Journal of Process Control 32 (2015) 51–63

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Resilient plant monitoring systems: Techniques, analysis, design, and performance evaluation

Humberto E. Garcia [a], Semyon M. Meerkov [b,*], Maruthi T. Ravichandran [b]

[a] Idaho National Laboratory, P.O. Box 1625, Idaho Falls, ID 83415-3675, USA
[b] Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:
Received 19 February 2015
Received in revised form 1 May 2015
Accepted 4 May 2015
Available online 29 May 2015

Keywords:
Sensor networks
Malicious attacks
Data quality acquisition
Process variable and plant condition assessment
Rational controllers
Decomposition with knowledge fusion
Resilient monitoring of power plants

ABSTRACT

Resilient monitoring systems (RMS) are sensor networks that degrade gracefully under malicious attacks on their sensors, causing them to project misleading information. This paper develops techniques to ensure resiliency, namely: active data quality acquisition, process variable and plant condition assessments, sensor network adaptation, and plant decomposition with knowledge fusion. Based on these techniques, we design a RMS for power plants and investigate its performance under various cyber-physical attacks. In all scenarios considered, the system offers effective protection against misleading information and identifies the plant condition – normal or anomalous – in a reliable and timely manner.

1. Introduction

Resilient plant monitoring systems is a relatively new area of research. In this section, we briefly characterize these systems, describe a specific scenario addressed, and outline the techniques developed in this work.

1.1. What is a resilient plant monitoring system?

Plant monitoring systems are wired or wireless sensor networks intended to measure process variables (e.g., temperature, pressure, flow rates, etc.), analyze them, and inform the plant operator about the plant conditions – normal or anomalous. Based on this information, the operator takes corrective actions, if needed. When some of the sensors are captured by an attacker, forcing them to project misleading information (possibly, statistically unrelated to the actual values of process variables), the identified plant conditions could be erroneous. This may lead to wrong actions on the part of the operator and, possibly, a disaster. To prevent this situation, the monitoring system must possess a capability of autonomously identifying the attacked sensors and mitigating their effect (by discounting or disregarding completely the data they project). Although the loss of sensors may lead to *degradation* of plant condition assessment, in a well-designed system this degradation should be "proportional" to the severity of the attack, i.e., *graceful*. Plant monitoring systems that possess such a property are referred to as *resilient*.

This paper is devoted to developing techniques that can be used to ensure resiliency, analyzing their properties and, on this basis, designing and evaluating the performance of a resilient monitoring system (RMS). A specific application, in terms of which the development is carried out, is a simplified model of a power plant, although a similar approach can be used for other applications as well.

1.2. Scenario and problem addressed

Briefly, the scenario considered in this paper is as follows:

- The monitored plant process variables, $\mathbf{V}_i$, $i = 1, \ldots, M$, are characterized by probability density functions (pdf's) $f_{\tilde{V}_i}(\tilde{v}_i)$, $i = 1, \ldots, M$. In practice, the *status* of the process variables is often characterized as being Normal (N) or Anomalous (A). The latter could be, for instance, Low (L) or High (H). In this case, $f_{\tilde{V}_i}(\tilde{v}_i)$ induces a random event with the outcomes in $\{L_{V_i}, N_{V_i}, H_{V_i}\}$, $i = 1, \ldots, M$. With

* Corresponding author. Tel.: +1 7347636349.
E-mail addresses: Humberto.Garcia@inl.gov (H.E. Garcia), smm@umich.edu (S.M. Meerkov), marutrav@umich.edu (M.T. Ravichandran).

http://dx.doi.org/10.1016/j.jprocont.2015.05.001
0959-1524/

a slight abuse of terminology, we refer to this event (and similar events throughout this paper) as a discrete random variable, V_i, $i = 1, \ldots M$, with the probability mass function (pmf), $p[V_i]$, defined on the universal set $\Sigma_{V_i} = \{L_{V_i}, N_{V_i}, H_{V_i}\}$, $i = 1, \ldots, M$.

- The plant, **G**, is also characterized by its status, which is a discrete random variable, G, with the pmf $p[G]$ defined by the pmf's of process variables and taking values on $\Sigma_G = \{N_G, A_G\}$, where N_G and A_G denote the normal and anomalous plant statuses, respectively. Depending on the plant, the anomalous status can be further characterized by specific anomalies, e.g., boiler insulation damaged, turbine malfunctioning, etc. In each status, plant dynamics may be different, e.g., described by different transfer functions.
- Each process variable, $\mathbf{V}_i$, is monitored by a sensor, $\mathbf{S}_i$ (multiple sensors of a process variable are also considered in the sequel). If a sensor is under attack, its projected data may have a pdf, $f_{\tilde{S}_i}(\tilde{s}_i)$, statistically unrelated to $f_{\tilde{V}_i}(\tilde{v}_i)$. In this situation, utilizing the sensor data in order to assess the process variable may lead to a pmf, $\hat{p}[V_i]$, qualitatively different from $p[V_i]$. For instance, $\hat{p}[V_i]$ may indicate that the process variable is Normal, while in reality it is Low or High.
- The plant status assessment is based on the process variable assessments, $\hat{p}[V_i]$, $i = 1, \ldots, M$, and is quantified by a pmf denoted as $\hat{p}[G]$, $G \in \{N_G, A_G\}$. Since, as indicated above, the process variable assessments may be erroneous, $\hat{p}[G]$ may be quite different from the actual $p[G]$ and, thus, lead to erroneous actions by the plant operator.

In this scenario, the *optimal* resilient monitoring system must be able to identify the status of the plant, **G**, in such a manner that the "distance" between the estimated and the actual pmf's, $\hat{p}[G]$ and $p[G]$, is minimized, as quantified by an appropriate measure of distance between the two pmf's. While this paper is not intended to resolve this issue, the problem addressed here is: *design a plant monitoring system that degrades gracefully under an attack (i.e., is resilient), and demonstrate that it performs favorably in comparison with a non-resilient one (as quantified by a measure of resiliency based on the Kullback-Leibler divergence* [1]).

1.3. Related literature

The literature related to the topic of this paper can be classified into five groups. The first one is devoted to foundational issues, where the problems of resilient monitoring and control are motivated and formulated, [2–6]. The second group includes publications on control-theoretic methods for attack identification and alleviation, [7–11]. In these publications, the authors consider LTI systems with a given state space realization (A, B, C, D) and disturbances interpreted as attack vectors. The problem addressed is to identify the attack and, if possible, mitigate its effect, for instance, by designing a controller that makes the closed-loop system invariant with respect to the disturbance attack. The main difference of the current work is that the plant may be either normal or anomalous (i.e., described by several state space realizations), and the problem is to identify which plant status indeed takes place, in spite of the misleading information projected by the sensors.

The third group consists of publications on fault tolerant control, [12–14]. In these works, it is assumed that a closed-loop system has multiple sensors and actuators, some of which could be faulty due to natural or malicious causes. The typical problem here is to determine the conditions (e.g., the number of sensors and actuators) under which the closed-loop system performance is maintained without degradation. The difference of the current work is that, although multiple sensors may be present, the goal is to determine the status of the plant and, if otherwise impossible, tolerate degradation.

The fourth group consists of research on monitoring communication channels in order to capture anomalous traffic and correlate it with a possible attack, [15–19]. In terms of the current work, this implies the identification of *DQ*. While the results of these publications may be useful for resilient plant monitoring, they do not provide methods for process variable and plant condition assessment pursued in the current work.

The fifth group consists of papers on identification of and protection against data injection attacks intended to mislead state estimation algorithms, [20–24]. The emphasis of the research here is on determining optimal positions of "known-secure" sensors, which prevent the damage of the attack, or on utilizing game-theoretic approaches as quantitative techniques for risk management.

Our preliminary results on RMS have been reported in conference presentations [25–29] and summarized in article [30]. The current paper, along with reviewing and extending these results, introduces a decentralized RMS based on plant decomposition with knowledge fusion, as a means for combating the curse of dimensionality. As a result, we design a decentralized RMS for power plants and investigate its performance under various cyber-physical attack scenarios. Note that an ideologically similar but technically different approach to resilient monitoring of a chemical reactor system has been reported in a recent paper [31].

1.4. Contributions of this work: Techniques developed and resilient monitoring system designed

The techniques developed in this work are as follows:

- The "trustworthiness" of a sensor is quantified by a parameter referred to as *data quality* (*DQ*), which takes values on [0, 1], with 1 indicating that the sensor is totally trustworthy and 0 not trustworthy at all. To identify *DQ*, we develop an *active data quality acquisition procedure*, whereby probing signals are applied to process variables, and the level of disagreement between the anticipated and the actual response of the sensors is used to quantify their *DQ*'s.
- The estimates of process variables pmf's, $\hat{p}[V_i]$, $i = 1, \ldots, M$, are calculated based on the data projected by the sensors and their *DQ*'s. Since *DQ* is not a statistical quantity, classical statistics cannot be used for this purpose. Therefore, we introduce a model of the *DQ*'s effect on the coupling between sensors data and process variables and, using this model, develop the so-called *h-procedure* (which is a modified stochastic approximation algorithm [32]). Analyzing this procedure, we show that it converges to a steady state defined by the *DQ*'s. Specifically, if $DQ = 1$, it converges to the actual process variable pmf; as *DQ* tends to 0, the steady state of the h-procedure converges to a uniform pmf, implying that in this limit the sensor measurements carry no information at all. For all other *DQ*'s, the conditional pmf of V_i given the sensor data is an affine function of *DQ*. When multiple sensors monitor a process variable, the *Dempster-Shafer rule* [33] is used to combine the steady states of the h-procedures associated with each sensor.
- The estimate of the plant status pmf, $\hat{p}[G]$, is calculated based on the statistical plant model (typically given as a set of conditional pmf's $P[V_i|G]$, $i = 1, \ldots, M$, or a joint conditional pmf $P[V_1, V_2, \ldots, V_M|G]$), the estimates of the process variables pmf's, $\hat{p}[V_i]$, $i = 1, \ldots, M$, and the *Jeffrey rule* [34].
- The above assessments are carried out at each state of the sensor network, where the state is a vector of 1's and 0's, with 1 indicating that the corresponding sensor is taken into account for process variable assessment and 0 that it is not. The quality of each state is quantified by the entropy (i.e., the level of uncertainty) of either $\hat{p}[G]$ or $\hat{p}[V_i]$. The adaptation of the sensor network to the optimal state, i.e., the state with the smallest