0 | H_0)$, whereas in (1) it contains $P(\widehat{H} = H_1 | H_0)$. The value of $\hat{p}_0$ given by (2) embeds all network characteristics (namely, the topology, the sensor SNR, and the level of communication noise) which determine the expression of $\widehat{H}$.

The binary entropy of the random variable $\widehat{H}$ is [6]

$$\mathcal{H}_{\mathrm{e}}(\widehat{H}) = \mathcal{H}_{\mathrm{e}}(\hat{p}_0) \triangleq \hat{p}_0 \log_2 \frac{1}{\hat{p}_0} + (1 - \hat{p}_0) \log_2 \frac{1}{1 - \hat{p}_0}.$$

The mutual information of the BIBO sensor network, denoted as $I(H; \widehat{H})$, can then be written as [6, ch. 2]

$$I(H; \widehat{H}) = \mathcal{H}_{\mathrm{e}}(\widehat{H}) - \mathcal{H}_{\mathrm{e}}(\widehat{H} | H)$$

where $\mathcal{H}_{\mathrm{e}}(\widehat{H} | H)$ is the conditional entropy of $\widehat{H}$ given H and can be written as

$$\begin{aligned} \mathcal{H}_{\mathrm{e}}(\widehat{H} | H) &= P(\widehat{H} = H_0 | H) \log_2 \frac{1}{P(\widehat{H} = H_0 | H)} \\ &\quad + P(\widehat{H} = H_1 | H) \log_2 \frac{1}{P(\widehat{H} = H_1 | H)}. \end{aligned}$$

After a few manipulations, the mutual information becomes

$$\begin{aligned} I(H; \widehat{H}) &= \mathcal{H}_{\mathrm{e}}\left(p_0(1 - p_{10}) + (1 - p_0) p_{01}\right) \\ &\quad - p_0 \mathcal{H}_{\mathrm{e}}(p_{10}) - (1 - p_0) \mathcal{H}_{\mathrm{e}}(p_{01}) \end{aligned} \quad (3)$$

where $p_{ij} \triangleq P(\widehat{H} = H_i | H_j)$, $i, j \in \{0, 1\}$, $i \neq j$, are the component conditional probabilities in (1). Note that I is then a function of the sensor SNR. Therefore, P_{e} can be investigated as a parameterized (in $\mathrm{SNR}_{\mathrm{sensor}}$) function of I.

Although all the results presented in the following are obtained in a scenario with $N = 16$ sensors and $p_0 = 0.5$ (binary phenomenon with equally likely outcomes), the same framework can be directly applied with other values of N and p_0. In Fig. 2, the probability of decision error is shown, as a function of the mutual information, for the following topologies: no clustering (circles), uniform clustering (triangles), and non-uniform clustering (pluses, 14-1-1 configuration). The communication links are ideal. As one can see, the operational points associated with different network topologies

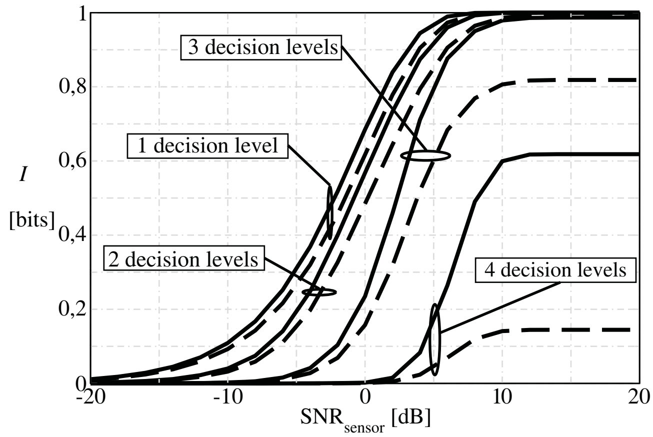

Fig. 3. Mutual information, as a function of the sensor SNR, in a sensor network with uniform clustering and noisy communication links with two possible values of the cross-over probability: (i) $p = 0.01$ (solid lines) and (ii) $p = 0.05$ (dashed lines).

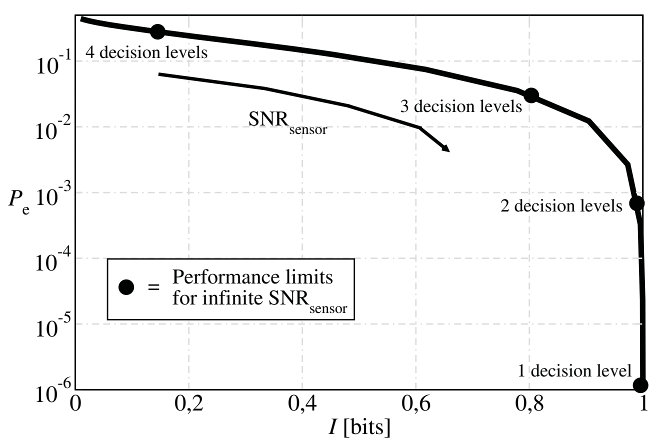

Fig. 4. Probability of decision error, as a function of the mutual information, in a scenario with uniform clustering and noisy communication links ($p = 0.05$). The limiting ($\mathrm{SNR}_{\mathrm{sensor}} \rightarrow \infty$) operational point with various number of decision levels are shown.

lie on the same curve. In other words, for a given value of the mutual information, the probability of decision error is fixed. In particular, for a given network topology, the position of the network operational point over the curve in Fig. 2 is associated with a specific sensor SNR. However, the same sensor SNR corresponds to different values of the mutual information in clustered and non-clustered scenarios—in Fig. 2, a few representative points, associated with two sensor SNRs, are indicated. As one can see, for a given sensor SNR, the mutual information is highest (and the probability of decision error lowest) with no clustering and the loss with non-uniform clustering is higher than with uniform clustering. Similar curves can be derived for other scenarios, e.g., for a large number of sensors, with more than two decision levels, and in the presence of noisy communication links between sensors and first-level FCs, as will be shown in the following (Fig. 4). In all cases, the information-theoretic characterization of the network behavior does not change: for a fixed value of the mutual information, the probability of decision error is uniquely determined.

In Fig. 3, the mutual information is shown, as a function of the sensor SNR, in a network with uniform clustering (with all possible numbers of decision levels). The communication links between the sensors and the first-level FCs are noisy, with a cross-over probability p equal to either 0.01 (solid lines) or 0.05 (dashed lines).

As one can see, the presence of noise in the communication links limits the maximum achievable mutual information, i.e., the maximum information transfer rate across the network. This phenomenon is more pronounced the larger is the number of decision levels. In fact, in this case the information loss across the network is the highest possible.

In Fig. 4, the probability of decision error is shown, as a function of the mutual information, in a scenario with uniform clustering. Communication links between sensors and first-level FCs are noisy, with cross-over probability $p = 0.05$.

The limiting (for $\mathrm{SNR}_{\mathrm{sensor}} \rightarrow \infty$) operational points on the $P_{\mathrm{e}} - I$ curve (already introduced in Fig. 2), corresponding to all possible numbers of decision levels (1, 2, 3, and 4, respectively), are shown. For a given number of decision levels, the system operational point moves from the position corresponding to $I = 0$ (for very low values of $\mathrm{SNR}_{\mathrm{sensor}}$) to the limiting position, which is asymptotically approached for $\mathrm{SNR}_{\mathrm{sensor}} \rightarrow \infty$. Therefore, the results in Fig. 4 show clearly the limitations, from a practical viewpoint, introduced by a hierarchical network architecture with decentralized detection and multi-level fusion.

IV. Concluding Remarks

In this letter, we have characterized, from an information-theoretic perspective, the behavior of clustered sensor networks with decentralized detection of a binary phenomenon in the presence of multi-level majority-like information fusion. By modelling the network as a BIBO system, it has been shown that its operational point lies on a single $P_{\mathrm{e}} - I$ curve, regardless of the network configuration. In particular, the position of the operational point depends on the network characteristics, namely, the phenomenon (its a priori probability distribution), clustering (uniform or non-uniform), the number of decision levels, and the (observation and communication) noise level. Therefore, our simple approach allows to directly compare different networking schemes.

References

[1] J. N. Tsitsiklis, *Advanced Statistical Signal Processing*. Greenwich, CT: JAI Press, 1993, vol. 2, ch. Decentralized Detection, pp. 297–344; editors: H. V. Poor and J. B. Thomas.

[2] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors–part I: fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[3] I. Y. Hoballah and P. K. Varshney, "An information theoretic approach to the distributed detection problem," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 988–994, Sep. 1989.

[4] K. Yamasaki and T. Ohtsuki, "Design of energy-efficient wireless sensor networks with censoring, on-off, and censoring and on-off sensors based on mutual information," in *Proc. IEEE Vehicular Tech. Conf. (VTC)*, vol. 2, Stockholm, Sweden, May 2005, pp. 1312–1316.

[5] G. Ferrari, M. Martalò, and R. Pagliari, "Decentralized detection in clustered sensor networks," *IEEE Trans. Aerosp. Electron. Syst.*, to appear, 2010. Available: http://www.tlc.unipr.it/martalo/privfold/preprint.pdf.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, Inc., 1991.