2019 18th European Control Conference (ECC)
Napoli, Italy, June 25-28, 2019

On the Estimation of Signal Attacks: A Dual Rate SD Control Framework

Nabil H. Hirzallah, Petros G. Voulgaris and Naira Hovakimyan

***Abstract*— We consider the problem of estimating signal attacks injected into the actuators or sensors of control systems, assuming the attack is detectable, i.e., it can be seen at the output. We show that there exists a trade-off between attack rejection and control, and that the estimator design depends on the controller used. We use dual rate sampling to enhance detectability of the attacks and we provide different methods to design the estimator. The first method is by solving a model matching problem subject to causality constraints. The second method exploits dual rate sampling to accurately reconstruct the unknown input. The third method is using a dual rate unknown input observer. We provide conditions on the existence of these estimators, and show that dual rate unknown input observers always exist if the multirate system does not have a zero at 1.**

I. INTRODUCTION

The problem of security of control systems under signal attacks has caught a lot of attention lately after the several attacks on critical infrastructure world wide [1]. In our previous work [2] we presented the conditions for the existence of unbounded stealthy attacks in terms of system structure and zero/pole locations. We showed that studying the problem in a sampled-data (SD) framework is important because sampling may introduce additional unstable zeros that render the system vulnerable. We also showed that dual rate control removes unstable zeros and can detect stealthy actuator attacks. In [3] we showed that the computation of stealthy worst attacks is a convex maximization problem, which can be converted to a series of LP problems if the attack is bounded in time. We also presented an iterative controller design method to reduce the effect of worst attacks.

In this paper, we consider the problem of estimating signal attacks on actuators and/or sensors of control systems using the available measurements. The estimated attack signal will help the operator decide whether it is a persistent intelligent attack or just a nominal disturbance. First, we show that there is a coupling between attack estimation and rejection, and that a trade-off exists between their individual performances. The quality of the estimate depends on the performance of the attack rejection controller. In particular, the faster and better we reject the attack, the worse is the attack estimate. This is of course assuming the attack can be detected or seen from the outputs used for estimation.

Nabil H. Hirzallah is a PhD candidate with the Electrical and Computer Engineering Department, University of Illinois, Urbana, IL. hirzall2@illinois.edu

Petros G. Voulgaris is with the Aerospace Engineering Department, University of Illinois, Urbana, IL. voulgari@illinois.edu

Naira Hovakimyan is with the Mechanical Engineering Department, University of Illinois, Urbana, IL. nhovakim@illinois.edu

This work was supported in part by the National Science Foundation under NSF awards CMMI-1663460, ECCS-1739732, CCF-1717154.

Next we consider multirate (MR) sampling to estimate the injected attack d. In particular, we consider the case where we have two sets of sensors measuring the output. The first set is sampled at the same rate of the hold device, and is used to provide input for the feedback controller creating a single rate control system. The second set is secure and is sampled at a higher frequency, and is used for attack detection and estimation. This architecture is practical for different applications such as wireless networked control systems, where the sensor measurements are sent over wireless (unsecured) networks to the control center, and the control signals are sent back to the physical plant again over wireless networks. A local estimator that has access to some of the measurements over hard-wired secure lines can be built to generate the attack estimates in this kind of scenario. The faster sampling loop is needed so that all unbounded attacks are detectable (i.e. removes the unstable zeros) [2], and to allow for the design of a certain class of observers as will be discussed later. Furthermore, we want to estimate the attack at a faster rate than control so that we can isolate the attack and limit the damage as fast as possible. In addition to detecting unbounded attacks, removing the unstable zeros is essential because they limit the achievable estimation performance. The attack estimation problem is similar to the unknown input observer (UIO) problem discussed in [4], [5], [6], [7], [8], [9] in which such an observer exists if and only if the system is strongly detectable, i.e., all zeros are strictly stable. Multirate sampling guarantees that the system has at most one non-minimum phase zero and is located at $\lambda = 1$, and under specific conditions, multrirate sampling can remove all zeros in the lifted domain. Conditions when a zero at $\lambda = 1$ exists in the MR scheme can be found in [2]. After introducing dual rate sampling for attack estimation, we introduce a few estimator design methods utilizing the dual rate property. In particular, we show that UIOs always exist if the dual rate system does not have a zero at $\lambda = 1$. In addition, the observer provides an estimate of the attack with a delay of a single time-step only. This result is significant because single rate observers do not exist most of the time due to the hard conditions for their existence [7], or they may exist but estimation is delayed (system must be strongly detectable) [4].

Some standard notation we use is as follows: For a sequence of real $n \times m$ dimensional real matrices $G = \{G_k\}_{k\in\mathbb{Z}_+}$ we denote its λ-transform $G(\lambda) := \sum_{k=0}^{\infty} G_k\lambda^k$. For a λ-transform $x(\lambda)$ of a sequence x of n-dimensional vectors $||x(\lambda)|| = ||x||$. Finally, we will call system G "tall" when $y = Gu$ and the dimension of the output vector y is at least equal to the dimension of the input u; otherwise we

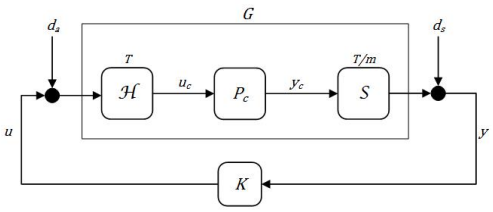

Fig. 1: A dual rate SD system.

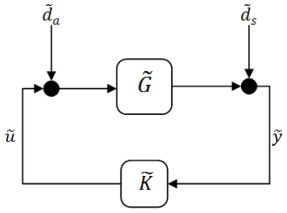

Fig. 2: The lifted system.

will call G "fat". We will also be using the standard notions for zeros of a LTI system G (e.g., [10].)

II. Preliminaries

This section reviews the results in our previous work [2] needed for the work in this paper. We consider the physical, continuous-time, LTI plant $P_c = [A_c, B_c, C_c, D_c]$ of Figure 1 that is controlled by a digital controller K using the standard zero order hold and sampling devices $\mathcal{H}$ and $\mathcal{S}$ respectively. In the sampled-data (SD) scheme of Figure 1 (temporarily without any disturbances) the digital controller input $u = \{u(k)\}$ converts to the continuous time input $u_c(t) = (\mathcal{H}u)(t) = u(k)$ for $kT \leq t < (k+1)T$ where T is the hold period. The output is sampled with period T/m where m is a sufficiently large integer, i.e., $y(k) = (\mathcal{S}_m y_c)(t) := y_c(kT/m)$. To this end, let the corresponding discrete-time system mapping u to y be

$$G = \mathcal{S}_m P_c \mathcal{H}.$$

For this MR discrete system we have that

$$\Lambda^m G = G\Lambda,$$

where Λ is the 1-step right shift operator on discrete sequences $\{x(k)\}$, i.e., $(\Lambda x)(k+1) = x(k)$ with $(\Lambda x)(0) = 0$. Using standard lifting techniques (e.g., [11]) one can obtain a shift invariant (LTI) description $\tilde{G}$ of the discrete dynamics by grouping the plant input and output signals as $\tilde{u}(k) = u(k)$ and $\tilde{y}(k) = [y_c'(kT/m)\ y_c'((k+1)T/m) \dots y_c'((k+m-1)T/m)]'$ (similarly for $\tilde{d}_a$ and $\tilde{d}_s$.) A state space description for $\tilde{G}$ can be obtained from the original system.

Define state space matrices

$$A := e^{A_c T/m} \in \mathbb{R}^{n\times n}, \quad B := \int_0^{T/m} e^{A_c \tau} B_c \mathrm{d}\tau \in \mathbb{R}^{n\times n_u},$$
$$C := C_c \in \mathbb{R}^{n_y \times n}, \qquad D := D_c \in \mathbb{R}^{n_y \times n_u}.$$

And assume that $|CB| \neq 0$, then

$$\tilde{G} = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right], \tag{1}$$

where

$$\tilde{A} = A^m \in \mathbb{R}^{n\times n}, \tilde{B} = \sum_{k=0}^{m-1} A^k B \in \mathbb{R}^{n\times n_u},$$

$$\tilde{C} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{m-1} \end{bmatrix}, \ \tilde{D} = \begin{bmatrix} D \\ CB + D \\ \vdots \\ C\sum_{k=0}^{m-2} A^k B + D \end{bmatrix},$$

such that $\tilde{C} \in \mathbb{R}^{mn_y \times n}$ and $\tilde{D} \in \mathbb{R}^{mn_y \times n_u}$. We consider now the closed loop in the lifted domain in Figure 2 where the controller is $\tilde{K}$. To this end, the integer m is chosen such that the following assumptions are satisfied.

Assumption 1. *The matrix B is full column rank.*

Assumption 2. *The matrix $\mathcal{O} := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{m-2} \end{bmatrix}$ is full column rank.*

The first assumption is standard and holds generically if B_c is full column rank in the continuous system. The second assumption holds for large enough m, in particular $m = n + 1$, if the pair (A, C) is observable. It can also hold however, even with a small m generically. Also, if Assumption 2 holds, $\tilde{G}$ is a tall system. Then the following lemma from [2] characterizes the zeros of $\tilde{G}$.

Lemma 3. *Consider the lifted system $\tilde{G}$ as in (1) together with Assumptions (1) and (2). Then $\tilde{G}$ has at most one non-minimum phase zero and is located at $\lambda = 1$.*

In [2] we showed that if $\tilde{G}$ has a non-minimum phase zero, then this zero is located at $\lambda = 1$, and its multiplicity is 1. As a result, dual rate control renders the system secure against unbounded stealthy actuator attacks. This applies to any P_c of any structure. For the case when P_c is tall and has no zeros at the origin, [12] (Theorem 1) states that $\tilde{G}$ has no zeros at all for almost all $m \in \mathcal{R}$ such that $m > 1$. In our MR scheme in [2] and in this paper, we only consider m to be an integer.

III. Attack Estimation-Rejection Trade-off

In the absence of zero dynamics attack possibilities, and for single rate systems, we investigate the trade-off between the ability to control the damage that an attack d inflicts versus the ability to estimate d. In other words, we would like to investigate whether one can trade control performance for extra ability to estimate the attack signal d. A relevant problem to study how to design a controller K jointly with a filter F to reject as well as estimate d can be cast as

$$\min_{K,F} \|d \mapsto z\|, \quad z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix},$$

where z_1 relates to performance in terms of disturbance rejection, e.g., $z_1 = \begin{bmatrix} W_1 y \\ u \end{bmatrix}$ and z_2 relates to attack estimation, i.e., $z_2 = W_2(d - \hat{d})$, where $\hat{d}$ is the estimated

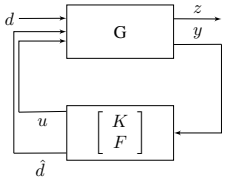

Fig. 3: General block diagram to reject and estimate d.

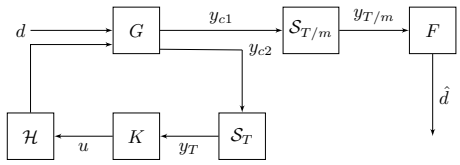

Fig. 4: General block diagram to reject and estimate d with secured sensors.

attack and $W_{1,2}$ are weights, as seen in the general block diagram in Figure 3, where G is a general discrete LTI system. This type of problem is convex in any norm when the Youla parametrization of all stabilizing controllers [13] is employed. The input-output map of the system in Figure 3 can be described as:

$$\begin{bmatrix} z_1 \\ z_2 \\ y \end{bmatrix} = \left[\begin{array}{c|cc} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ \hline G_{31} & G_{32} & G_{33} \end{array}\right] \begin{bmatrix} d \\ u \\ \hat{d} \end{bmatrix},$$

where G_{32} is the open loop discrete time LTI plant. For z_2 defined as $z_2 = W(d - \hat{d})$ and z_1 does not depend on $\hat{d}$, we have $G_{13} = G_{22} = G_{33} = 0,\ G_{21} = W,\ G_{23} = -W$. For actuator-only attacks, we have

$$G_{31} = G_{32},$$

while for sensor-only attacks

$$G_{31} = I.$$

The remaining maps G_{11} and G_{12} depend on how z_1 is defined. The input-output map is now more sparse and can be described by:

$$\begin{bmatrix} z_1 \\ z_2 \\ y \end{bmatrix} = \left[\begin{array}{c|cc} G_{11} & G_{12} & 0 \\ W & 0 & -W \\ \hline G_{31} & G_{32} & 0 \end{array}\right] \begin{bmatrix} d \\ u \\ \hat{d} \end{bmatrix}.$$

The closed loop map T_{zd} can then be found and is

$$\begin{bmatrix} T_{z_1 d} \\ T_{z_2 d} \end{bmatrix} = \begin{bmatrix} G_{11} \\ W \end{bmatrix} + \begin{bmatrix} G_{12} & 0 \\ 0 & -W \end{bmatrix} \begin{bmatrix} K \\ F \end{bmatrix} \left(I - \begin{bmatrix} G_{32} & 0 \end{bmatrix} \begin{bmatrix} K \\ F \end{bmatrix} \right)^{-1} G_{31}$$
$$= \begin{bmatrix} G_{11} + G_{12}K(I - G_{32}K)^{-1}G_{31} \\ W - WF(I - G_{32}K)^{-1}G_{31} \end{bmatrix}.$$

It is easy to see that minimizing $\|T_{z_1 d}\|$ depends only on finding the optimal K and can be solved as a model matching problem. On the other hand, minimizing $\|T_{z_2 d}\|$ depends on finding the optimal K and F simultaneously. By inspecting T_{zd}, keeping $\|T_{z_2 d}\|$ small is achieved by making $|F(I - G_{32}K)^{-1}G_{31}| \approx I$ for all frequencies. On the other hand, it is well known that $|K|$ has to be large for good disturbance rejection [14]. As a result, a trade-off between good estimation and good disturbance rejection exists. An example can be found in [15].

IV. Estimation Via Multirate Sampling

A. Motivation and Control Loop Architecture

Next we consider multirate (MR) sampling to estimate the injected attack d. In particular, we consider the case where we have two sets of sensors measuring the output. The first set is sampled at the same rate of the hold device, and is used to provide input for the feedback controller creating a single rate control system. The second set is secure and is sampled at a higher frequency, and is used for attack detection and estimation as seen in Figure 4, where G is the continuous-time LTI general input-output map. This architecture is practical for different applications such as wireless networked control systems. A local estimator that has access to some of the measurements over hard-wired secure lines can be built to generate the attack estimates in this kind of scenario. Higher sampling rate for the detection loop is necessary to detect stealthy unbounded attacks [2], and to make it harder for attackers to design stealthy bounded attacks [3]. In addition, MR sampling removes unstable zeros (except for possibly one zero at $\lambda = 1$) in the map from the attack signal d to the monitored signals (y and possibly u).

Remark 4. *The control loop in the architecture in Figure 4 can also be dual rate. What is important is to have the output feeding the estimation loop sampled at a sufficiently higher rate than that at which the attack is injected into the system.*

B. Estimator Design

In this section we present a few control methods for the design of the estimator F for a fixed controller K, for the architecture in Figure 4.

1) Model Matching

First we find the mapping from d to the measurements

$$\begin{bmatrix} \tilde{y} \\ y_T \end{bmatrix} = G_d \begin{bmatrix} d \\ u \end{bmatrix} = \begin{bmatrix} L\mathcal{S}_{T/m}G_{11}\mathcal{H} & L\mathcal{S}_{T/m}G_{12}\mathcal{H} \\ L\mathcal{S}_T G_{21}\mathcal{H} & L\mathcal{S}_T G_{22}\mathcal{H} \end{bmatrix} \begin{bmatrix} d \\ u \end{bmatrix},$$

where $\tilde{y}(k) = [y'_{c1}(kT/m)\ y'_{c1}((k+1)T/m) \dots y'_{c1}((k+m-1)T/m)]'$, $y_T(k) = y_{c2}(kT)$, G_{11} is the mapping from d to y_{c1}, G_{12} is the mapping from u to y_{c1} and G_{21} is the mapping from d to y_{c2}, G_{22} is the mapping from u to y_{c2}, y_{c1} and y_{c2} are the continuous-time measurements feeding $\mathcal{S}_T/m$ and $\mathcal{S}_T$ respectively, and L is the lifting operator. G_{11} may represent actuator attacks or sensor attacks as explained in section III. In view of the above, let G be controllable, observable and have the following representation:

$$G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right], \tag{2}$$

then

$$G_d = \left[\begin{array}{c|cc} A_d & B_{1d} & B_{2d} \\ \hline \tilde{C}_1 & \tilde{D}_{11} & \tilde{D}_{12} \\ C_2 & D_{21} & 0 \end{array}\right],$$

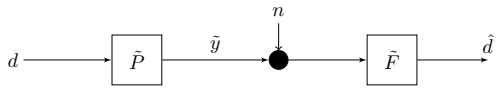

Fig. 5: Block diagram for estimator design in the lifted domain. $\tilde{P}$ is dual rate, lifted and augmented with a controller for stabilization.

where

$$\tilde{C}_1 = \begin{bmatrix} C_1 \\ C_1A_f \\ \vdots \\ C_1A_f^{m-1} \end{bmatrix}, \quad \tilde{D}_{11} = \begin{bmatrix} D_{11} \\ D_{11} + C_1B_{1f} \\ \vdots \\ D_{11} + C_1 \sum_{k=0}^{m-2} A_f^k B_{1f} \end{bmatrix},$$

$$\tilde{D}_{12} = \begin{bmatrix} D_{12} \\ D_{12} + C_1B_{2f} \\ \vdots \\ D_{12} + C_1 \sum_{k=0}^{m-2} A_f^k B_{2f} \end{bmatrix},$$

and

$$A_d := e^{AT}, \qquad A_f := e^{AT/m},$$
$$B_{1d} := \int_0^T e^{A\tau} B_1 \mathrm{d}\tau, \qquad B_{2d} := \int_0^T e^{A\tau} B_2 \mathrm{d}\tau,$$
$$B_{1f} := \int_0^{T/m} e^{A\tau} B_1 \mathrm{d}\tau, \quad B_{2f} := \int_0^{T/m} e^{A\tau} B_2 \mathrm{d}\tau$$

Now for a given controller K with state space

$$K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right],$$

the input-output map from d to $y_{T/m}$ is described as

$$\tilde{P} = \left[\begin{array}{cc|c} A_d + B_{2d}D_KC_2 & B_{2d}C_K & B_{1d} + B_{2d}D_KD_{21} \\ B_KC_2 & A_K & B_KD_{21} \\ \hline \tilde{C}_1 + \tilde{D}_{12}D_KC_2 & \tilde{D}_{12}C_K & \tilde{D}_{11} + \tilde{D}_{12}D_KD_{21} \end{array}\right], \quad (3)$$

as seen in Figure 5, where n is sensor noise. The problem of finding the best $\hat{d}$ (in some sense) can now be cast as

$$\min_F \left\| W - W\tilde{F}\tilde{P} \right\|,$$

or in the case of noisy measurements

$$\min_F \left\| W(I \;\; 0) - W\tilde{F}(\tilde{P} \;\; I) \right\|$$

such that F is stable (to minimize noise amplification) and causal. Since we are solving the problem in the lifted domain, the causality of F is guaranteed by enforcing the constraint that $F(0)$ is block lower triangular. Several methods to solve this synthesis problem can found in the literature such as in [16], [17].

2) Unknown Input Reconstruction

In this section we seek to exploit dual rate sampling to accurately reconstruct the unknown input (attack) d injected in the system in Figures 4 and 5, as well as the initial condition $x(0)$. In particular, we consider the relationship between the states and input from one end and the output of the system from another end. This relationship has been studied for single rate systems in the context of strong observability in the literature [5], [9], [8]. We consider the state space description $\tilde{P}$ in (3). We assume for now without loss of generality that $K = 0$; we also assume that the measurements are noise free, as a result $\tilde{P}$ reduces to

$$\tilde{P} = \left[\begin{array}{c|c} A_d & B_{1d} \\ \hline \tilde{C}_1 & \tilde{D}_{11} \end{array}\right]. \quad (4)$$

The lifted output of $\tilde{P}$ can be described as

$$\begin{bmatrix} y(0) \\ y(T/m) \\ y(2T/m) \\ \vdots \\ y(\frac{(m-1)T}{m}) \end{bmatrix} = \underbrace{\begin{bmatrix} C_1 \\ C_1A_f \\ C_1A_f^2 \\ \vdots \\ C_1A_f^{m-1} \end{bmatrix}}_{\tilde{C}_1} x(0) + \underbrace{\begin{bmatrix} D_{11} \\ D_{11} + C_1B_{1f} \\ D_{11} + C_1B_{1f} + C_1A_fB_{1f} \\ \vdots \\ D_{11} + C_1 \sum_{k=0}^{m-2} A_f^k B_{1f} \end{bmatrix}}_{\tilde{D}_{11}} d(0). \quad (5)$$

From the above equation, we can deduce that $x(0)$ and $d(0)$ can be recovered without delay with respect to the original single rate system if and only if

$$\begin{bmatrix} \tilde{C}_1 & \tilde{D}_{11} \end{bmatrix} \quad (6)$$

is full column rank. A necessary condition for (6) to have full column rank is that $\tilde{P}$ be strongly observable, i.e., have no invariant zeros [5], [9].

Strong observability of $\tilde{P}$ is guaranteed if G_{11} is tall and does not have a zero at the origin for a sufficiently large m (Theorem 1 in [12]), given that (A_f, C_1) is observable. Choosing m to satisfy Assumptions 1 and 2 of section II is one choice. Strong observability of $\tilde{P}$ alone does not imply that (6) is full column rank; however, (6) can be made to have full column rank by choosing m sufficiently large [18]. The idea is to add more linearly independent rows to (6) by sampling faster until (6) is tall. The added rows are linearly independent because m satisfies Assumptions 1 and 2, assuming $|C_1B_{1d}| \neq 0$ as mentioned in section II.

Remark 5. *The attack and the states are reconstructed with no delay with respect to original single rate period T. Still, m samples are needed within T, so in actual continuous-time the delay is T sec (or one sample period of the original single rate). In contrast, for single rate systems the delay can be up to nT sec where n is the dimension of A_d (assuming the observer does exist using single rate control, i.e., the single rate system is strongly observable).*

Remark 6. *As long as we choose m to make $\tilde{P}$ strongly observable, then we can still reconstruct d using delayed measurements (i.e., $\tilde{y}(0), \tilde{y}(1), \ldots, \tilde{y}(n)$, where n is the dimension of A_d) even if* (6) *is not full column rank. Details*

about this scheme for single rate systems can be found in [5].

a) Example - Automatic Voltage Regulator (AVR)

The open loop state space representation of a linearized single rate system AVR after discretization at a sample rate $T = 0.5$ sec is

$$A_d = \begin{bmatrix} 0.0105 & 0.3949 & 3.86 & 2.869 \\ -0.0057 & -0.1817 & -1.369 & -0.587 \\ 0.00117 & 0.03359 & 0.1793 & -0.4597 \\ 0.00092 & 0.03197 & 0.3163 & 0.8918 \end{bmatrix}, B_d = \begin{bmatrix} -0.005738 \\ 0.001174 \\ 0.0009193 \\ 0.0002165 \end{bmatrix}, \quad (7)$$
$$C_d = \begin{bmatrix} 0 & 0 & 0 & 5000 \end{bmatrix}, D_d = \begin{bmatrix} 0 \end{bmatrix},$$

which has an unstable zero at $\lambda = -0.7045$. Assume we want to estimate an actuator attack d. Since the system has an unstable zero, we know that we cannot reconstruct attacks even if we use an arbitrary large number of measurements. Choosing $m = 2$ resulting in $T/m = 0.25$ sec removes the unstable zero when viewed from the lifted domain. The resulting open loop state space representation after lifting is

$$\tilde{A} = A_d,\ \tilde{B} = B_d,$$
$$\tilde{C} = \begin{bmatrix} 0 & 0 & 0 & 5000 \\ 2.185 & 86.13 & 1092 & 4902 \end{bmatrix}, \tilde{D} = \begin{bmatrix} 0 \\ 0.196 \end{bmatrix}. \quad (8)$$

Although the open loop system is strongly observable using $m = 2$, $[\tilde{C} \quad \tilde{D}]$ is not full column rank, and we cannot reconstruct d without delay. Next if we select $m = 5$, the resulting $\tilde{C}$ and $\tilde{D}$ matrices become

$$\tilde{C} = \begin{bmatrix} 0 & 0 & 0 & 5000 \\ 0.38 & 21.38 & 491.24 & 4994.4 \\ 1.53 & 64.35 & 917.65 & 4948.5 \\ 2.80 & 106.05 & 1238.6 & 4839.5 \\ 3.86 & 138.22 & 1454.3e & 4671.5 \end{bmatrix}, \tilde{D} = \begin{bmatrix} 0 \\ 0.011 \\ 0.10 \\ 0.32 \\ 0.66 \end{bmatrix}.$$

Now $[\tilde{C} \quad \tilde{D}]$ is full column rank and the attack along with $x(0)$ can be reconstructed without delay.

This concludes how to reconstruct d using dual rate sampling. The case where K is augmented in $\tilde{P}$ as in (3) can be handled similarly as long as K does not introduce a zero at $\lambda = 1$ in the closed loop map from d to $y_{T/m}$.

3) Unknown Input Observer

In the previous section we saw how to reconstruct d and $x(0)$ given that $\tilde{P}$ is sampled faster than the rate at which the input is feeding the system. However, the method involved inverting a matrix with high dimensions, which might be computationally expensive. A cheaper alternative is to design a dual rate unknown input observer that estimates the states of the system asymptotically, and then estimates the attack d using the state estimates. The theory for single rate unknown input observers is well studied and can be found in [5], [4], [6], [7] and the references therein. In this section, we will extend the theory to design a dual rate observer to estimate the attack in Figure 4. Dual rate unknown input observers were briefly mentioned in [12], however, the authors assumed D_{11} and D_{12} to be equal to zero in (2), which changes the analysis and the conditions for existence of the observer and how the attack is estimated. We consider the state space description $\tilde{P}$ in (3), assuming without loss of generality that $K = 0$ and that the measurements are noise free. $\tilde{P}$ is then represented by

$$\begin{aligned} x(k+1) &= A_d x(k) + B_{1d} d(k) \\ \tilde{y}(k) &= \tilde{C}_1 x(k) + \tilde{D}_{11} d(k). \end{aligned} \quad (9)$$

We consider an observer of the form

$$\hat{x}(k+1) = E\hat{x}(k) + L\tilde{y}(k), \quad (10)$$

where E and L are matrices to be designed.

Definition 7. *The system* (10) *is said to be an unknown input dual rate observer with rate T/m if $\hat{x}(k) - x(k) \to 0$ as $k \to \infty$, regardless of the input.*

We note that the observer in (10) does not depend on the input to the system (9). To choose the observer matrices E and L, we examine the estimation error

$$\begin{aligned} e(k+1) &= \hat{x}(k+1) - x(k+1) \\ &= E\hat{x}(k) + L\tilde{y}(k) - A_d x(k) - B_{1d} d(k) \\ &= Ee(k) + (E - A_d + L\tilde{C}_1)x(k) \\ &\quad + (L\tilde{D}_{11} - B_{1d})d(k). \end{aligned}$$

In order to force the error to go to zero, regardless of the values of $x(k)$ and $d(k)$, E and L must simultaneously satisfy

$$L\tilde{D}_{11} = B_{1d} \quad (11)$$

$$E = A_d - L\tilde{C}_1 \quad (12)$$

such that E is stable. There exists a matrix L that satisfies (11) if and only if B_{1d} is in the space spanned by the rows of $\tilde{D}_{11}$, which is equivalent to

$$rank\left(\begin{bmatrix} B_{1d} \\ \tilde{D}_{11} \end{bmatrix}\right) = rank(\tilde{D}_{11}). \quad (13)$$

Necessary and sufficient conditions for the existence of E and L that satisfy (11) and (12) are that $\tilde{P}$ is strongly detectable (i.e., all zeros of $\tilde{P}$ are strictly stable), and that (13) holds.

The strong detectability condition is inherited from the conditions of existence of UIO for single rate systems. We know that using dual rate sampling guarantees that at most one zero exists and is at $\lambda = 1$; therefore, checking $\tilde{P}$ for this zero is sufficient to check for the strong detectability of $\tilde{P}$, as long as Assumptions 1 and 2 are met. Furthermore, strong observability of $\tilde{P}$, which is a more strict property, is guaranteed if G_{11} is tall and does not have a zero at the origin, as long as Assumptions 1 and 2 are met. Now to ensure the solvability of (11), m is chosen large enough until (13) holds.

Once a state observer is constructed, we can obtain an estimate of the attack by first rearranging (9) as

$$\begin{bmatrix} x(k+1) - A_d x(k) \\ \tilde{y}(k) - \tilde{C}_1 x(k) \end{bmatrix} = \begin{bmatrix} B_{1d} \\ \tilde{D}_{11} \end{bmatrix} d(k). \quad (14)$$

Since both Assumption 1 and (13) hold, then $\begin{bmatrix} B_{1d} \\ \tilde{D}_{11} \end{bmatrix}$ is full column rank, and there exists a matrix R such that

$$R \begin{bmatrix} B_{1d} \\ \tilde{D}_{11} \end{bmatrix} = I,$$

where I has the appropriate dimension. Left multiplying (14) by R and using $\hat{x}(k)$ instead of $x(k)$, we find the estimate of the attack to be

$$\hat{d}(k) = R \begin{bmatrix} \hat{x}(k+1) - A_d \hat{x}(k) \\ \tilde{y}(k) - \tilde{C}_1 \hat{x}(k) \end{bmatrix}.$$

Since

$$e(k) \to 0 \text{ as } k \to \infty,$$

$\hat{d}(k)$ will asymptotically approach $d(k)$. (one sample period of the original single rate) Note that there is a single time-step delay (one T) in computing the attack estimate. In case of single rate sampling, there will be at most $n+1$ time-steps delay ($(n+1)T$) where n is the dimension of the vector x in (9), if the observer exists (i.e. if the single rate system is strongly detectable) [5].

a) Example - Automatic Voltage Regulator (AVR)

We revisit the AVR example in section IV-B.2.a. Since the system has an unstable zero, we know that we cannot construct a single rate unknown input observer to estimate actuator attacks even if we use an arbitrary large number of measurements. Next we know from the previous section that dual rate sampling at a rate of $T/m = 0.25$ sec removes the unstable zero when viewed from the lifted domain. The resulting open loop state space representation after lifting is presented in (8), and we can see that the condition $rank\left(\begin{bmatrix} B_d \\ \tilde{D} \end{bmatrix}\right) = rank(\tilde{D})$ is satisfied. We construct a dual rate unknown input observer of the form (10), i.e.,

$$\hat{x}(k+1) = E\hat{x}(k) + L\tilde{y}(k),$$

where E and L satisfy (11), (12) and (8). Using MATLAB solver, we find E and L to be

$$E = \begin{bmatrix} 0.074 & 2.91 & 35.70 & -3413.41 \\ -0.013 & -0.47 & -5.00 & 9003.32 \\ 0.013 & 0.51 & 6.21 & -734.21 \\ 0.0063 & 0.25 & 3.02 & -5.42 \end{bmatrix}, \; L = \begin{bmatrix} 0.7118 & -0.0292 \\ -1.8040 & 0.0033 \\ 0.1522 & -0.0055 \\ 0.0037 & -0.0025 \end{bmatrix}.$$

We note that for the above AVR example, sampling faster using $m = 2$ was sufficient to estimate the states and the attack asymptotically, while in section IV-B.2.a, we saw that $m = 5$ was needed to accurately reconstruct the states and the attack in each period T. This observation makes sense as it means more measurements are needed for accurate estimation in each period T versus asymptotic estimation.

V. Conclusion

We posed the problem of estimating signal attacks injected into the actuators or sensors of control systems. We showed that there exists a trade-off between attack rejection and estimation, and that the estimator design depends on the controller used. We used dual rate sampling to enhance the detectability of the attack and we provided three methods to design the estimator. Method 1 solves a model matching problem subject to causality constraints. Method 2 exploits dual rate sampling to accurately reconstruct the unknown input. Method 3 uses a dual rate unknown input observer. Using dual rate sampling, necessary and sufficient rank and zero location conditions to check the existence of the observers in methods 2 and 3 were provided. Once these conditions are satisfied, then the attack can be estimated with at most a single time-step delay. This work shows again the importance of studying the security problem in the SD framework, and the power of using dual rate sampling to defend against signal attacks. A future research direction is to study dual rate unknown input observers when noise is present in the measurements. Optimal single rate delayed UIOs were discussed in [6] (minimizing mean square error), but to our knowledge no results exist for the dual rate UIOs.

References

[1] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs," *IEEE Control Systems*, vol. 35, no. 1, pp. 93–109, Feb 2015.

[2] M. Naghnaeian, N. Hirzallah, and P. G. Voulgaris, "Dual rate control for security in cyber-physical systems," in *2015 54th IEEE Conference on Decision and Control*, December 2015, pp. 1415–1420.

[3] N. H. Hirzallah and P. G. Voulgaris, "On the computation of worst attacks: a lp framework," in *2018 Annual American Control Conference (ACC)*, June 2018, pp. 4527–4532.

[4] S. Sundaram and C. N. Hadjicostis, "Delayed observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 52, no. 2, pp. 334–339, Feb 2007.

[5] S. Sundaram, *Fault-Tolerant and Secure Control Systems*. Online, Univesity of Waterloo., 2012.

[6] S. Sundaram and C. N. Hadjicostis, "Optimal state estimators for linear systems with unknown inputs," in *Proceedings of the 45th IEEE Conference on Decision and Control*, Dec 2006, pp. 4763–4768.

[7] J. Chen and R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Springer Publishing Company, Incorporated, 2012.

[8] W. Kratz, "Characterization of strong observability and construction of an observer," *Linear Algebra and Its Applications*, vol. 221, pp. 31 – 40, 1995.

[9] M. Hautus, "Strong detectability and observers," *Linear Algebra and Its Applications*, vol. 50, pp. 353 – 368, 1983.

[10] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice-Hall, 1995.

[11] T. Chen and B. Francis, *Optimal Sampled-Data Control Systems*. Springer, 1995.

[12] T. Mita, Y. Chida, Y. Kaku, and H. Numasato, "Two-delay robust digital control and its applications-avoiding the problem on unstable limiting zeros," *IEEE Transactions on Automatic Control*, vol. 35, no. 8, pp. 962–970, Aug 1990.

[13] X. Qi, M. V. Salapaka, P. G. Voulgaris, and M. Khammash, "Structured optimal and robust control with multiple criteria: a convex solution," *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1623–1640, Oct 2004.

[14] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*. USA: John Wiley & Sons, Inc., 2005.

[15] N. H. Hirzallah, "Security of cyber-physical systems: A control-theoretic perspective," Ph.D. dissertation, Electrical and Computer Eng., Univ. of Illinois at Urbana-Champaign, 2018.

[16] P. G. Voulgaris and B. Bamieh, "Optimal H_∞ control of hybrid multirate systems," in *Proceedings of the 31st IEEE Conference on Decision and Control*, December 1992, pp. 457–462.

[17] M. F. Sagfors, H. T. Toivonen, and B. Lennartson, "H_∞ control of multirate sampled-data systems: A state-space approach," *Automatica*, vol. 34, no. 4, pp. 415 – 428, 1998.

[18] T. Hagiwara and M. Araki, "Design of a stable state feedback controller based on the multirate sampling of the plant output," *IEEE Transactions on Automatic Control*, vol. 33, no. 9, pp. 812–819, September 1988.