Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 48-27 (2015) 260–266

Model Invalidation for Switched Affine Systems with Applications to Fault and Anomaly Detection

Farshad Harirchi and Necmiye Ozay

Electrical Engineering and Computer Science Department
University of Michigan, Ann Arbor, MI 48109 USA
(e-mail: harirchi,necmiye@umich.edu).

Abstract: In this paper, the model (in)validation problem is addressed for the class of switched state space models. We pose the model invalidation problem as a mixed-integer linear program and solve it using the state-of-the-art MILP solvers. Model invalidation is mainly utilized to build trust in the models obtained from system identification. However, we turn our attention to solve another important class of problems using model invalidation approach proposed in this paper. It is shown that the model invalidation approach can be utilized to detect any general fault in cyber-physical systems. Moreover, it is illustrated that knowing the fault model can reduce the complexity of fault detection approach proposed here, if the fault and system model satisfy certain conditions.

Keywords: model invalidation, switched systems, fault detection, radiant systems.

1. INTRODUCTION

Cyber-physical systems are combinations of physical processes and embedded computers that collect data from these processes through sensors and control these processes in a closed loop manner. With the increase in data acquisition and storage capacity and the decrease in sensor costs, it is possible to collect large amounts of data during the operation of complex cyber-physical systems. For instance, "a four-engine jumbo jet can create 640 terabytes of data in just one crossing of the Atlantic Ocean" Rajah (2014). As discussed in Sznaier et al. (2014), this exponential growth in the data collection capabilities is a major challenge for systems and control community. Sensor-/information-rich networked cyber-physical systems, from air traffic or energy networks to smart buildings, are getting tightly integrated into our daily lives. As such, their safety-criticality increases. For such systems, it is crucial to detect faults or anomalies in real-time to support the decision-making process and to prevent potential large-scale failures. It is also important to obtain accurate models for these systems that can be used both for control design and later for monitoring the system.

Switched affine state-space models provide a convenient means to model many cyber-physical systems. In this paper, we consider two problems related to switched affine models: (i) model invalidation; (ii) fault and anomaly detection. In model invalidation problem, one starts with a family of models (i.e., a priori or admissible model set) and experimental input/output data collected from a system (i.e., a finite execution trace) and tries to determine whether the experimental data can be generated by one of the models in the initial model family. It was originally proposed as a way to build trust in models obtained through a system identification step or discard/improve them before using these models in robust control design Smith and Doyle (1992). There are more recent results applying model invalidation ideas for nonlinear systems Prajna (2006), and switched auto-regressive models Cheng et al. (2012); Ozay et al. (2010, 2014). Moreover, it is demonstrated via examples in Ozay et al. (2014) that model invalidation algorithms can be used for anomaly detection, where anomaly is roughly treated as anything that cannot be explained by the a priori model set. In this paper, we formalize the connection between anomaly/fault detection and model invalidation. Moreover, we show how model invalidation algorithms can be used in a receding horizon manner when fault models exist.

Fault detection techniques are developed in different communities Miljkovic (2011). A category of fault detection approaches is signal and data processing based, which utilizes techniques from pattern recognition Diallo et al. (2005) and spectrum analysis Isermann (2005, 2006) or simple algorithms of trend or limit checking of the signal Verron et al. (2010). Fault detection in control community has been investigated from the process model perspective. The methods developed in this community are based on residual generation and evaluation (see e.g., Chow and Willsky (1984); De Persis and Isidori (2001); Frank and Ding (1997)). Another set of approaches is based on designing the state or output observers and using the estimation error, or innovation, as the residual for the detection of the fault Frank (1990). Parameter estimation techniques have been also utilized as residual generators where the difference between physical parameters of the system and the estimated parameters from data are used as residuals Venkatasubramanian et al. (2003). Finally, utilizing neural networks for black box modeling of static or dynamical

⋆ This work is supported in part by DARPA grant N66001-14-1-4045.

2405-8963 © 2015, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2015.11.185

systems and comparing the output of that model with the experimental data is another approach for fault detection Isermann (2006). In this work, we propose a new approach to fault detection from a controls perspective, which is based on model invalidation techniques for switched state-space models.

In Ozay et al. (2014), a model invalidation approach for switched auto-regressive models is proposed based on polynomial optimization and relaxation technique. Unavailability of switching sequence for measurement and noise in the input/output measurements render the model invalidation for switched systems challenging Ozay et al. (2014). The model invalidation problem setup we consider in the present paper is closely related to that in Ozay et al. (2014), however there are two important differences. First, we consider switched state-space models as opposed to switched autoregressive models. Arguably, when modeling a system using first principles, state-space representation is quite natural. Therefore, state-space models are more commonly used for modeling cyber-physical systems and developing model invalidation techniques for this class of systems is important. Second, we proceed by recasting the model invalidation problem as the feasibility check of a mixed integer linear programming (MILP) problem as opposed to the convex relaxation approach. Although from a complexity point of view convex relaxations are appealing, we experimentally observed that state-of-the-art MILP solvers (e.g., CPLEX (2009)) perform favorably on average without the need for a relaxation.

1.1 Contributions and Structure

The contributions of this paper are: (i) to propose a model invalidation approach for the class of switched affine (SWA) models, (ii) to utilize model invalidation as a tool for anomaly and fault detection in cyber-physical systems, and (iii) to apply proposed method on the fault detection of radiant systems in smart buildings.

The structure of the paper is as follows: The proposed approach for model invalidation of switched affine models is discussed in Section 2. In Section 3, model-based fault and anomaly detection as well as the relation to model invalidation is described. Finally, academic and practical examples are provided in Section 4. The paper, then is concluded with discussion and future directions in Section 5.

1.2 Notation

Let $\mathbf{x} \in \mathbb{R}^n$ denote a vector and $\mathbf{x}^i$ indicate its ith element. Also, let $\mathbf{M} \in \mathbb{R}^{n\times m}$ represent a matrix and $\mathbf{M}^{i,j}$ indicate the element on ith row and jth column of the matrix $\mathbf{M}$. I_n denotes the identity matrix of size n. The infinity norm of a vector $\mathbf{x}$ is denoted by $\|\mathbf{x}\|_\infty \doteq \max_i \mathbf{x}^i$. The set of positive integers up to n is denoted by $\mathbb{Z}_n^+$, and the set of non-negative integers up to n is denoted by $\mathbb{Z}_n^0$.

2. SWA MODEL (IN)VALIDATION

In this section, we state the invalidation problem, and provide a tractable approach to solve it.

2.1 Problem Definition

We consider switched affine (SWA) systems of the form:

$$G = (\mathcal{X}, \mathcal{E}, \mathcal{U}, \{G_i\}_{i=1}^s) \tag{1}$$

where $\mathcal{X} \subset \mathbb{R}^n$ is the set of states, $\mathcal{E} \subset \mathbb{R}^{n_y}$ is the set of measurement noise values, $\mathcal{U} \subset \mathbb{R}^{n_u}$ is the set of inputs and $\{G_i\}_{i=1}^s$ is a collection of s modes where for all $i \in \mathbb{Z}_s^+$, the i^{th} mode is an affine model $G_i = (\mathbf{A}_i, \mathbf{B}_i, \mathbf{f}_i, \mathbf{C}_i, \mathbf{D}_i)$.The evolution of G is governed by:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}_{\sigma(k)}\mathbf{x}(k) + \mathbf{B}_{\sigma(k)}\mathbf{u}(k) + \mathbf{f}_{\sigma(k)}, \\ \mathbf{y}(k) &= \mathbf{C}_{\sigma(k)}\mathbf{x}(k) + \mathbf{D}_{\sigma(k)}\mathbf{u}(k) + \boldsymbol{\eta}(k), \end{aligned} \tag{2}$$

where $\mathbf{x}(k) \in \mathcal{X}$ is the state, $\mathbf{u}(k) \in \mathcal{U}$ is the control input, $\mathbf{y}(k) \in \mathbb{R}^{n_y}$ is the output, and $\boldsymbol{\eta}(k) \in \mathcal{E}$ is the measurement noise at time k. Here, $\sigma(k) \in \mathbb{Z}_s^+$ indicates the active mode at time k, that is, if $\sigma(k) = i$ the state evolves with respect to the dynamics of G_i. Throughout the paper we take $\mathcal{X}$, $\mathcal{E}$ and $\mathcal{U}$ to be infinity norm balls. That is, we let $\mathcal{X} = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq M\}$, $\mathcal{E} = \{\boldsymbol{\eta} \mid \|\boldsymbol{\eta}\|_\infty \leq \epsilon\}$ and $\mathcal{U} = \{\mathbf{u} \mid \|\mathbf{u}\|_\infty \leq U\}$, where M, ϵ and U are given constants. Usually physical constraints on the system impose the bounds M and U. If no such bound is known, they can be taken to be infinite. On the other hand, the bound ϵ on the measurement noise value is based on the accuracy of the sensors and always assumed to be finite.

Remark 1. We do not consider process noise in the SWA system defined above, but the results in this paper can be extended to the systems with process noise, simply by adding variables to the problem.

In order to state the model invalidation problem, we first define the behavior of an SWA system.

Definition 1. The *N-truncated behavior* associated with an SWA system G is the set of all length-$N+1$ input-output trajectories compatible with G, given by the set

$$\mathcal{B}_{swa}^N(G) := \big\{\{\mathbf{u}(k), \mathbf{y}(k)\}_{k=0}^N, \mid \mathbf{u}(k) \in \mathcal{U} \text{ and } \exists \mathbf{x}(k) \in \mathcal{X}, \\ \sigma(k) \in \mathbb{Z}_s^+, \boldsymbol{\eta}(k) \in \mathcal{E} \text{ for } k = 0, \ldots, N \text{ s.t. (2) holds}\big\}.$$

With slight abuse of terminology, we will call $\mathcal{B}_{swa}^N(G)$ just the *behavior* of the system G.

Now we can state the model invalidation problem for SWA systems. Roughly speaking, given an input-output data sequence and a switched affine model, model invalidation problem is to determine whether or not the data is compatible with the model. This can be formally stated in terms of behaviors as follows:

Problem 1. Given $\big\{\mathbf{u}(k), \mathbf{y}(k)\big\}_{k=0}^N$, an input-output sequence, and a switched affine model G, determine whether or not the input-output sequence is contained in the behavior of G. That is, whether or not the following is true

$$\big\{\mathbf{u}(k), \mathbf{y}(k)\big\}_{k=0}^N \in \mathcal{B}_{swa}^N(G). \tag{3}$$