INL/CON-11-21607
PREPRINT

‘Known Secure Sensor Measurements’ for Critical Infrastructure Systems: Detecting Falsification of System State

SERENE 2011

Miles McQueen
Annarita Giani

September 2011

The INL is a
U.S. Department of Energy
National Laboratory
operated by
Battelle Energy Alliance

'Known Secure Sensor Measurements' for Critical Infrastructure Systems: Detecting Falsification of System State

Miles McQueen and Annarita Giani*

Idaho National Laboratory and University of California at Berkeley

Abstract. This paper describes a first investigation on a low cost and low false alarm, reliable mechanism for detecting manipulation of critical physical processes and falsification of system state. We call this novel mechanism *Known Secure Sensor Measurements* (KSSM). The method moves beyond analysis of network traffic and host based state information, in fact it uses physical measurements of the process being controlled to detect falsification of state. KSSM is intended to be incorporated into the design of new, resilient, cost effective critical infrastructure control systems. It can also be included in incremental upgrades of already installed systems for enhanced resilience. KSSM is based on *known secure* physical measurements for assessing the likelihood of an attack and will demonstrate a practical approach to creating, transmitting, and using the known secure measurements for detection.

Keywords: Secure measurements, system state, control systems, cyber security, critical infrastructure.

1 Introduction

Software is a primary component of control systems which are used to operate our critical infrastructures such as the energy, water, chemical, transportation, and critical manufacturing sectors. To attain system resilience against disturbances it is critical to develop methods to enhance reliable state awareness. This discussion paper proposes a novel integrated hardware-software approach for quickly and effectively detecting adversarial manipulation of the physical infrastructure and the attacker's attempted falsification of system state. This detection mechanism is a step towards early response to disruptions. We present now a first investigation and the main concept and ideas. We plan to explore new automated adaptive responses based on this methodology.

The detection of network attacks has been studied for years. Many commercial Intrusion Detection Systems (IDSs) are used to provide defense in depth for IT systems. Due to the large amount of traffic in most IT systems, the issue of lowering the false positive rate without increasing the false negative rate cannot be overcome [1], [2]. This is a limit of current IDS systems.

A high false alarm rate is unacceptable for control systems due to the high costs required to respond to alerts. Discussions in the control system community focused on the idea that IDSs for control systems may have far fewer false alarms than that experienced by IT systems due to the greater regularity and reduced complexity of the messaging [3]. While work continues in this area, it is yet to be proven valid in anything other than artificial settings. Further, a continued focus on the network seems to ignore advantages that critical infrastructure control systems may provide for detecting intrusions.

One benefit is that control systems operate physical processes with several associated

* Miles.McQueen@inl.gov, agiani@eecs.berkeley.edu

measurements. These measurements are used locally for real-time control and reported back to the operators for monitoring and control. If an adversary wishes to manipulate the process without early detection by an operator, the attacker must deceive the operator regarding the current system state. This would involve the modification of measured values being sent back to the operators or the injection of false measurements. If the measurements could be protected, the system state could be known by the operator with high confidence, permitting effective action. There are known cryptographic techniques for assuring the security of a message so that manipulation, injection of a new message, or replay of a previous message will be detected [5, 4]. Unfortunately, many control systems have limited processor cycles available at the sensor for encrypting, sensor power may be limited, and in some cases the communication bandwidth will be in short supply. Consequently, a comprehensive solution for control systems which makes full use of standard cryptographic techniques is not generally practical.
We propose a low-cost technique that uses physical process measurements to build an intrusion detection engine faster and more accurate than what is currently feasible. Encrypting all measurements is not reasonable due to limited control system resources. Further, one cannot necessarily afford the time to encrypt before forwarding even a single message without negatively impacting the timely reception of the message at the controller. But the use of encryption is needed to provide enough known secure data to an engine to enhance attack detection. So the technique must be low cost, make limited demands on system resources, and must have flexible timing requirements.

2 Technical Objectives and Research Approach

Our objective is to investigate the value of KSSM for effective detection of unauthorized process manipulation and falsification of system state.

2.1 Targeted Facilities

We consider critical infrastructure control systems that currently lack robust cryptographic techniques, and have limited communication bandwidth and computational resources. Due to the long lifetimes of most control systems this description applies to most infrastructures. Thus, we anticipate a significant benefit from the deployment of the KSSM technique.
A hypothetical hybrid energy plant is shown in figure 1. This figure represents a hybrid energy production facility with three abstract layers. The lowest layer is the physical process which consists of a set of production units each of which consists of reactors, tanks, gas flows, coolers, heaters, valves and like physical components. The information layer, in the middle, is responsible for communication. The sensors in the physical layer communicate with control devices and commands are sent to the edge controllers that drive actuator behavior. The highest layer is represented by the primary functions of plant control, and security threat monitoring and alarming. These highest level functions make use of real time data feeds from the physical plant up through the communication layer, and may also make use of information derived over time through initial monitoring of the system (e.g. passive network discovery).
We assume that the attacker can compromise any of the components in the information layer without being detected as long as the attacker does not modify the sensor signals being transmitted back to the controller and the control room. KSSM is not designed to detect the system process exceeding its operational performance envelope, normal system monitoring is expected to detect that situation. Rather, KSSM is designed to reliably detect any attempt to falsify system state through manipulation of one or more of the sensor measurements being reported back to the control room.

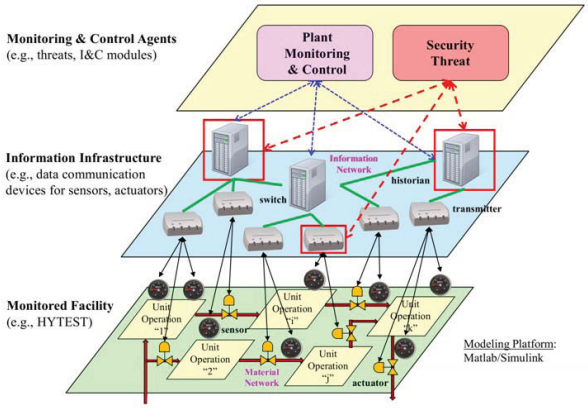

Fig. 1. Hybrid energy production facility.

2.2 KSSM System Hypotheses

We created the following four hypotheses to stay focused on the core issues in conceptualizing, designing, and validating a prototype of a KSSM system.

H1. A small set of sensor measurements, which are known to be secure, can significantly aid the operator and detection engine in more quickly and accurately identifying a cyber attack. A hybrid energy model, simulation, and detection engine will be used as an experimental foundation for measuring automated and human detection of process failures and falsification of system state when KSSM data is available.

H2. Some known secure measurements from randomly chosen sets of sensors providing data within selected time frames will harden the process against covert cyber attacks attempting to blind the operator and KSSM detection engine. Neither the enhanced operator effectiveness nor enhanced detection engine performance(gained from using a fixed set of known secure data) will be degraded by the changing and diverse sets of sensors selected for providing the known secure data.

H3. It is possible to create a very low cost, limited bandwidth, and highly secure measurement capture and communication channel for transmitting $k_i\%$ $(0 < k_i < 100)$ of a chosen sensor's physical measurements, end to end, from sensor to detection engine for analysis. The channel will involve adaptation of known cryptographic protocols to provide message and measurement integrity, and detection of replay attacks. Tradeoffs between cryptographic computational requirements at the sensor, power restrictions of a sensor, network bandwidth limitations, and the speed and accuracy of detection will be assessed in selecting specific cryptographic techniques for KSSM systems and establishing each KSSM sensor's appropriate value for k_i.

H4. Heuristics for selecting the set of sensors providing known secure sensor measurements can be developed which allow for the results of this research to be easily adapted for use in the design, implementation, and configuration of many diverse industrial control systems and infrastructures.

In a KSSM enabled infrastructure, the attacker will be unable to reliably falsify the process state to the control room operators.

2.3 KSSM Sensor

In any technique intended to protect against an intelligent adversary, one or more components must be trusted. We assume that the cryptographic sensor module, which includes

the hardware or software which receives and encrypts the physical measurements and identifiers gained directly from the sensor hardware (e.g. AD converter), is trusted.
Figure 2 represents a KSSM hardware enabled sensor. The signal from the AD converter is tapped off and available to the secure encryption module. This module, at some randomized time δ after the unencrypted measurement M_i is sent, forwards the encrypted version E_i of the measurement value to the KSSM detection module running on a control room computer. Whether or not the encrypted version of the plaintext measurement is sent depends on whether that particular sensor is currently selected by the KSSM control room module, and whether the secure encryption module selects it as one of the k_i of measurements for which a dual encrypted value will be expected. We note that these sensor functions may also be implemented in software and reside in the sensor or closest computational edge point. The KSSM detection algorithm in the control room, which must also be trusted, will compare the two versions of the measured value, unencrypted and encrypted, and trigger an alarm if there is any difference. For the exposition of the idea in this paper, we are making simplifying assumptions related to reliable transport of measurements, both plaintext and ciphertext, and to the reliability of the sensor and encryption hardware.

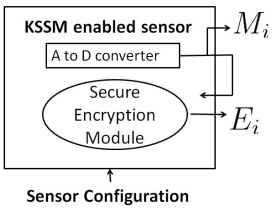

Fig. 2. Sample KSSM enabled sensor.

3 KSSM and Attacker Scenarios

While not required, we assume that all process sensors are KSSM hardware or software enabled, all encryption modules are secure from cyber attack, and the KSSM control room module, including the detection engine, are secure. Any other component in the system, including the entire information infrastructure layer may be assumed to be compromised by an attacker.

3.1 Attack Scenarios

In this section we present different possibilities for the system and the attacker.
Figure 3 presents two scenarios. The system on scenario 1 has no KSSM sensors available. The adversary has compromised choke points in the communication network and is behaving as a man in the middle by preventing all valid sensor signals to the control room and replacing them with corrupted, signals C_i. This deception could be done, as Stuxnet partially demonstrated, through collection and then replay of sensor measurement data. The attacker is now able to manipulate the process as desired while the operators remain completely unaware. This level of attack may be undetectable without KSSM and will leave the operators completely blind to the actual system state. KSSM sensors are available in scenario 2 and the attacker is choosing to corrupt only those signals which he knows are not providing encrypted values back to the detection engine. If the attacker

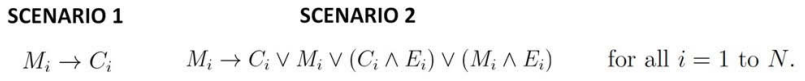

Fig. 3. M_i is the measurement from sensor i. C_i is the corrupted version of measurement M_i. E_i is the encrypted version of measurement M_i. $\vee$ and $\wedge$ are the logic disjunction and conjunction.

corrupted the signals for which an encrypted version was sent at a later time then the corruption would be instantly recognized by the KSSM detector. Given that the encrypted signals are sent at some δ time after the unencrypted signal the attacker can only know which sensors will provide the encrypted signal by observing the network traffic for some period of time.
The attack in figure 4 consists in corrupting signals and blocking some of the encrypted versions. In fact attack would be easily detectable if the encrypted version of the measurement reached the detection engine and be compared against the corrupted version.

A way to make the above attack more difficult is to *periodically and unpredictably*

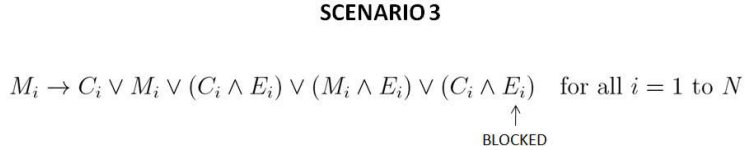

Fig. 4. Attacker identifies which sensors are providing encrypted versions of measurements. During the attack only a few of the sensors are blocked.

modify the subset of sensors providing encrypted values. This ongoing and unpredictable selection of new sensors (and deselection of others) may be based on current system state, or communication network topology (for example let us not select our sensors such that their measurements all go through the same router).

3.2 KSSM Control Room Module

The KSSM module residing in the control room is represented in figure 5. It is responsible for modifying the subset of KSSM-enabled sensors which perform encryption, and is also responsible for detecting attacks. Many functions are needed to provide these capabilities and we will very briefly describe only the highest level functions.

The *system analyzer* receives input from network discovery tools which can both reside on the system and operate in real time, or can be one time only devices used during a phase such as system acceptance testing. It develops simplified models of the communication network to aid the sensor selection function in choosing smart subsets of sensors.

The *signal analyzer* is responsible for analyzing the sensor measurements that are provided to the control room, and alarming when appropriate. When encrypted and associated unencrypted values do not match then an alarm will be set; if some number of requested encrypted values do not arrive in a timely fashion, and are distributed over a variety of communication paths then it may be appropriate to raise an alarm based on probabilistic assessment of likely communication and sensor failures. Other conditions and analyses for alarming need to be explored.

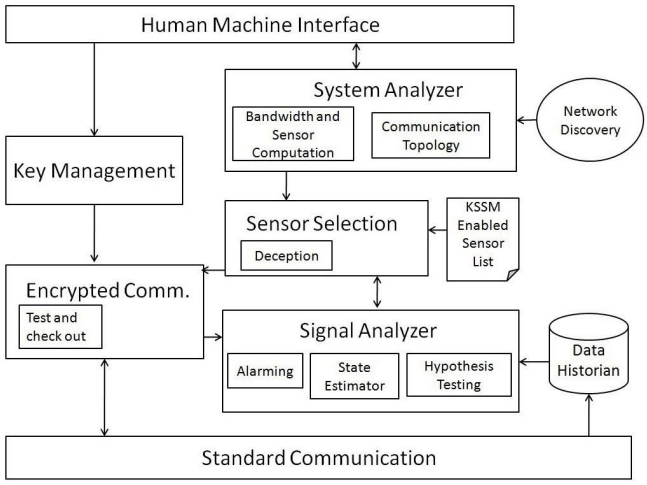

Fig. 5. Block diagram of KSSM module residing in the control room.

The *sensor selection algorithm* will incorporate what is known about the communication topology and the failure rates of all components within the system. The failure rates may be based on empirical data or models built into the algorithm. Further some understanding of the limits on computation cycles available, sensor power restrictions, and limitations in communication bandwidth will be incorporated to aid, not only the selection of a new subset of sensors for KSSM but also the selection for each chosen sensor of the k_i of measurements that will be encrypted and forwarded. Sensor selection and the percent of measurements for which dual encrypted values are required may also be made selectable by the operators so that they have control in limiting the sensor processor cycles, sensor power consumption, and communication bandwidth utilized by KSSM

The *cryptographic functions* will be adopted from currently well understood cryptographic components and systems. The KSSM-enabled sensor list is needed so that sensor selection can accommodate systems that are slowly being upgraded with KSSM-enabled sensors. And the KSSM user interface will be separate from all other devices in the control room in order to provide as much hardening against attack as possible.

4 Conclusion

The concept of Known Secure Sensor Measurements has been presented. The idea has been evaluated against potential attacker behavior and seems to have merit in providing early and reliable detection of attacker's possible attempts at falsification of process state. We are proceeding in our analysis and design of KSSM systems, and intend to initially validate through simulations in which effectiveness will be assessed through measuring the reduced time to detect falsification of system state. Eventually KSSM capability will be demonstrated on live control systems responsible for running our critical infrastructure facilities.

References

1. Stefan Axelsson: *The base-rate fallacy and the difficulty of intrusion detection.* ACM Transactions on Information and System Security (TISSEC), vol.3 n.3, pp.186–205, (1981)

2. David J. Fried and Isaac Graf and Joshua W. Haines and Kristopher R. Kendall and David Mcclung and Dan Weber and Seth E. Webster and Dan Wyschogrod and Robert K. Cunningham and Marc A. Zissman: *Evaluating intrusion detection systems: The 1998 darpa off-line intrusion detection evaluation*. In Proceedings of the 2000 DARPA Information Survivability Conference and Exposition, pp.12-26, (2000)
3. Ondrej Linda, Todd Vollmer, Milos Manic: *Neural Network Based Intrusion Detection System for Critical Infrastructures*. In Proceedings of International Joint Conference on Neural Networks, pp.1827–1834, (2009)
4. Mark Stamp: *Information Security*. 2nd edition. John Wiley and Sons, Chapters 3-5, and 9, (2011)
5. Niels Ferguson and Bruce Schneier and Tadayoshi Kohno: *Cryptography Engineering*. Chapters 3- 7, (2010)