Application of Dynamic System Models and State Estimation Technology to the Cyber Security of Physical Systems

Barry Horowitz, Kate Pierce
Systems and Information Engineering
University of Virginia
Charlottesville, VA, USA
Email: bh8e@virginia.edu, kmp7ef@virginia.edu

As advances in technology permit automatic control of more and more of the functions of physical systems, the opportunity for cyber attacks that include exploitation of such automation capabilities becomes a greater risk. For example, in the 2010 Stuxnet attack an embedded infection in control systems was used to successfully damage a large number of nuclear power related centrifuges in Iran. While the application of perimeter security technologies have been applied to help manage the likelihood of SCADA-based cyber attackers exploiting highly automated physical systems, successful attacks have occurred, and furthermore, perimeter solutions do not address important classes of insider and supply chain initiated attacks. As a result, it has been recognized that perimeter security needs to be augmented by other approaches for addressing potential cyber attacks [1].

Frequently, as a means for added operational assurance, highly automated physical systems include the presentation of system status information that permits human operators to take controlling actions when the automated system appears to be operating in an out-of-normal manner. For example, the operation of a turbine may be automatically controlled, but operators can observe critical information regarding the turbines operation, such as vibration levels, temperature, and rotation rate. If the operator observes measurements that are outside the designated region of proper operation, specific manual actions can be required of the operator in order to avoid undesirable consequences [2]. However, as was the case in the Stuxnet attacks, the cyber attacker can not only manipulate a physical systems performance through infections in its control system, but can also manipulate data presented to operators; data that can, when utilized within standard operating procedures, either stimulate inappropriate control actions or prevent needed control actions on an operators part. In the case of the turbine example, a successful cyber attack can result in indications to operators that would imply that all is well when it is not, or indications that would call for disruptive operator action when, in reality, none is required (e.g., unnecessarily shutting down the turbine). Note that it is quite typical for operator displays to be designed for simplicity, so that critical manual actions will not be delayed by human limitations related to viewing and interpreting too much information. As a result, physical systems typically include measurement and collection of information that could conceivably be used, but is not, for automation override decision making. For example, driving an automobile involves a driver monitoring a few of the many available engine state measurements that could be made available for viewing, but could confuse the driver while offering little, if any, benefit.

This paper presents an approach for addressing cyber attacks on physical systems that include purposeful manipulation of operator displays. The presented approach involves embedding security functions within the physical system being protected; functions that can be the basis for detection of inconsistent system dynamics data derived from measurements within the system that is being protected. In particular, the use of dynamic mathematical models of physical systems in combination with state estimation techniques is suggested as the basis for system architectures that can be employed to detect situations where information displays for system operators are being manipulated as part of a cyber attack. One can divide the states of a physical system into 3 classes: 1) those that are presented to operators for control purposes, but are considered as least trusted from a cyber security viewpoint; 2) those that can be measured and analyzed in segregated equipment from the equipment being used for measuring, analyzing and displaying of least trusted states, but are not used for operator assistance and are considered as more trusted; and 3) those that are not measured. The paper draws on dynamic state estimation techniques to develop, when feasible, estimates of the least trusted states values, and the variances of these estimates, from measurements of more trusted states. Systems that satisfy control system conditions for being observable satisfy the sufficient conditions for this cyber security solution. The paper shows how these estimates can provide the basis for detecting attacks on automatic control systems that include manipulation of data presented to operators, and how this approach can be used to manage system restoration. Theoretical results are presented for a range of physical system models, and a specific system model for an electrical generator is used to illustrate the performance one can achieve in an actual
application, including calibration of expected performance in terms of missed detections, false detections and delay time for detections of cyber attacks.

This material is based upon work supported in part by the U.S. Department of Defense through the Systems Engineering Research Center (SERC) under Contract H98230-08-D-0171. SERC is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

REFERENCES