Data-Driven Strategies for selective data transmission in sensor networks

Giorgio Battistelli, Alessio Benavoli, and Luigi Chisci

***Abstract*— Energy efficiency is a crucial issue for any task involving wireless sensor networks. The present paper addresses nonlinear state estimation over a centralized sensor network, i.e. a set of sensor nodes communicating with a central information fusion unit, and proposes smart data-driven strategies by which sensors decide which data transmit to the central unit so as to reduce data communication, and thus avoid congestion problems as well as prolong the network lifetime, while providing enhanced performance with respect to periodic transmission. Both measurement and estimate transmission strategies are developed. To cope with nonlinear sensors that cannot fully observe the state, suitable nonlinear observability decompositions are employed. A bearing-only tracking simulation case-study is presented in order to demonstrate the effectiveness of the proposed approach.**

I. Introduction

The recent and rapid advances in WSN (Wireless Sensor Network) technology pose challenging estimation issues related to the possibility of optimally exploiting the distributed information provided by the WSN while preserving as much as possible the limited energy resources of the wireless sensor nodes and, thus, prolonging the network lifetime. Since, as well known, data communication represents by far, for a sensor node, the most energy-consuming task, the idea of controlling data transmission so as to achieve a trade-off between communication costs and estimation performance has recently received a certain attention in the literature (see [1], [2], [3], [4], [5] and the references therein). This can be done either in a centralized or in a distributed way. In the former setting, the fusion node selects only a subset of the available sensors to receive the information [6], [7]. Conversely, in a distributed setting, each sensor node decides whether or not the data should be transmitted only on the grounds of the locally available information.

Recent work [5] has concerned a network architecture, in which the sensor nodes - equipped with processing capabilities - transmit data (either raw measurements or computed estimates) to the central fusion node, providing estimation strategies that properly balance data processing in the WSN nodes and data communication from the sensor nodes to the central unit. The idea in [5] is to control transmission in sensor nodes by selectively transmitting data (measurements or estimates) whose distance, in a suitably weighted norm, from a properly defined prediction computed on the basis of information available to both the sensor node and the fusion unit exceeds a given threshold chosen according to the desired transmission rate. The data-driven selective transmission strategies proposed in [5] provide significant performance improvements with respect to periodic transmission operating at the same rate but are unable to cope with the presence of nonlinear sensors that cannot fully observe the target process (e.g. angle-only or range-only or Doppler-only position sensors employed for target localization/tracking). This paper extends the data-driven measurement and estimate transmission strategies to general nonlinear systems making use of nonlinear observability decompositions at the sensor level.

A. Problem Formulation

The present paper addresses estimation of the state of a discrete-time dynamical system

$$x_{k+1} = f(x_k) + w_k \tag{1}$$

given measurements collected from multiple sensors

$$y_k^i = h^i(x_k) + v_k^i\,, \quad i = 1, \ldots, s \tag{2}$$

under a limitation on the communication rate from each remote sensor unit to a central information fusion unit. Each remote sensor collects noisy measurements of the given system, can process them to find filtered estimates and transmits, at a reduced rate, either measurements or estimates to the fusion node. The fusion node, on the basis of the data received from the remote sensors, should provide, in the best possible way, an estimate of the system state.

In the foregoing, we formalize the concept of *communication strategy (CS)* with fixed rate α^i for sensor i. To this end, let us introduce for each sensor i binary variables c_k^i such that $c_k^i = 1$ if sensor i transmits at time k or $c_k^i = 0$ otherwise. Then, a decision mechanism with rate $\alpha^i \in (0,1)$ can be formally defined as any, deterministic or stochastic, mechanism of generating c_k^i such that

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} \mathbb{E}\left\{c_k^i\right\} \;=\; \alpha^i \tag{3}$$

where $\mathbb{E}\{\cdot\}$ denotes the expectation operator. This indicates that, for each sensor, the averaged number of data transmissions per time unit is constrained to take a value α^i. Such a constraint can be used to model all those practical situations in which the sensing units and the monitoring unit are remotely dislocated with respect to each other and the communication rate between them is severely limited.

Alessio Benavoli acknowledges financial support from the Swiss NSF grant n. 200020-137680/1.

G. Battistelli and L. Chisci are with Dipartimento di Sistemi e Informatica, DSI-Università di Firenze, Via Santa Marta 3, 50139 Firenze, Italy {battistelli, chisci}@dsi.unifi.it

A. Benavoli is with Istituto "Dalle Molle" di Studi sull'Intelligenza Artificiale, 6928 Manno-Lugano, Switzerland alessio@idsia.ch

Notations:: $\mathbb{Z}_+$ is the set of positive integers; given a square matrix M, $\mathrm{tr}(M)$ denotes its trace; given a positive definite matrix M, $\mathcal{E}_M \stackrel{\triangle}{=} \{\zeta : \zeta' M \zeta \leq 1\}$ denotes the ellipsoid centered at the origin associated with M; $\|\zeta\|_M \stackrel{\triangle}{=} (\zeta' M \zeta)^{1/2}$ is the weighted norm of the vector ζ; given matrices $M^1, \ldots, M^s$, $\mathrm{diag}(M^1, \ldots, M^s)$ denotes the block-diagonal matrix whose diagonal blocks are $M^1, \ldots, M^s$; while $\mathrm{col}(M^1, \ldots, M^s)$ is the matrix obtained by stacking $M^1, \ldots, M^s$ one on top of the other.

II. Data-driven strategies for measurement transmission

In the spirit of the results of [5], it is supposed that each sensor i transmits its local measurement y_k^i on the basis of a data-driven strategy of the type

$$c_k^i = \begin{cases} 0\,, & \text{if } \|y_k^i - \tilde{y}_k^i\|^2_{W_k^i} \leq \delta^i \\ 1\,, & \text{otherwise} \end{cases} \tag{4}$$

where the vectors $\tilde{y}_k^i, k \in \mathbb{Z}_+$, the positive definite weight matrices $W_k^i, k \in \mathbb{Z}_+$, and the positive reals δ^i have to be chosen so as to ensure that the transmission rate constraint (3) is satisfied.

Hereafter, the process noise w_k and the measurement noises v_k^i, $i = 1, \ldots, s$ in (1)-(2) are supposed to be zero-mean stochastic processes with $\mathbb{E}\{w_k w_t'\} = Q\,\delta_{kt}$, $\mathbb{E}\{v_k^i w_t'\} = 0$, and $\mathbb{E}\{v_k^i (v_t^j)'\} = R^i\,\delta_{ij}\,\delta_{kt}$ where δ_{ij} denotes the Kronecker delta.

A. Operations at fusion node

One of the advantages of a data-driven strategy over non data driven strategies (e.g., periodic) is that, even when no transmission is performed from the sensor node, the fusion node knows that the measurement y_k^i belongs to the ellipsoid $\delta^i \mathcal{E}^i_{W_k}$. Such information can be fruitfully exploited by suitably modifying the nonlinear filtering algorithm implemented in the fusion node. In fact, it has been shown in [5] that the case of no transmission from sensor i at time k can be treated as if a virtual measurement $z_k^i = \tilde{y}_k^i$ were generated by the measurement channel

$$z_k^i = h^i(x_k) + v_k^i - u_k^i \tag{5}$$

where u_k^i is uniformly distributed in the ellipsoid $\delta^i \mathcal{E}^i_{W_k}$ and uncorrelated with v_k^i. Note that the $v_k^i - u_k^i$ has zero-mean and covariance matrix[1]

$$\begin{aligned} \mathbb{E}\{(v_k^i - u_k^i)(v_k^i - u_k^i)'\} &= \mathbb{E}\{v_k^i (v_k^i)'\} + \mathbb{E}\{u_k^i (u_k^i)'\} \\ &= R^i + \frac{\delta^i}{m_i + 2} (W_k^i)^{-1}, \end{aligned}$$

where $m_i \stackrel{\triangle}{=} \dim(y_k^i)$. As discussed in [5], this is possible from a Bayesian point of view in that, regardless of the choice of $\tilde{y}_k^i$, the posterior PDF of x_k conditioned to the fact that no data have been received (i.e., $c_k^i = 0$) coincides with the posterior PDF of x_k conditioned to the fact that a measurement $z_k^i = \tilde{y}_k^i$ has been originated from a virtual measurement channel (5).

In view of the foregoing considerations, by defining

$$\begin{aligned} z_k^i &\stackrel{\triangle}{=} c_k^i y_k^i + (1 - c_k^i)\tilde{y}_k^i \\ R_k^i &\stackrel{\triangle}{=} R^i + (1 - c_k^i)\frac{\delta^i}{m_i + 2}(W_k^i)^{-1} \end{aligned} \tag{6}$$

and, accordingly,

$$\begin{aligned} z_k &\stackrel{\triangle}{=} \mathrm{col}\left(z_k^1, \ldots, z_k^s\right), \\ R_k &\stackrel{\triangle}{=} \mathrm{diag}\left(R_k^1, \ldots, R_k^s\right), \\ h(\cdot) &\stackrel{\triangle}{=} \mathrm{col}\left(h^1(\cdot), \ldots, h^s(\cdot)\right), \end{aligned}$$

the estimate $\hat{x}_{k|k}$ at the fusion node can be computed by means of the following recursion

$$\begin{aligned} \left(\hat{x}_{k|k}, P_{k|k}\right) &= \mathrm{upd}\left(\hat{x}_{k|k-1}, P_{k|k-1}, z_k, R_k, h\right), \\ \left(\hat{x}_{k+1|k}, P_{k+1|k}\right) &= \mathrm{pred}\left(\hat{x}_{k|k}, P_{k|k}, Q, f\right). \end{aligned} \tag{7}$$

Here and in the following, $\mathrm{upd}(\cdot)$ denotes the updating step of some nonlinear filtering algorithm; its first and second arguments are the predicted estimate and covariance, its third argument is the measurement vector, its fourth argument is the covariance of the additive measurement noise, and its last argument is the (non)linear measurement function. Similarly, $\mathrm{pred}(\cdot)$ denotes the prediction step of the nonlinear filtering algorithm; its first two arguments are the updated estimate and covariance, its third argument is the covariance of the additive process noise, and the last argument is the (non)linear function which relates the state at time k with the state at time $k+1$.

Notice that the filtering steps $\mathrm{pred}(\cdot)$ and $\mathrm{upd}(\cdot)$ can be performed via any nonlinear filter such as for instance the Extended Kalman filter, the Unscented Kalman filter (UKF), or a particle filter [8], [9]. UKF represents a good tradeoff between accuracy and computational cost and, for this reason, will be adopted in the subsequent developments.

B. Operations at sensor node

As should be evident, the performance of transmission strategy (4) may be significantly affected by the specific mechanism used for generating the quantities $\tilde{y}_k^i$ and W_k^i. In this connection, in view of the discussion of Section 2.1 of [5], a sensible choice can be obtained by tailoring the vector $\tilde{y}_k^i$ and the weight matrix W_k^i to the prior distribution of the measurement y_k^i so that $\tilde{y}_k^i$ coincides with the predicted value of y_k^i and W_k^i is proportional to the inverse innovation covariance. In this way, supposing that the PDF of y_k^i is approximately characterized by a radial symmetry, the volume of the non-transmission region is close to being minimal (see Proposition 1 of [5]). Further, since the vector $\tilde{y}_k^i$ is used in the fusion node in order to define the vector z_k and compute the estimate $\hat{x}_{k|k}$, the mechanism for generating the vectors $\tilde{y}_k^i$ and the weight matrices W_k^i must use only information available to both the sensor i and the fusion node. A last important requirement that should be fulfilled

[1] Recall that a uniform random variable taking value in an ellipsoid $\mathcal{E}_M \subset \mathbb{R}^m$ has covariance matrix equal to $[(m+2)M]^{-1}$.

is that the volume of the non-transmission region does not grow unbounded with time. In fact, in a linear setting, it has been shown that this is a sufficient condition for preserving the stability of the estimation error at the fusion node for any positive transmission rate [5]. Summing up, taking into account such desired properties, a reasonable choice is

$$\tilde{y}_k^i = \hat{y}_{k|k-1}^i\,, \quad (W_k^i)^{-1} = \frac{1}{\operatorname{tr}(S_k^i)}\, S_k^i \tag{8}$$

where $\hat{y}_{k|k-1}^i$ is a prediction of y_k^i based on the information shared by the fusion node and sensor i up to time $t-1$ and S_k^i is the corresponding covariance. The normalization factor $\operatorname{tr}(S_k^i)$ is a way for ensuring that the volume of the non-transmission region is always bounded in that $\frac{1}{\operatorname{tr}(S_k^i)}\, S_k^i \leq I$. Of course other choices for the normalization factor can be devised (e.g., the spectral radius of S_k^i). The main reason for preferring the trace is that it can be computed with little computational effort from the remote sensors.

The main difficulty in computing the prediction $\hat{y}_{k|k-1}^i$ is that in general, even assuming collective observability from the whole set of sensors, the state vector need not be observable from a single sensor. This means that, locally in a remote sensor, it may not be possible to construct a stable filter for estimating the whole state of system (1). Nevertheless, it may still be possible to compute the prediction $\hat{y}_{k|k-1}^i$ by resorting to an observability decomposition of the state space.

To this end, consider the noise-free system

$$x_{k+1} = f(x_k) \tag{9}$$
$$y_k^i = h^i(x_k) \tag{10}$$

and consider the set of functions

$$\Theta \stackrel{\triangle}{=} \{h^i(x),\, h^i \circ f^N(x),\; N = 1, 2 \ldots\}.$$

where f^N denotes the N-th composition of f, i.e. the function f composed with itself N times. As well known, under suitable assumptions, observability of system (9) from the output y_k^i corresponds to the fact that $\dim(\operatorname{span} d\Theta) = n$ almost everywhere, where $d\Theta$ is the differential of Θ and n is the system order. If, instead, $\operatorname{span} d\Theta$ has constant dimension $n^{i,o} < n$ in the set of interest then observability does not hold, but it might be possible to perform a state space decomposition with respect to the system observability properties. Hereafter we shall assume that, for all the sensors for which observability does not hold, such a decomposition exists or, more precisely, that there exists a change of coordinates $q = T^i(x)$ such that system (9) can be rewritten as

$$q_{k+1}^{i,o} = \tilde{f}^{i,o}(q_k^{i,o}) \tag{11}$$
$$q_{k+1}^{i,\bar{o}} = \tilde{f}^{i,\bar{o}}(q_k^{i,o}, q_k^{i,\bar{o}}) \tag{12}$$
$$y_k^i = \tilde{h}^i(q_k^{i,o}) \tag{13}$$

where $q^{i,o}$ has dimension $n^{i,o}$, $q^{i,\bar{o}}$ has dimension $n - n^{i,o}$, $q = \operatorname{col}\left(q^{i,o}, q^{i,\bar{o}}\right) = \operatorname{col}\left(T^{i,o}(x), T^{i,\bar{o}}(x)\right)$, and the subsystem made up by (11) and (13) is observable by construction. Notice that, in general, the observability decomposition will be different for different sensors, hence the dependence on the index i. Discussions on the existence of such a decomposition in a discrete-time setting can be found, for instance, in [10], [11]. As evident from (11)-(13), in the noise-free case, the output y_k^i depends only on the observable part of the system. Thus one can exploit this fact and construct a filter for the observable subsystem in order to obtain a prediction $\hat{y}_{k|k-1}^i$ of the measurement y_k^i.

In practice, turning back our attention to the noisy case (9), each sensor will consider only the subsystem

$$q_{k+1}^{i,o} = \tilde{f}^{i,o}(q_k^{i,o}) + g^i(q_k^{i,o}, q_k^{i,\bar{o}}, w_k) \tag{14}$$
$$y_k^i = \tilde{h}^i(q_k^{i,o}) + v_k^i \tag{15}$$

where the function $g(\cdot)$ has the property that $g(\cdot,\cdot,0) = 0$. While, in general, in the presence of noises there is not a clear decoupling between observable and nonobservable states, several techniques can all the same be devised for constructing a filter for (14)-(15) whenever $q_k^{i,\bar{o}}$ belongs to a (known) compact set, which is usually the case when dealing with state estimation for physical systems. For example, the simplest alternative consists in treating the term $g^i(q_k^{i,o}, q_k^{i,\bar{o}}, w_k)$ as an additive noise vector. Alternatively, a more accurate approach amounts to considering the vector $q_k^{i,\bar{o}}$ as a slowly time-varying unknown parameter and resorting to robustness arguments. Notice that when the functions $f(\cdot)$ and $h^i(\cdot)$ are smooth then the function $g^i(\cdot)$ will be smooth as well, thus the contribution of the second term in the right hand side of (14) will in general be small for small disturbances w_k.

Taking into account subsystem (14)-(15), the operations for computing the prediction $\hat{y}_{k|k-1}^i$ in node i are as follows

$$\begin{aligned} &\left(\hat{q}_{k|k}^{i,o}, P_{k|k}^i\right) = \operatorname{upd}(\hat{q}_{k|k-1}^{i,o}, P_{k|k-1}^i, z_k^i, R_k^i, \tilde{h}^i) \\ &\left(\hat{q}_{k+1|k}^{i,o}, P_{k+1|k}^i\right) = \operatorname{pred}(\hat{q}_{k|k}^{i,o}, P_{k|k}^i, Q, \tilde{f}^{o,i}, g^i) \\ &\left(\hat{y}_{k+1|k}^i, S_k^i\right) = \operatorname{proj}(\hat{q}_{k+1|k}^{i,o}, P_{k+1|k}^i, \tilde{h}^i) \end{aligned} \tag{16}$$

Here $\operatorname{proj}(\cdot)$ denotes the function that projects the predicted estimate and relative covariance on the measurement space; its first two arguments are the predicted estimate and covariance and its last argument is the (non)linear function that projects the observable state in the measurement space. Such a projection can, for example, be computed by means of the unscented transform [8]. Notice that, with a little abuse of notation, we have augmented the arguments of $\operatorname{pred}(\cdot)$ by including also the (non)linear function g^i. Notice also that the proposed filter does not use the true measurement y_k^i and covariance R but, instead, uses the virtual measurement z_k^i and its covariance R_k^i as defined in (6). This choice is motivated by the fact that, as discussed in the previous section, the filter (16) should be also runned in the fusion node in order to compute the virtual measurement z_k and, consequently, the fused estimate $\hat{x}_{k|k}$.

C. An example: bearing only tracking

In this section, the case of bearing only tracking is analyzed in order to show how the foregoing discussion applies to a context of practical interest. The problem is that of estimating the state of an object whose motion is described by the kinematic nearly-constant velocity model

$$x_{k+1} = \begin{bmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k \; + \; w_k,$$

where Δ is the sampling interval; the unknown state vector is given by the position and velocity components along the coordinate axes (ξ, η), i.e., $x = \mathrm{col}(\xi, \eta, \dot{\xi}, \dot{\eta})$. The state has to be estimated given measurements collected from a network of direction-of-arrival (DOA) sensors

$$y_k^i = \mathrm{atan2}\left(\eta_k - \eta^i, \xi_k - \xi^i\right) + v_k^i\,, \quad i = 1, \ldots, s \qquad (17)$$

where $\mathrm{atan2}(\cdot)$ denotes the four-quadrant inverse tangent and (ξ^i, η^i) is the position of the i-th sensor.

As well known, it is not possible to observe the whole state from a single angular sensor. To see this, note that the N-th composition of $f(\cdot)$ takes the form

$$f^N(x) = \mathrm{col}\left(\xi + N\Delta v_x, \eta + N\Delta v_y, v_x, v_y\right)$$

and, consequently,

$$\begin{aligned} d\left(h^i \circ f^N\right) &= \frac{1}{(\xi^{N,i})^2 + (\eta^{N,i})^2} \\ &\times \left(-\eta^{N,i} d\xi + \xi^{N,i} d\eta - N\Delta\eta^{N,i} d\dot{\xi} + N\Delta\xi^{N,i} d\dot{\eta}\right) \end{aligned}$$

where

$$\xi^{N,i} \triangleq \xi + N\Delta v_x - \xi^i\,, \quad \eta^{N,i} \triangleq \eta + N\Delta v_y - \eta^i\,.$$

It is an easy matter to see that dh^i, $d(h^i \circ f)$, and $d(h^i \circ f^2)$ are mutually independent but, for any $N \geq 3$, one can write

$$d(h \circ f^N) = \beta_1^{N,i}(x) dh^i + \beta_2^{N,i}(x) d(h^i \circ f) + \beta_3^{N,i}(x) d(h^i \circ f^2)$$

where

$$\begin{aligned} \beta_1^{N,i}(x) &\triangleq \frac{\xi^{0,i} + \eta^{0,i}}{\xi^{N,i} + \eta^{N,i}} \left(1 + \frac{N^2 - 3N}{2}\right) \\ \beta_2^{N,i}(x) &\triangleq \frac{\xi^{1,i} + \eta^{1,i}}{\xi^{N,i} + \eta^{N,i}} \left(2N - N^2\right) \\ \beta_3^{N,i}(x) &\triangleq \frac{\xi^{2,i} + \eta^{2,i}}{\xi^{N,i} + \eta^{N,i}} \, \frac{N^2 - N}{2}\,. \end{aligned}$$

Hence, the dimension of $\mathrm{span}(d\Theta)$ is 3. As suggested in [10], in general the differential $d\Theta$ can be a useful starting point for constructing the observability decomposition. However, in this case, this is not necessary since the decomposition becomes apparent when the system is rewritten in terms of modified polar (MP) coordinates [9], [12], [13]. The MP state vector is defined as

$$q^i = \mathrm{col}\left(\dot{\theta}^i,\, \frac{\dot{r}^i}{r^i},\, \theta^i,\, \frac{1}{r^i}\right)\,. \qquad (18)$$

where θ^i denotes the bearing-angle measured from the ξ-axis counterclockwise, $\dot{\theta}^i$ is the angular velocity, r^i is the range from sensor position, and $\dot{r}^i$ the range-rate. Figure 1 shows the target-sensor geometry for the MP coordinates system.

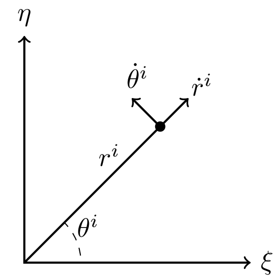

Fig. 1. MP Coordinate System sensor-target geometry

The Cartesian state vector x_k and the MP state vector q_k^i are related at all times by the following nonlinear one-to-one transformation

$$q^i = \begin{bmatrix} \dfrac{(\xi - \xi^i)\dot{\eta} - (\eta - \eta^i)\dot{\xi}}{(\xi - \xi^i)^2 + (\eta - \eta^i)^2} \\ \dfrac{(\xi - \xi^i)\dot{\xi} + (\eta - \eta^i)\dot{\eta}}{(\xi - \xi^i)^2 + (\eta - \eta^i)^2} \\ \mathrm{atan2}\left(\eta - \eta^i, \xi - \xi^i\right) \\ 1/\sqrt{(\xi - \xi^i)^2 + (\eta - \eta^i)^2} \end{bmatrix}\,. \qquad (19)$$

It can be shown that only the first three components of the vector q^i are observable from a single non-moving DOA sensor [13]. In fact, by defining $q^{i,o} \triangleq \mathrm{col}(\dot{\theta}^i,\, \dot{r}^i/r^i,\, \theta^i)$ and $q^{i,\bar{o}} \triangleq 1/r^i$, the dynamics of $q^{i,o}$ can be written as in (14) and, of course, the measurement equations take the form

$$y_k^i = C q^{i,o} + v_k^i\,.$$

where $C \triangleq [0 \quad 0 \quad 1]$. Further, in this particular case, the function $g(\cdot)$ in (14) takes the form $g^i(q^{i,\bar{o}}, q^{i,\bar{o}} w_k)$. Thus, taking into account the fact that all bearing sensors have a limited and known range for detecting the presence of an object, the quantity $q^{i,\bar{o}} w_k$ can be treated as a disturbance with unknown but bounded covariance.

Notice that, notwithstanding the lack of observability from a single sensor, the target state can be reconstructed at the fusion node provided that the number of sensors s that transmit their measurements is greater than or equal 2. Actually two sensors may not be sufficient for observability if target and sensors are aligned.

III. Data-driven strategies for estimate transmission

Let us now consider the case of estimate transmission, supposing that each sensor i transmits, instead of its local measurement y_k^i, a local estimate. Taking into account the development of Section II-B, only the component of the state that is locally observable is estimated and sent to the fusion node. This means that each sensor computes an estimate $\hat{q}_{k|k}^{i,o}$ of the state $q_k^{i,o}$ of the observable subsystem by means of the nonlinear filtering recursion

$$\begin{array}{l}\left(\hat{q}^{i,o}_{k|k}, P^i_{k|k}\right) = \mathrm{upd}(\hat{q}^{i,o}_{k|k-1}, P^i_{k|k-1}, y^i_k, R^i, \tilde{h}^i) \\ \left(\hat{q}^{i,o}_{k+1|k}, P^i_{k+1|k}\right) = \mathrm{pred}(\hat{q}^{i,o}_{k|k}, P^i_{k|k}, Q, \tilde{f}^{o,i}, g^i)\end{array} \tag{20}$$

Comparing such a recursion with (20), one can see that here the true measurement y^i_k is considered instead of the virtual measurement z^i_k. This means that the estimate $\hat{q}^{i,o}_{k|k}$ is computed using all the locally available information. Then, the local estimate is transmitted according to a data-driven strategy of the form

$$c^i_k = \begin{cases} 0\,, & \text{if } \|\hat{q}^{i,o}_{k|k} - \tilde{q}^{i,o}_k\|^2_{W^i_k} \le \delta^i, \\ 1\,, & \text{otherwise,}\end{cases} \tag{21}$$

where the vectors $\tilde{q}^{i,o}_k, k \in \mathbb{Z}_+$, the positive definite weight matrices $W^i_k, k \in \mathbb{Z}_+$, and the positive reals δ^i have to be chosen so as to ensure that the transmission rate constraint (3) is satisfied.

Also in this case, the general principles discussed at the beginning of Section II-B can be used as a guideline for choosing the vectors $\tilde{q}^{i,o}_k, k \in \mathbb{Z}_+$ and the positive definite weight matrices $W^i_k, k \in \mathbb{Z}_+$. Accordingly, the idea is to set $\tilde{q}^{i,o}_k$ equal to the best prediction of $q^{i,o}_k$ that can be computed on the basis of the information that is common to both sensor i and the fusion node up to time $k-1$. In particular, if one denotes by $n^i_k \ge 0$ the number of time instants elapsed from the last transmission of sensor i, i.e., n^i_k is such that $c^i_{k-n^i_k} = 1$ and $c^i_{k-1} = \cdots = c^i_{k-n^i_k+1} = 0$, then a natural choice corresponds to setting $\tilde{q}^{i,o}_k = \hat{q}^{i,o}_{k|k-n^i_k}$. Such a prediction can be recursively computed as

$$\left(\hat{q}^{i,o}_{k|k-n^i_k}, P^i_{k|k-n^i_k}\right) = \begin{cases} \mathrm{pred}(\hat{q}^{i,o}_{k-1|k-n^i_{k-1}}, P^i_{k-1|k-n^i_{k-1}}, Q, \tilde{f}^{o,i}, g^i) & \text{if } c^i_{k-1} = 0 \\ \hat{q}^{i,o}_{k|k-1}, P^i_{k|k-1} & \text{if } c^i_{k-1} = 1\end{cases}$$

As for the weight matrix W^i_k, one can compute, for example through the unscented transform, an approximation, say Λ^i_k, of the covariance of $\hat{q}^{i,o}_{k|k} - \hat{q}^{i,o}_{k|k-n^i_k}$ and then set

$$(W^i_k)^{-1} = \frac{1}{\mathrm{tr}(\Lambda^i_k)}\, \Lambda^i_k$$

where, as in (8), the normalization is a means for ensuring that the volume of the non-transmission region remains uniformly bounded.

As for the derivation of the fused estimate $\hat{x}_{k|k}$ in the fusion node, we adopt an approach based on the idea of interpreting each local estimate $\hat{q}^{i,o}_{k|k}$ as a measurement z^i_k of the true state x_k collected through the (virtual) measurement channel

$$z^i_k = T^{i,o}(x_k) + \epsilon^i_k \tag{22}$$

where, the estimation error $\epsilon^i_k \triangleq \hat{q}^{i,o}_{k|k} - T^{i,o}(x_k)$ plays the role of a (virtual) measurement noise with covariance equal to the estimation error covariance $P^i_{k|k}$. While, equation (22) can be directly applied only for the indices i for which an estimate has been received (i.e., for which $c^i_k = 1$), one can once again exploit the fact that, for a data-driven transmission strategy, even in case of no transmission it is known that the estimate $\hat{x}^i_{k|k}$ belongs to a certain ellipsoid $\delta^i \mathcal{E}_{W_k}$. Thus, ne can treat the case of no transmission by replacing (22) with

$$z^i_k = T^{i,o}(x_k) + \epsilon^i_k - u^i_k \tag{23}$$

where $z^i_k \triangleq \tilde{q}^{i,o}_k$ and u^i_k is uniformly distributed in the ellipsoid $\delta^i \mathcal{E}_{W_k}$. Summing up, by defining

$$\begin{aligned} z^i_k &\triangleq c^i_k \hat{q}^{i,o}_{k|k} + (1-c^i_k)\tilde{q}^{i,o}_k,\ i = 1,\dots,s\,, \\ R^i_k &\triangleq P^i_{k|k} + (1-c^i_k)\frac{\delta^i}{n^{i,o}+2}(W^i_k)^{-1},\ i = 1,\dots,s\,, \\ z_k &\triangleq \mathrm{col}\left(z^1_k,\dots,z^s_k\right)\,, \\ R_k &\triangleq \mathrm{diag}\left(R^1_k,\dots,R^s_k\right), \\ h(\cdot) &\triangleq \mathrm{col}\left(T^{1,o}(\cdot),\dots,T^{s,o}(\cdot)\right)\,, \end{aligned}$$

the estimate $\hat{x}_{k|k}$ at the fusion node can be computed by means of a recursion which is formally analogous to (7).

Finally, observe that, because of the common process noise, the estimates $\hat{q}^{i,o}_{k|k}$ are actually dependent. The naive independence assumption has been adopted since standard techniques for fusing correlated information (e.g., covariance intersection) cannot be applied to the nonlinear and non-fully observable case. The development of alternative fusion rules (possibly based on the moving horizon estimation paradigm [14]) will be subject of future research.

IV. Simulation results

The goal of this section is to evaluate the performance of the proposed transmission strategies when applied to the network of bearing sensors described in section II-C. The covariance matrix of process noise has been assumed equal to $Q = \zeta Q_0$ with

$$Q_0 = \begin{bmatrix} \Delta^3/3 & 0 & \Delta^2/2 & 0 \\ 0 & \Delta^3/3 & 0 & \Delta^2/2 \\ \Delta^2/2 & \Delta & 0 & 0 \\ 0 & 0 & \Delta^2/2 & \Delta \end{bmatrix},$$

$\zeta > 0$ and sampling interval $\Delta = 0.1\,s$. The following transmission strategies have been compared:

- **Data-Driven Measurement Transmission (DDMT)**: the strategy of Section II; the thresholds δ^i have been tuned so as to obtain the desired transmission rates α^i.
- **Periodic Measurement Transmission (PMT)**: sensors transmit periodically their measurements once every $1/\alpha^i$ time instants, the phase shift among the sensors being random.
- **Data-Driven Estimate Transmission (DDET)**: the strategy of Section III; the thresholds δ^i have been tuned so as to obtain the desired transmission rates α^i.
- **Periodic Estimate Transmission (PET)**: sensors transmit periodically their estimates once every $1/\alpha^i$ time instants, the phase shift among the sensors being random.

The comparison among the four strategies has been carried out via Monte Carlo simulations with independent runs obtained for random generated trajectories by varying the measurement and process noises realizations. The *time averaged square error* (TASE) at the fusion node F, $TASE = \frac{1}{T}\sum_{k=1}^{T}(x_k - \hat{x}_{k|k})'(x_k - \hat{x}_{k|k})$, has been computed for each simulation run and for each strategy. Then, the mean and the maximum of the TASE over the Monte Carlo runs have been considered as performance indices. The simulation time T and the number of Monte Carlo runs have been chosen equal to 600 and 1000, respectively.

Consider a network with eight sensors located at $(-16.67 + 16.67i, 0)$ for $i = 0, 1, \ldots, 7$. A target moves in the first quadrant and the target's trajectory is generated so as to guarantee observability at each time instant from the overall sensor network. In all simulations the transmission rate α^i has been fixed to 0.1 for all sensors and transmission strategies. The target initial state is $[60, 60, 0.1, 0.1]'$, the variance of the measurement noise $R = 3\ 10^{-4}\ rad$ (1° standard deviation), the variance of the process noise $\zeta = 10^{-3}$. The simulation results are shown in Table I. It can be noticed that transmitting the estimates is better than transmitting the measurements and data-driven strategies are better than the periodic ones. These results are in agreement with those obtained in [5] by the same strategies in the linear case. The relative percentage of TASE reduction of data-driven w.r.t. periodic strategies for both cases, transmitting the estimates and transmitting the measurements, have also been reported in Table I to help in the comparison.

A difference between the linear case treated in [5] and bearing-only tracking is that the transmission rate depends on the target trajectory. In particular, we have noticed that for the data-driven strategies (especially for DDET) the transmission rate increases at the decrease of the distance between target and sensor. In other words, sensors transmit more frequently when the distance target-sensor decreases. This means that data-driven strategies perform a natural sensor-scheduling in the nonlinear case. Sensors that are closer to the target (and, thus, sensors that have a more accurate estimate of the state) transmit more frequently. This is an interesting feature of data-driven strategies that we intend to investigate in future work. Notice that, for a fair comparison between the strategies with a constant transmission rate (the periodic ones) and the data-driven strategies, all simulations have been performed by generating trajectories that ensure that the average transmission rate α^i is equal to the fixed value 0.1 for all sensors and transmission strategies (i.e., the target is always "far enough" from the sensors).

	TASE	PMT	DMT	PET	DDET
$s = 2$	avg	0.8532	0.2660 (68%)	0.4483	0.1997 (55%)
	max	3.6511	1.0250 (71%)	3.3815	0.7300 (79%)
$s = 4$	avg	0.5864	0.1819 (68%)	0.4213	0.1500 (63%)
	max	2.6720	0.6188 (76%)	2.9492	0.6238 (78%)
$s = 6$	avg	0.4367	0.1364 (68%)	0.4081	0.1260 (69%)
	max	1.5589	0.5010 (67%)	3.1153	0.6301 (78%)
$s = 8$	avg	0.3825	0.1065 (72%)	0.3633	0.1011 (72%)
	max	1.3968	0.3510 (74%)	2.4125	0.4182 (82%)

TABLE I

CASE 2: MEDIUM PROCESS NOISE VARIANCE

V. CONCLUSIONS

The paper has provided a contribution towards energy-efficient management of a centralized wireless sensor network employed for estimating the state of a dynamical system. An interesting solution to this problem has been provided in [5] but this solution is not applicable in certain situations of practical interest, like e.g. angle-only or rangle-only or Doppler-only target tracking, wherein there are nonlinear sensors that cannot fully observe the system. This paper has filled this practical gap, providing novel data-driven transmission strategies that can be used for general collectively observable (i.e. observable from the whole set of sensors but non necessarily from a single one) nonlinear systems. Future work along this line will concern: (1) centralized fusion techniques which do not rely on the independence assumption among local estimates, and (2) nonlinear observability decompositions for other types of sensors (e.g., Doppler-only).

REFERENCES

[1] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. of IEEE, Special Issue on Technology of Networked Control Systems*, vol. 95, pp. 138–162, 2007.

[2] Y. S. Suh, V. H. Nguyen, and Y. S. Ro, "Modified Kalman filter for networked monitoring systems employing a send-on-delta method," *Automatica*, vol. 43, pp. 332–338, 2007.

[3] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "Stochastic sensor scheduling for energy constrained estimation in multi-hop wireless sensor networks," *IEEE Transactions on Automatic Control*, vol. 56, pp. 2489 –2495, oct. 2011.

[4] G. Battistelli, A. Benavoli, and L. Chisci, "State estimation with remote sensors and intermittent transmissions," *Systems & Control Letters*, vol. 61, no. 1, pp. 155 – 164, 2012.

[5] G. Battistelli, A. Benavoli, and L. Chisci, "Data-driven communication for state estimation with sensor networks," *Automatica*, vol. 48, no. 5, pp. 926–935, 2012.

[6] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, pp. 251–260, 2006.

[7] L. Zuo, R. Niu, and P. Varshney, "Posterior CRLB based sensor selection for target tracking in sensor networks," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 2, pp. II–1041 –II–1044, april 2007.

[8] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, pp. 477 –482, mar 2000.

[9] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter: particle filters for tracking applications*. Artech House Publishers, 2004.

[10] F. Albertini and D. D'Alessandro, "Remarks on the observability of nonlinear discrete time systems," in *System Modeling and Optimization* (J. D. et al., ed.), London: Chapman and Hall, 1996.

[11] U. Kotta and K. Schlacher, "Possible non-integrability of observable space for discrete-time nonlinear control systems," in *Proceedings of the 17th IFAC World Congress*, pp. 9852–9856, 2008.

[12] V. Aidala and S. Hammel, "Utilization of modified polar coordinates for bearings-only tracking," *IEEE Transactions on Automatic Control*, vol. 28, pp. 283 – 294, mar 1983.

[13] A. Farina, "Target tracking with bearings only measurements," *Signal Processing*, vol. 78, no. 1, pp. 61 – 78, 1999.

[14] A. Alessandri, M. Baglietto, and G. Battistelli, "Moving-horizon state estimation for nonlinear discrete-time systems: new stability results and approximation schemes," *Automatica*, vol. 44, no. 7, pp. 1753 – 1765, 2008.