2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

On Myopic Strategies For Resource Constrained Informative Sampling

Hamid Emadi[1], Sourabh Bhattacharya[1]

Abstract— **In this work, we investigate data-driven sensor scheduling policies under resource constraints. We consider a scenario in which data is being generated by n discrete-time dynamical processes. At each instant, $m < n$ measurements can be collected. Based on an auto-regressive (AR) prediction model, we define two loss functions, namely, the prediction error and the measurement error associated with the missing measurements. The problem of choosing the measurements at each step is formulated as an optimization problem over the two loss functions. For $n = 2$ and $m = 1$, we explicitly derive the optimal measurement policy, and present an analytical expression for the decision rule. We extend the technique adopted in the aforementioned problem to derive an algorithmic procedure to address the case when m and n are arbitrary. We present simulations to illustrate the performance of the proposed algorithm, and compare it to the case when the parameters of the AR predictor are updated at each step based on the measurements.**

I. Introduction

Modern sensor networks are plagued with the problem of *data deluge* [1]. In order to circumvent the aforementioned challenge, techniques are being developed to enable sensors to capture the salient features of the underlying spatio-temporal phenomenon. Techniques from different disciplines, for example, sequential experimental design [2], active sensing [3], adaptive sensing/sampling [4], sensor scheduling [5] or controlled sensing [6] enable such smart data collection procedures. In this paper, we investigate the problem of minimizing the overall prediction error for several time-series data when there is a constraint on the maximum number of measurements that can be collected at each step. Such problems often arise in mobile sensor networks when the number of sensors are not sufficient to cover the entire observation field.

Our paper is closely related to the area of *controlled sensing*. In controlled sensing, the final goal is to develop a decision policy to take observations in order to infer accurately the state of the system or forecast the future state of the system [6]. In contrast to classical control systems, where the control primarily affects the states of the system, in controlled sensing, the control affects only the observations. In case of a mobile sensor network, controlled sensing can be used to develop sensor routing policies. For a static sensor network, it can be used to develop sensor activation and scheduling policies. Sensor scheduling arises when multiple sensors that provide information with different relevance/quality and observation costs are present or when a single sensor can be operated in multiple modes. Sensor scheduling is used in several applications, for example, surveillance [7], target tracking [8], coverage [9], to name a few. An extensive survey of several sensor scheduling policies is provided in [10].

In the past, there has been some work that addresses scheduling in linear spatially distributed systems [11]. Solving these problems leads to a deterministic optimal control because the evolution of the covariance is deterministic. Consequently, the evolution of information metric, which typically depends on the covariance, given a sequence of controls is also deterministic. These problems can be solved on-line by using Model Predictive Control (MPC) [11]. However in nonlinear dynamical systems, MPC is not very efficient because it results in a Stochastic Dynamic Programming (SDP) which is computationally expensive.

The optimal sensor scheduling problem can be considered as a Partially Observed Markov Decision Process (POMDP) [12]. Such problems are difficult to solve in general, and solution techniques are efficient for small input sizes. In [13], authors consider a sensor scheduling problem as a POMDP on the Information State (I-state), and provide an on-line receding horizon approach to solution of POMDP. This method results in reducing the variance of the gradient estimates. However this method leads to suboptimal solution because of the non-convex structure. [14] modifies an on-line receding horizon approach to the case of multiple sensing. In [9], authors propose an algorithm to estimate a process, based on the idea of optimizing the expected steady state performance when sensors switch randomly according to some optimal probability distribution. In [9], [7], selection of sensors are based on transition probabilities described by Markov process. In [15], authors consider a sequential estimation problem with two agents, one of which is an estimator and the other one is an observer. They propose a scheduling policy for coordination between these agents to minimize the estimation error. In [2], authors consider the problem of sensor selection for binary sequential hypothesis testing with multiple sensors in which at every time step, one out of multiple sensors is selected to take an observation. In [8], the authors examine the problem of target tracking subject to limited system resources from a sensor scheduling perspective.

The observation control policy can either be open loop [16], [17], [18], which is fixed and determined a priori, or closed loop policy [19], [20], [6], [21], [5], in which case the previous observations affect the current decision. In this paper, our focus is on temporal observation control in which the control at each time is closed loop.

In general, the notion of optimality is asymptotic in sensor

[1]Department of Mechanical Engineering Iowa State University, Ames, IA {emadi,sbhattac}@iastate.edu

scheduling problems, i.e, the number of observations are infinite almost surely [5], [20], [21]. In [5], authors pose an online sensor selection problem as an infinite horizon dynamic program, and provide non-asymptotic optimal strategy when the average number of observations is finite. In contrast, we consider a myopic/greedy strategy for sensor selection in which a one-step look-ahead policy is adopted to minimize the loss function. This leads to computationally feasible sensor scheduling strategies which can be implemented in practical scenarios involving autonomous mobile sensor platforms which have severe constraints in terms of the communication and computation.

In this work, we consider two different objective functions that model the error in measurement and prediction due to missing measurements. An intrinsic feature of time series is that adjacent observations are dependent. Therefore, loss functions can be computed in terms of the parameters of the time series, and previous observations. Consequently, the set of sensors that need to be activated can be determined efficiently. We show that our approach is computationally efficient. The organization of the paper is as follows. In Section II, we present the problem formulation. In section III, we present the optimal policy for the loss functions, and derive an explicit decision rule for two time series. In section IV, we simulate the proposed approach. Finally, Section V presents our conclusion and avenues for future research.

II. Problem Statement

Consider a network of n static sensors $\mathcal{S} = \{S_1, \ldots, S_n\}$ connected to a fusion center $\mathcal{F}$. Each sensor is capable of measuring samples generated by a discrete-time exogenous process. Let $\mathcal{Y}_i = \{Y_{i,k}\}$ denote the discrete-time random process that models the exogenous process generating the data at sensor S_i, where $Y_{i,k}$ is a random variable. Let $y_{i,k}$ denote the realization of the random variable $Y_{i,k}$. We assume that $\mathcal{Y}_i$ is a weakly stationary process with mean μ_i, and the jth autocovariance $\gamma_{i,j}$. Let $u_{i,k} \in \{0, 1\}$ represent the possible actions of S_i at step k. If $u_{i,k} = 1$, then S_i takes a measurement (i.e., $y_{i,k}$), and transmits it to $\mathcal{F}$. If $u_{i,k} = 0$, then S_i does not take any measurement, and hence there is no transmission to $\mathcal{F}$.

We assume that $\mathcal{F}$ uses a predictor for the following purposes: (i) To substitute for the missing data from the sensors (ii) To predict for the future value of the samples. In this paper, we investigate the scenario when the predictor is a p^{th} autoregressive model (denoted as $AR(p)$), and satisfies the following equation:

$$\begin{aligned} \hat{Y}_{i,k} &= c_i + \phi_{i,1} Y_{i,k-1} + \cdots + \phi_{i,p} Y_{i,k-p} + \epsilon_{i,k}, \quad (1) \\ &= \boldsymbol{b}_i . [1, Y_{i,k-1}, \ldots, Y_{i,k-p}]^T + \epsilon_{i,k}, \end{aligned}$$

where $\boldsymbol{b}_i = [c_i, \phi_{i,1}, \ldots, \phi_{i,p}]^T$, and $\epsilon_{i,k}$ is a white noise sequence which satisfies $\mathbb{E}[\epsilon_{i,k}] = 0$, $\mathbb{E}[\epsilon_{i,k}^2] = \sigma_i^2$, and $\mathbb{E}[\epsilon_{i,k}\epsilon_{i,k+k'}] = 0$ for all $k' \neq 0$. In $AR(p)$ model, mean value and covariance of time series are given as follows [22]:

$$\mu_i = \frac{c_i}{1 - \phi_{i,1} - \cdots - \phi_{i,p}}, \quad (2)$$

$$\gamma_{i,j} = \begin{cases} \phi_{i,1}\gamma_{i,j-1} + \cdots + \phi_{i,p}\gamma_{i,j-p} & j = 1, 2, \ldots \\ \phi_{i,1}\gamma_{i,1} + \cdots + \phi_{i,p}\gamma_{i,p} + \sigma_i^2 & j = 0 \end{cases}.$$

$\boldsymbol{b}$ can be computed by a least squared regression which leads to the following form [22],

$$\boldsymbol{b_i} = \Big(\sum_{k=p+1}^{T-1} \boldsymbol{y_{i,k}}^T \boldsymbol{y_{i,k}}\Big)^{-1} \Big(\sum_{k=p+1}^{T-1} \boldsymbol{y_{i,k}}^T y_{i,k}\Big), \quad (3)$$

where $\boldsymbol{y_{i,k}}$ is a row vector, defined as $[1, y_{i,k-1}, \ldots, y_{i,k-p}]$, and T represents a pre-determined time-horizon over which the data is collected.

Based on the data collected till step k, $\mathcal{F}$ sends an activation signal to $m < n$ sensors at the next step $k+1$ (i.e., $\sum_{i=1}^{n} u_{i,k+1} = m$). The choice of sensors at the next step is based on minimizing a cost function that models the error incurred in the estimation of missing data, and prediction of the future data. Next, we provide an elaborate description of the cost functions considered in this paper.

A. Prediction Error

Let $\mathcal{M}(k)$ be an index set containing the indices of the sensors that are active at time k (i.e. $|\mathcal{M}(k)| = m$). The objective is to formulate a cost function that models the error in predicting the samples at round $k+1$ due to measurement constraints at step k. In order to minimize the aforementioned cost, $\mathcal{F}$ needs to compute $\mathcal{M}(k)$ at stage $k-1$ based on the data received from the sensors till stage $k-1$. Let $\mathcal{N} = \{1, \ldots, n\}$. The following equation provides an expression for the mean square prediction error at step $k+1$:

$$\begin{aligned} f(\boldsymbol{b}, \boldsymbol{\sigma}, \boldsymbol{y}) = \mathbb{E}[&\sum_{i \in \mathcal{M}(k)} (Y_{i,k+1} - \hat{Y}_{i,k+1})^2 \quad (4) \\ &+ \sum_{i \in \mathcal{N} \setminus \mathcal{M}(k)} (Y_{i,k+1} - \hat{Y}_{i,k+1})^2], \end{aligned}$$

where $\boldsymbol{b}, \boldsymbol{\sigma}, \boldsymbol{y}$ are concatenation of $\boldsymbol{b_i}$,σ_i ,and $\boldsymbol{y_{i,k}}$ for $i = 1, \ldots, n$, respectively. The first summation in (4) represents the prediction error for sensors in $\mathcal{M}(k)$, in which $\hat{Y}_{i,k+1}$ can be written in terms of $Y_{i,k}$ which is a random variable with mean and covariance expressed in (2). The second summation represents the prediction error for inactive sensors at stage $k-1$. Note that $\hat{Y}_{i,k+1}$ in second summation is also a random variable due to noise in prediction model. Therefore, the loss function can be written as

$$\begin{aligned} f &= \mathbb{E}[\sum_{i \in \mathcal{M}(t)} (Y_{t+1}^i - \underbrace{\hat{Y}_{i,k+1}}_{\boldsymbol{b}_i . [1, Y_{i,k}, \ldots, Y_{i,k-p+1}]^T + \epsilon_{i,k+1}})^2 \\ &+ \sum_{i \in \mathcal{N} \setminus \mathcal{M}(t)} (Y_{t+1}^i - \hat{Y}_{i,k+1})^2]. \quad (5) \end{aligned}$$

B. Measurement error

The objective here is to model the error incurred from missing samples at the next stage due to measurement constraints. The following cost function provides the mean square measurement error:

$$f(\boldsymbol{b}, \boldsymbol{\sigma}, \boldsymbol{y}) = \mathbb{E}[\sum_{i \in \mathcal{N} \setminus \mathcal{M}(k)} (Y_{i,k} - \hat{Y}_{i,k})^2], \quad (6)$$

where $\hat{Y}_{i,t}$ is obtained from the predictor.

The objective of the fusion center is to arrive at a sensor activation scheme at each stage k which minimizes the above loss functions. This leads to the following problem:

$$\text{Compute } \mathcal{M}(k)^* = \arg\min_{\mathcal{M}(k)} f(\boldsymbol{b}, \boldsymbol{\sigma}, \boldsymbol{y}) \quad (7)$$

III. 1-STEP OPTIMAL POLICY

In this section, we initially consider the problem of optimal temporal sampling for minimizing the prediction error. First, we develop a decision policy for $n = 2$ and $p = 1$, and subsequently, the general case is analyzed. Next, we consider the measurement error.

A. Prediction error

$(n = 2, m = 1)$: In this case, we compute the loss function for the scenario in which measurements are acquired for only one time series at each step. Let f_1 and f_2 denote the loss functions when $u_{1,k} = 1$ and $u_{2,k} = 1$, respectively. Moreover, only one of the sensors was active before stage k. Without loss of generality, assume that $u_{1,k-1} = 1$, which implies that $u_{2,k-1} = 0$. If $u_{2,k-2} = 1$, then loss functions can be written in the following form

$$\begin{aligned} f_1 &= g_1(\sigma_1^2) + g_3(c_2, \phi_2, \sigma_2^2, y_{2,k-2}), \\ f_2 &= g_2(c_1, \phi_1, \sigma_1^2, y_{1,k-1}) + g_1(\sigma_2^2), \end{aligned}$$

where

$$g_1(\sigma_i^2) = 2\sigma_i^2,$$

$$g_2(\boldsymbol{b_i}, \sigma_i^2, y_{i,k-1}) = \phi_i^4 (\frac{c_i}{\phi_i - 1} + y_{i,k-1})^2 + \sigma_i^2 \frac{\phi_i^4 - 2}{\phi_i^2 - 1}$$

$$g_3(\boldsymbol{b_i}, \sigma_i^2, y_{i,k-2}) = \phi_i^6 (\frac{c_i}{\phi_i - 1} + y_{i,k-1})^2 + \sigma_i^2 \frac{\phi_i^6 - 2}{\phi_i^2 - 1}$$

We can determine a decision rule in the space of $y_{1,k-1}, y_{2,k-1}$ based on the sign of $f_1 - f_2$. In other words, $f_1 - f_2 = 0$ specifies the curve in the space of $y_{1,k-1}, y_{2,k-1}$ that separates the set of values of $(y_{1,k-1}, y_{2,k-1})$ for which $u_{1,k} = 1$ from the set of values of $(y_{1,k-1}, y_{2,k-1})$ for which $u_{2,k} = 1$. We call the aforementioned curve as the *separator*. In this case, the separator is a hyperbola with the following equation:

$$\begin{aligned} & \phi_1^4 (y_{1,k-1} - \frac{c_1}{1-\phi_1})^2 + \phi_2^6 (y_{2,k-2} - \frac{c_2}{1-\phi_2})^2 \\ = & -\frac{\phi_1^2 \sigma_1^2 (\phi_1^2 - 2)}{1 - \phi_1^2} + \frac{\phi_2^2 \sigma_2^2 (\phi_2^4 - 2)}{1 - \phi_2^2}. \end{aligned} \quad (8)$$

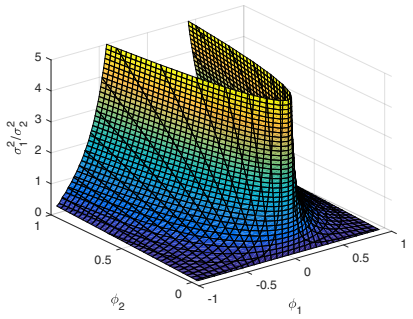


Fig. 1: Separator

Next, assume that $u_{1,k-1} = 1$ and $u_{1,k-2} = 1$. Therefore, $u_{2,k-1} = 0$ and $u_{2,k-2} = 0$ which leads to the following:

$$\begin{aligned} f_1 &= g_1(\sigma_1^2) + g_4(\sigma_2^2, \phi_2) \\ f_2 &= g_2(c_1, \phi_1, \sigma_1^2, y_{1,k-2}) + g_1(\sigma_2^2), \end{aligned}$$

where

$$g_4(\sigma_2^2, \phi_2) = 2\sigma_2^2(\phi_2^2 + 1). \quad (9)$$

In this case, comparing f_1, f_2 results in the following separator

$$\phi_1^4 (y_{1,k-1} - \frac{c_1}{1-\phi_1})^2 = \frac{\phi_1^2 \sigma_1^2 (\phi_1^2 - 2)}{1 - \phi_1^2} + 2\sigma_2^2 \phi_2. \quad (10)$$

Right hand side of the above equation can be evaluated for different values of ϕ_1, ϕ_2 and $\frac{\sigma_1^2}{\sigma_2^2}$. Figure 1, shows the surface that separates the positive and negative values of the right hand side of (10). When this expression is negative, it implies that $u_{1,k} = 1$. On the other hand, for positive values, there is a threshold policy to determine the next active sensor. By threshold policy we mean For instance, if $y_{1,k-1}$ lies in the following interval, then $u_{2,k} = 1$

$$-D \leq (y_{1,k-1} - \frac{c_1}{1-\phi_1}) \leq D$$

$$D = \frac{1}{\phi_1^2} \sqrt{(\frac{\phi_1^2 \sigma_1^2 (\phi_1^2 - 2)}{1 - \phi_1^2} + 2\sigma_2^2 \phi_2)}$$

$(n > 2, m = 1)$: In this case, we assume that at each stage only one measurement is taken. Let l_1, l_2 denote the index of time series in which $u_{l_1,k-1} = 1, u_{l_2,k-2} = 1$, and $l = \mathcal{M}(k)$. For the case $l_1 = l_2 \neq l$, the loss function can be formulated as follows:

$$f_l = g_1(\sigma_l^2) + g_2(c_{l_1}, \phi_{l_1}, \sigma_{l_1}^2, y_{l_1,k-1}) + \sum_{\mathcal{N} \setminus \{l, l_1\}} g_4(\sigma_i^2, \phi_i),$$

and, when $l_1 = l_2 = l$, we obtain the following:

$$f_l = g_1(\sigma_l^2) + \sum_{i \in \mathcal{N} \setminus \{l\}} g_4(\sigma_i^2, \phi_i).$$

For the case $l_1 \neq l_2$, we obtain the following:

$$\begin{aligned} f_l = g_1(\sigma_l^2) + g_2(c_{l_1}, \phi_{l_1}, \sigma_{l_1}^2, y_{l_1,k-1}) \\ + g_3(c_{l_2}, \phi_{l_2}, \sigma_{l_2}^2, y_{l_2,k-2}) + \sum_{i \in \mathcal{N} \setminus \{l, l_1, l_2\}} g_4(\sigma_i^2, \phi_i). \end{aligned}$$

Comparing any two pairs of f_l and $f_{l'}$ for $l, l' \in \mathcal{N}$ leads to the decision rule which is either hyperbola or threshold policy, explained in the previous subsection. **General Case($n > m$)**: In this case, we assume that m measurements are collected at each step. There exists $\binom{n}{m}$ loss functions associated to sampling from n time series. The loss function is given by the following expression:

$$\begin{aligned} f &= \sum_{i\in\mathcal{M}(k)} g_1(\sigma_i^2) + \sum_{i\in\mathcal{M}_1(k)} g_2(c_i,\phi_i,\sigma_i^2,y_{i,k-1}) \\ &+ \sum_{i\in\mathcal{M}_2(k)} g_3(c_i,\phi_i,\sigma_i^2,y_{i,k-2}) + \sum_{i\in\mathcal{M}_3(k)} g_4(\sigma_i^2,\phi_i), \end{aligned}$$

where $\mathcal{M}(k)$ is the set of active sensors at time step k, and $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ are defined as

1) $\mathcal{M}_1(k)$: $\mathcal{M}(k-1) \setminus \mathcal{M}(k)$.
2) $\mathcal{M}_2(k)$: $\mathcal{M}(k-2) \setminus (\mathcal{M}_1(k) \cup \mathcal{M}(k))$.
3) $\mathcal{M}_3(k)$: $\mathcal{N} \setminus \mathcal{M}_2(k)$.

In order to determine the optimal $u_{i,k}$, we need to compare $\binom{n}{m}$ loss functions. Since all terms in the loss function are disjoint, we can reduce the number of cases to n. Let L_i denote a term in the loss function corresponding to ith time series, when $u_{i,k} = 0$. Therefore, the loss function can be written in the following form:

$$f = \sum_{i\in\mathcal{M}(k)} g_1(\sigma_i^2) + \sum_{i\in\mathcal{N}\setminus\mathcal{M}(k)} L_i.$$

By adding and subtracting $\sum_{i\in\mathcal{M}} L_i$, f can be written in the following form:

$$f = \sum_{i\in\mathcal{M}(k)} (g_1(\sigma_i^2) - L_i) + \sum_{i\in\mathcal{N}} L_i.$$

Since the second term in the above equation is constant, minimizing the first term over all possible sets of $\mathcal{M}(k)$ is equivalent to minimizing the loss function. Moreover, the first term contains sum of m disjoint terms. Therefore, it is equivalent to finding the m lowest values of $h_i = g_1(\sigma_i^2) - L_i$ among n time series. In other words, fusion center $\mathcal{F}$ computes h_i for $i \in \mathcal{N}$, and put $u_{i,k} = 1$ for the corresponding indices of m lowest values, and $\mathcal{M}(k)$ is the set of indices for which $u_{i,k} = 1$.

B. Measurement Error

In this case, we define the loss function as the mean square prediction error for the missing measurements i,e., series for which $u_{i,k} = 0$. Therefore, f has the following form:

$$f = \mathbb{E}[\sum_{i\in\mathcal{N}\setminus\mathcal{M}} (Y_{i,k} - \hat{Y}_{i,k})^2]. \tag{11}$$

Substituting $\hat{Y}_{i,k}$ by AR(1) model leads to the following:

$$\begin{aligned} f &= \sum_{i\in\mathcal{M}_1} q_1(c_i,\phi_i,\sigma_i^2,y_{i,k-1}) \\ &+ \sum_{i\in\mathcal{M}_2} q_2(c_i,\phi_i,\sigma_i^2,y_{i,k-2}) + \sum_{i\in\mathcal{M}_3} q_3(c_i,\phi_i), \end{aligned} \tag{12}$$

where $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ are defined in previous section, and q_1, q_2 and q_3 are defined as follows:

$$\begin{aligned} q_1 &= \phi_i^2(\frac{c_i}{\phi_i - 1} + y_{i,k-1})^2 + \sigma_i^2 \frac{\phi_i^2 - 2}{\phi_i^2 - 1}, \\ q_2 &= \phi_i^4(\frac{c_i}{\phi_i - 1} + y_{i,k-2})^2 + \sigma_i^2 \frac{\phi_i^4 - 2}{\phi_i^2 - 1}, \\ q_3 &= 2\sigma_i^2. \end{aligned} \tag{13}$$

Since the approach is the same as previous loss function, we only present the general case for n, m.

General Case (n, m): There are $\binom{n}{m}$ loss functions associated with sampling m time series. The loss function is given by the following expression:

$$\begin{aligned} f &= \sum_{i\in\mathcal{M}_1} q_1(c_i,\phi_i,\sigma_i^2,y_{i,k-1}) \\ &+ \sum_{i\in\mathcal{M}_2} q_2(c_i,\phi_i,\sigma_i^2,y_{i,k-2}) + \sum_{i\in\mathcal{M}_3} q_3(\sigma_i^2), \end{aligned}$$

In order to find the most informative time series, $\binom{n}{m}$ loss functions need to be compared. However, we can reduce the number of cases to n since all terms are disjoint. f can be written in the following form by defining L_i as the loss function associated with the ith time series when $u_{i,k} = 0$

$$f = \sum_{i\in\mathcal{N}\setminus\mathcal{M}} L_i.$$

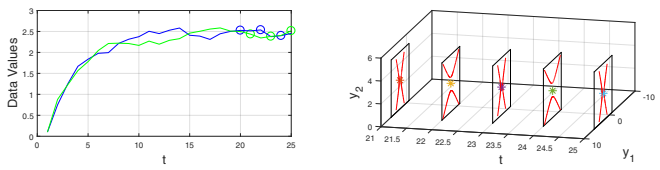
Minimizing the above function over all possible composition of $\mathcal{M}$ is equivalent to finding $n - m$ lowest values of L_i for $i \in \mathcal{N}$.

IV. Simulation

In this section, the simulation results for the proposed algorithms are presented. In all the simulations, we use the realization of auto regressive model. We assume that all time series are initially observed for some time to compute the parameters of the prediction model based on (3).

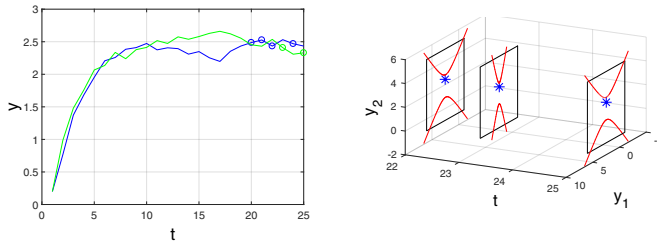
In our first simulation, we consider the case when $n = 2, m = 1$. We assume that Y_1 and Y_2 are initially observed for 20 steps. Our goal is to determine the active sensors for the next steps. In Figure 2a and 2b, the prediction error is considered as the loss function, and we use AR(1) as our prediction model. Active sensors can be determined by examining the location $y_{1,k-1}, y_{2,k-2}$ respect to the hyperbola (8), described in the previous sections. In Figure 2a, the informative time series is marked with a circle for next 5 steps. The decision policy and the data are depicted in Figure 2b in different steps. In Figure 3a and 3b, we consider the measurement loss function.

In the second simulation, we consider the 4th order AR model as prediction model in our loss functions. Figure 4a depicts the results for $n = 2$. Both time series are initially observed for 50 steps for estimating the parameters in the prediction model. Next, we pick the informative sample points for next 50 steps. Note that in this simulation we used the initial parameters (i.e. $\boldsymbol{b_i}$) which are estimated from



(a) (b)

Fig. 2: (a) Informative samples in two time series. Prediction error is loss function, and prediction model is AR(1). (b) Decision policy for different steps



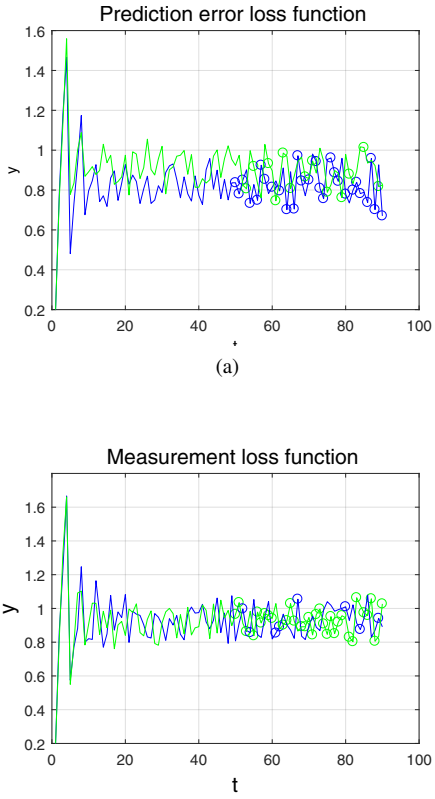
(a) (b)

Fig. 3: (a) Informative samples in two time series. Measurement loss function, and prediction model is AR(1). (b) Decision policy for different steps

the first 50 steps of the time series. We can also update b_i according to the current observed data. We call this scenario as learning parameters. We expect that the loss function after incorporating parameter learning should be lower. Figure 5a depicts the loss functions for learning vs non-learning scenarios (prediction error loss function, measurement error loss function). This implies that prediction and measurements are more close to each other.

Based on the simulation results, we propose the following hypothesis: For two series with close parameters, the dominant terms are the variances of white noises included in the time series (i.e σ_1^2, σ_2^2). This implies that when σ_1 is big enough compared to σ_2, one only needs to measure Y_1. On the other hand, when $\sigma_2 > \sigma_1$, the policy is reversed. From simulations, we observe that when the $|\sigma_1 - \sigma_2| > 0.1$, active sensors can be determined based on only σ_1 and σ_2. In other words, when $\sigma_1 - \sigma_2 > 0.1$, then $u_{1,k} = 1$, and when $\sigma_2 - \sigma_1 > 0.1$, then $u_{2,k} = 1$. Figure 6 shows three regions in the space of σ_1, σ_2. In the white region, the decision policy is only based on the value of σ_1, σ_2, and in the green region it depends on $\sigma_1, \sigma_2, b_1, b_2$. In other words, when the uncertainty is high, prediction is less accurate. Thus the loss function is larger. As the noise level in the time series is increased, the loss function increases. Figures 7b and 7a show the variation of the loss function for several σ_1 and σ_2 for different values of p.

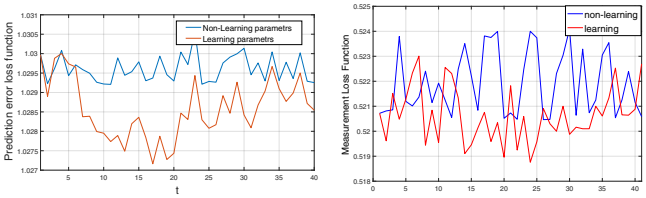
Next, we simulate the case $n = 8$ and $m = 3$ for the prediction error loss function. At each step, we compute the smallest three h_i's among all time series. Figure 8 shows the informative time series at each step with a circle.



(a)

(b)

Fig. 4: (a) Informative samples in two time series. Prediction error loss function, and prediction model is AR(4). (b) Measurement loss function, and prediction model is AR(4).



(a) (b)

Fig. 5: Loss function in learning vs non-learning parameters. (a) Prediction error loss function.(b) Measurement loss function

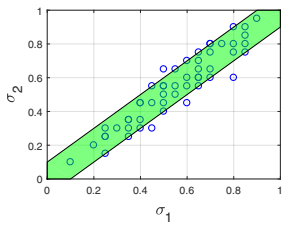
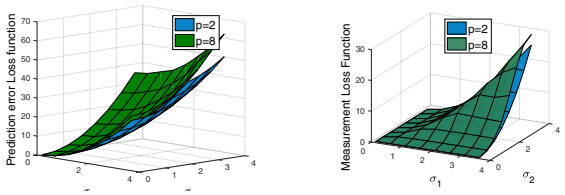


Fig. 6: In white regions the decision policy is only based on the value of σ_1, σ_2, and in the green region other parameters also matter.



(a) (b)

Fig. 7: Variation of loss function for different σ_1 and σ_2. (a) Prediction error loss function.(b) Measurement loss function

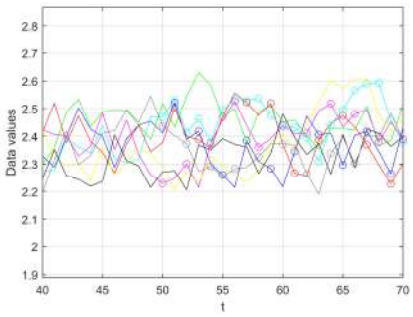


Fig. 8: Multiple Time Series (8,3)

V. Conclusion

In this work, a data-driven sensor scheduling technique is proposed to collect the data from n discrete-time dynamical processes. Based on an auto-regressive (AR) prediction model, we defined two loss functions, namely, the prediction error and the measurement error associated with the missing measurements. The problem of choosing the measurements at each step is formulated as an optimization problem over the two loss functions. Moreover, we derived the optimal measurement policy for the case of $n = 2, m = 1$. We presented simulations to illustrate the performance of the proposed algorithm, and compare it to the case when the parameters of the AR predictor are updated at each step based on the measurements.

An ongoing effort in our research is to investigate other prediction models which results in computationally efficient optimization problem. Moreover, we are interested in the case when time series are not independent, and using this property in the loss functions to select next measurements. This problem may provide solution for the sensor scheduling problem for the spatio-temporal phenomenon.

VI. Acknowledgment

This work was in part supported by USDA\ NIFA Award Number 2017-67021-25965.

References

[1] R. G. Baraniuk, "More is less: Signal processing and the data deluge," *Science*, vol. 331, no. 6018, pp. 717–719, 2011.

[2] C. Z. Bai and V. Gupta, "An on-line sensor selection algorithm for sprt with multiple sensors," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3532–3539, July 2017.

[3] T. H. Chung, V. Gupta, J. W. Burdick, and R. M. Murray, "On a decentralized active sensing strategy using mobile sensor platforms in a network," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 2. IEEE, 2004, pp. 1914–1919.

[4] H.-L. Choi, "Adaptive sampling and forecasting with mobile sensor networks," Ph.D. dissertation, Massachusetts Institute of Technology, 2009.

[5] C. Z. Bai and V. Gupta, "An on-line sensor selection algorithm for sprt with multiple sensors," in *53rd IEEE Conference on Decision and Control*, Dec 2014, pp. 6869–6874.

[6] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2451–2464, Oct 2013.

[7] K. Premkumar and A. Kumar, "Optimal sleep-wake scheduling for quickest intrusion detection using wireless sensor networks," in *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, April 2008, pp. 1400–1408.

[8] Y. He and E. K. Chong, "Sensor scheduling for target tracking: A monte carlo sampling approach," *Digital Signal Processing*, vol. 16, no. 5, pp. 533–545, 2006.

[9] 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, no. 2, pp. 251–260, 2006.

[10] A. O. Hero and D. Cochran, "Sensor management: Past, present, and future," *IEEE Sensors Journal*, vol. 11, no. 12, pp. 3064–3075, Dec 2011.

[11] A. Y. Khapalov, "Optimal measurement trajectories for distributed parameter systems," *Systems & control letters*, vol. 18, no. 6, pp. 467–477, 1992.

[12] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena scientific Belmont, MA, 1995, vol. 1, no. 2.

[13] S. Chakravorty and R. S. Erwin, "Information space receding horizon control," in *Adaptive Dynamic Programming And Reinforcement Learning (ADPRL), 2011 IEEE Symposium on*. IEEE, 2011, pp. 302–309.

[14] Z. Sunberg, S. Chakravorty, and R. S. Erwin, "Information space receding horizon control for multisensor tasking problems," *IEEE transactions on cybernetics*, vol. 46, no. 6, pp. 1325–1336, 2016.

[15] O. C. Imer and T. Basar, "Optimal estimation with limited measurements," *International Journal of Systems, Control and Communications*, vol. 2, no. 1-3, pp. 5–29, 2010.

[16] V. Srivastava, K. Plarre, and F. Bullo, "Randomized sensor selection in sequential hypothesis testing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2342–2354, May 2011.

[17] V. Katewa and V. Gupta, "On the optimality of sequential test with multiple sensors," in *2012 American Control Conference (ACC)*, June 2012, pp. 282–287.

[18] C. Z. Bai, V. Gupta, and Y. F. Huang, "Sequential hypothesis testing with off-line randomized sensor selection strategy," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 3269–3272.

[19] M. Naghshvar and T. Javidi, "Sequentiality and adaptivity gains in active hypothesis testing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 768–782, Oct 2013.

[20] M. Naghshvar, T. Javidi *et al.*, "Active sequential hypothesis testing," *The Annals of Statistics*, vol. 41, no. 6, pp. 2703–2738, 2013.

[21] V. Srivastava, K. Plarre, and F. Bullo, "Adaptive sensor selection in sequential hypothesis testing," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, Dec 2011, pp. 6284–6289.

[22] J. D. Hamilton, *Time series analysis*. Princeton university press Princeton, 1994, vol. 2.