Integrated Anomaly Detection for Cyber security of the Substations

ARTICLE *in* IEEE TRANSACTIONS ON SMART GRID · JUNE 2014

Impact Factor: 4.25 · DOI: 10.1109/TSG.2013.2294473

1 AUTHOR:

Junho Hong
ABB USCRC
7 PUBLICATIONS **59** CITATIONS

SEE PROFILE

Integrated Anomaly Detection for Cyber Security of the Substations

Junho Hong, *Student Member, IEEE*, Chen-Ching Liu, *Fellow, IEEE*, and Manimaran Govindarasu, *Senior Member, IEEE*

***Abstract*—Cyber intrusions to substations of a power grid are a source of vulnerability since most substations are unmanned and with limited protection of the physical security. In the worst case, simultaneous intrusions into multiple substations can lead to severe cascading events, causing catastrophic power outages. In this paper, an integrated Anomaly Detection System (ADS) is proposed which contains host- and network-based anomaly detection systems for the substations, and simultaneous anomaly detection for multiple substations. Potential scenarios of simultaneous intrusions into the substations have been simulated using a substation automation testbed. The host-based anomaly detection considers temporal anomalies in the substation facilities, e.g., user-interfaces, Intelligent Electronic Devices (IEDs) and circuit breakers. The malicious behaviors of substation automation based on multicast messages, e.g., Generic Object Oriented Substation Event (GOOSE) and Sampled Measured Value (SMV), are incorporated in the proposed network-based anomaly detection. The proposed simultaneous intrusion detection method is able to identify the same type of attacks at multiple substations and their locations. The result is a new integrated tool for detection and mitigation of cyber intrusions at a single substation or multiple substations of a power grid.**

***Index Terms*—Anomaly detection, cyber security of substations, GOOSE anomaly detection, SMV anomaly detection and intrusion detection.**

I. Introduction

A SMART GRID IS an enhanced power grid that generates, transmits, and uses electricity with the support of information and communications technology (ICT) for advanced remote control and automation [1], [2]. A smart grid has the potential to benefit power systems and customers, such as improved reliability, efficiency and reduced costs. For example, with advanced automation technology, a power grid can identify and isolate the faulted area(s) and restore unaffected areas by self-healing technologies [3]. Smart meters allow data acquisition from the customers to be conducted frequently and enable customer participation through various demand side response mechanisms [4]. Automation of the power grid includes substation and distribution automation. The subject of smart substations is a critical issue for the smart grid as it plays an important role in advanced monitoring and control of the power grids. The substation is installed with critical devices and communication networks such as IEDs, transformers, distribution feeders, circuit breakers, and communication systems. A smart substation enhances reliability and efficiency of operation, monitoring, control and protection [2].

Cyber security of substations has been recognized as a critical issue [5]. For example, well organized simultaneous cyber attacks to multiple substations can trigger a sequence of cascading events, leading to a system blackout [6], [7]. Therefore, an effective measure to address this issue is to prevent, detect, and mitigate malicious activities at the substations. Anomaly detection refers to the task of finding abnormal behaviors in data networks; it is a concept widely adopted in computer networks [8]. The term Intrusion Detection System (IDS) is also used for cyber security in a substation. The concept of the IDS was proposed by [9]. It monitors user access logs, file access logs, and system event logs to see if there is any anomaly in the host system. The work of [10] provides a model of an IDS that became a starting point of the recent IDSs. This model uses statistics for anomaly detection and an intrusion detection expert system (IDES). Typical approaches to intrusion detection are either network or host-based methods. A network-based IDS (NIDS) collects packets from a communication network and analyze network activities. References [11] and [12] propose network-based anomaly detection systems. A host-based IDS monitors a host system and generates alarms when anomalies and malicious activities are observed. The authors of [13] and [14] propose host-based anomaly detection. However, both network- and host-based intrusion detection methods have their own weaknesses. For example, host-based detection can fail to detect multiple hosts or applications. Network-based detection, on the other hand, can have a high rate of false alarms. In [15] and [16], the authors propose an integrated (or hybrid) anomaly detection system in order to compensate for the weakness of each system. The work of [17] proposes an intrusion detection system for IEC 61850 automated substations. A cyber-physical security vulnerability index has been proposed [18]. Temporal event construction based anomaly detection has been developed in the authors' previous work [19]. Reference [20] reports a framework for cyber-physical security. A system-level security design for power systems has been developed [21]. Cyber security technologies for anomaly detection at a substation are in an early stage of development. Technologies to detect anomalies for substation automation protocols are critically needed, such as GOOSE, SMV, and Manufacturing Message Specification (MMS).

Stuxnet, *Duqu* and *Flame* could be highly relevant cyber attacks (malwares) that are aimed at critical power infrastruc-

Manuscript received May 14, 2013; revised October 10, 2013; accepted December 03, 2013. Date of publication April 10, 2014; date of current version June 18, 2014. This research is sponsored by U.S. National Science Foundation, "Collaborative Research: Resiliency Against Coordinated Cyber Attacks on Power Grid." Grant 1202229. Paper no. TSG-00384-2013.

J. Hong is with the School of Electrical Engineering and Computer Science (EECS), Washington State University (WSU), Pullman, WA 99163 USA (e-mail: jhong@eecs.wsu.edu).

C.-C. Liu is with School of Electrical Engineering and Computer Science (EECS), Washington State University (WSU), Pullman, WA 99163 USA, and also with the School of Mechanical and Materials Engineering, University College Dublin, Ireland (e-mail: liu@eecs.wsu.edu).

M. Govindarasu is with Iowa State University, Ames, IA 50011 USA (e-mail: gmani@iastate.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2013.2294473

ture control systems [22]. Other cyber security concerns and potential threats to the power infrastructures have been reported by governments and other organizations, e.g., General Accounting Office (GAO), NIST, or Interagency Reports National Institute of Standards and Technology Interagency Report (NISTIR) and Department of Energy (DOE) [23]–[25]. In addition, substation automation standards existed before cyber security became a major concern for power grid. As a result, full security measures have not been incorporated in the open standards [26]. Multicast distribution techniques that are used for GOOSE and SMV enable an efficient communication mechanism; however, it also causes cyber security issues and vulnerabilities, e.g., open group membership and open access [27]. Due to the fast transmission time requirement (within 4 ms), most encryption techniques or other security measures that increase transmission delays may not be practical for GOOSE and SMV. Although the work of [26] proposes an authentication method through a digital signature, the performance test is yet to be performed. Current intrusion detection or anomaly detection methods do not normally support substation automation protocols, e.g., GOOSE and SMV; they are more focused on cyber attacks through Distributed Denial of Service attack (DDoS), and website and operating system (OS). Cyber intrusions related to GOOSE and SMV can cause serious damages. Intruder(s) can modify GOOSE control messages and trip circuit breakers in a substation. They can also send fabricated (and improper) protection coordination messages to other substations. A SMV message attack can generate fabricated analog values to a control center, leading to undesirable operations.

This paper is concerned with anomaly detection at a substation. An integrated method for host-based and network-based anomaly detection schemes is proposed. The host-based anomaly detection uses a systematic extraction technique for intrusion footprints that can be used to identify credible intrusion events within a substation, e.g., firewall, user-interface, IEDs, and circuit breakers. The network-based anomaly detection is focused on multicast messages in a substation network; it also detects, in a real-time environment, anomalies that demonstrate abnormal behaviors. The main contribution of this paper is a new method for 1) an integrated anomaly detection system for protection of IEC 61850 based substation automation system, e.g., IEDs, user-interface and firewall, and 2) a network-based anomaly detection algorithm that can be used to detect malicious activities of IEC 61850 based multicast protocols, e.g., GOOSE and SMV, across the substation network. Anomaly detection for multicast messages in a substation automation network is a new field of research for the power grids. In this research, a cyber security testbed has been developed and used to validate the proposed anomaly detection algorithms. Cyber intrusions are simulated using the testbed including protective IEDs. The test results demonstrate that proposed anomaly detection algorithms are effective for the detection of simulated attacks.

In the remaining of this paper, Section II describes cyber security vulnerabilities in a substation network. Section III includes algorithms for host- and network-based anomaly detection schemes. In Section IV, the network-based substation multicast messages are analyzed for anomaly detection. Section V provides the test results of the proposed anomaly detection system and the simultaneous intrusion detection at multiple substations. The conclusions and recommendations for future work are given in Section VI.

II. Cyber Security Vulnerability of a Substation

A power substation may consist of various types of equipment such as network devices, user-interface, server, global positioning system (GPS), firewall, IEDs, and remote access points. IEC 61850 based protocols are used by substation automation facilities, e.g., GOOSE, SMV, and MMS. GOOSE is used to send tripping signals from IEDs to circuit breakers. Sampled measured voltage and current values (SMV) are sent from a Merging Unit (MU) to an IED. Many devices are synchronized by GPS. MMS is used for monitoring, control and reporting between the user-interface and IEDs. Vulnerabilities of the substation network and mitigation of cyber attacks are critical subjects for anomaly detection. Remote access to a substation network from corporate offices or locations external to the substation is not uncommon for control and maintenance purposes. Dial-up, Virtual Private Network (VPN), and wireless are available mechanisms between remote access points and the substation Local Area Network (LAN) [28]. These access points are potential sources of cyber vulnerabilities. When remote access points have been compromised by an intruder, malicious attacks to operate circuit breakers and/or to access critical information, such as Substation Configuration Description (SCD), can be launched. IEDs may have a web server to allow remote configuration change and control. This paper assumes that the remote access point is the main intrusion point to the substations. An intruder may be able to access the substation network after the firewall is compromised. (S)he may capture, modify, and retransfer GOOSE packets and operate circuit breakers in a substation. The attacker may also send fabricated (and improper) GOOSE to other substations, causing unauthorized breaker operations. The consequence of a fabricated SV message attack can generate high current values to a control center and it may lead to an undesirable operation. After malicious activities or anomalies are detected in a substation network using the proposed integrated anomaly detection system, an intruder can be disconnected by collaboration between the IDS and firewall in the substation network. For a firewall, this can be achieved by dynamic rejection rules or disconnecting open ports. The proposed IADS uses anomaly and specification-based detection algorithms. Therefore, it is not able to detect unknown or intelligent attacks that are not defined in the algorithm. Periodic updates of the attack models will be important.

As illustrated in Fig. 1, possible intrusions to the substation communication network can originate from outside or inside a substation network, e.g.,

- From outside a substation network: Intrusions can originate from remote access points (A1) or a control center (A2) to the firewall and the substation communication network (A3). Once an intruder can access the substation communication network, (s)he can access other facilities in the substation.
- From inside a substation network: Intrusions can enter from the substation communication network (A3) or user-interface (A4) and then gain access to other facilities in the substation.

Here are examples on how an intrusion from inside and outside of a substation can be launched on a substation network:

Inside attack: if a USB is already infected by an attacker, it may be used to install malware on the substation user-interface. Then it may be used to open a predefined communication port or execute hacking tools.

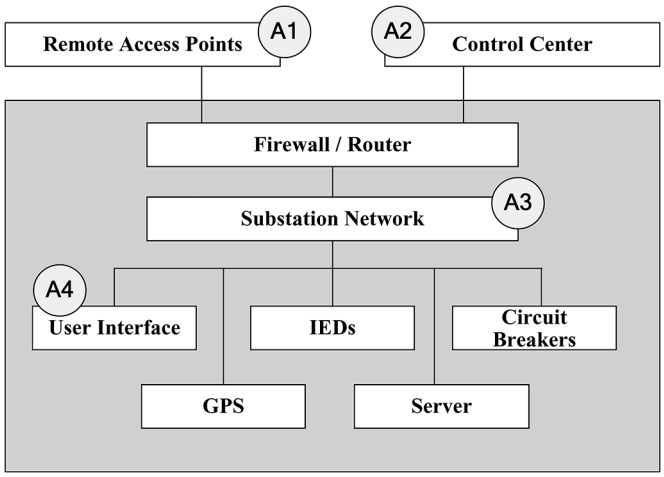

Fig. 1. Intrusion points in a substation automation system.

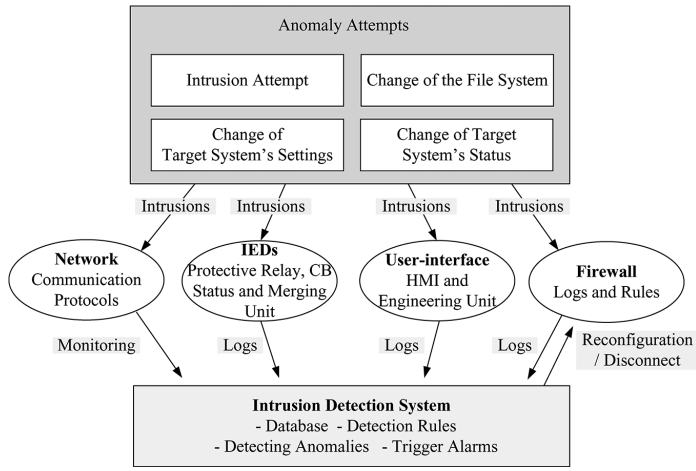

Fig. 2. Intrusion detection in a substation.

Outside attack: Remote access points may be used for maintenance, control or operation. Once an intruder compromises the access points, the attack may be able to pass the firewall and gain access to the substation ICT network.

Both inside and outside intrusions can be host-based or network-based attacks. A critical host-based attack is to compromise the user-interface machine. The user-interface system has the Human Machine Interface (HMI) and engineering tools that allow an operator or engineer to control, monitor or modify settings of the IEDs. A compromised user-interface can lead to undesirable operations of circuit breakers and settings for IEDs and transformer taps. Network-based intrusions can be conducted through packet monitoring, modification and replay attacks. Intruders can open circuit breakers by modifying GOOSE, SMV, and MMS messages in a substation network. Modification of Simple Network Time Protocol (SNTP) messages can disrupt time synchronization. Each of attacks may cause severe damages.

III. Anomaly Detection

Anomaly detection refers to finding patterns that indicate abnormal or unusual behaviors. It is a method for detection of cyber security intrusions [8] that requires data analysis and correlation of events.

As depicted in Fig. 2, intruders' behaviors generate logs across the substation-level networks, e.g., IEDs, firewall, user interface, and communication networks. For instance, the *Stuxnet* attack is based on: 1) intrusions, 2) changing the file system, 3) modifications of target system settings, and 4) altering the target system status [22]. If intruders try to compromise the substation targets, e.g., IEDs, networks, user interface and firewall, their behaviors will leave footprints in substation networks. Anomaly detection is performed based on logs of intruders' footprints.

A. Host-based Anomaly Detection

This section proposes a temporal anomaly detection method for host-based anomaly detection which is a generalization of the authors' previous work [19]. Generalizations from the authors' prior work are: 1) this paper proposes an integrated anomaly detection system, whereas [19] is concerned only with host-based anomaly detection in a substation, 2) this paper proposes a more efficient algorithm for attack similarity compared to the previous results, and 3) the generalized method incorporates a comprehensive set of substation logs and messages and extends the capability to scenarios involving multiple substations. The main assumption of the temporal anomaly detection for host-based anomaly detection is that the engineering software and hardware are able to generate system and security logs. For instance, if an intruder makes a wrong password attempt to IED or the user-interface, this action will generate a wrong password attempt flag. Similarly, if an intruder tries to copy or change a file in the user-interface, it will generate an unauthorized file change flag. The generalized method incorporates a comprehensive set of substation logs and messages and extends the capability to scenarios involving multiple substations. Temporal anomaly is used for host-based anomaly detection and can be determined from discrepancies between event logs from different time periods. As shown in Fig. 2, data logs at substation networks are used for the host-based anomaly detection algorithm.

The anomaly between two different time snapshots can be determined by a metric. The proposed technique is explained through an example. In Table I, the event log matrix $\mathbf{\Omega}$ with a dimension of 7 by 4, contains 7 rows of anomaly indicators at the same substation for 7 consecutive time instants. Each of the 4 columns represents a specific type of host-based anomaly indicator, i.e., ψ^a (intrusion attempt on user interface or IED), ψ^{cf} (change of the file system), ψ^{cs} (change of IED critical settings), and ψ^o (change of status of breakers/switches or transformer taps), respectively. If a specific type of anomaly is detected at time t, the value of the corresponding element in matrix $\mathbf{\Omega}$ will be changed from 0 (no anomaly) to 1 (anomaly). Detection of temporal anomalies is performed by comparing consecutive row vectors representing a sequence of time instants. The host-based ADS module imports the system and security logs from the user-interface, IEDs and firewalls at a predefined time. In this paper, the predefined polling time of system and security logs data is 10 seconds. An example of $\mathbf{\Omega}$ matrix describes a temporal anomaly detection for a sequence of 7 time instances of a substation A and B, and the time difference from t_1 to t_2 in $\mathbf{\Omega}$ matrix is 10 seconds. After subscribing to the logs, a data convertor module will change all temporal logs to binary values as shown in Table I (substation A) and Table II. For example, Table I (substation A) has the converted binary values from Table II. A detailed explanation is given in the following:

TABLE I
AN EXAMPLE OF TEMPORAL ANOMALY DETECTION IN SUBSTATIONS

Substation A					Substation B				
t_1	0	0	0	0	t_1	0	0	0	0
t_2	1	0	0	0	t_2	1	0	0	0
t_3	1	1	0	0	t_3	1	1	0	0
$\mathbf{\Omega} = t_4$	1	1	0	0	$\mathbf{\Omega}' = t_4$	1	1	0	0
t_5	1	1	0	0	t_5	1	1	0	1
t_6	1	1	1	1	t_6	1	1	1	1
t_7	1	1	1	1	t_7	1	1	1	1

TABLE II
SYSTEM LOGS OF A SUBSTATION A

Substation A				
o.	**Date**	**Time**	**Contents**	**Issue**
45	15.09.2013	10:28:33,560	IED 1	Wrong password attempt
46	15.09.2013	10:35:43,159	User-interface	Unauthorized file change
47	15.09.2013	11:02:04,368	IED 2	Unauthorized setting change
48	15.09.2013	11:03:14,270	Transformer 1	Unauthorized tap change

— At 10:20:000, there is no anomaly so t_1 is [0 0 0 0].
— At 10:30:000, ADS detects a wrong password attempt to IED 1 so t_2 is [1 0 0 0].
— At 10:40:000, ADS detects an unauthorized file change to the user-interface so t_3 is [1 1 0 0].
— At 10:50:000, there is no change so t_4 is [1 1 0 0].
— At 11:00:000, there is no change so t_5 is [1 1 0 0].
— At 11:10:000, ADS detects two anomalies, unauthorized setting change to IED 2 and unauthorized tap change to transformer 1 so t_6 is [1 1 1 1].
— At 11:20:000, there is no change so t_7 is [1 1 1 1].

An example of $\mathbf{\Omega}$ matrix describes a temporal anomaly detection for a sequence of 7 time instances of a substation A as shown in Table I.

An assumption of temporal anomaly detection for host-based anomaly detection is that the engineering software and hardware are able to generate system and security logs. For instance, if an intruder makes a wrong password attempt to IED or the user-interface, this action will generate a wrong password attempt flag. In the same manner, if an intruder tried to copy or change a file in the user-interface, it will generate an unauthorized file change flag. Some products have this log generating function but not all. If a specific type of anomaly is detected at time t, the value of the corresponding element in matrix $\mathbf{\Omega}$ will be changed from 0 (no anomaly) to 1 (anomaly). The binary number 1 (anomaly) will be kept until the operator resolves the issue and resets the integrated anomaly detection system. After resetting, all elements in matrix $\mathbf{\Omega}$ will be set to zero (no anomaly). The main reason to use binary values for temporal anomaly detection is to minimize the calculation time for simultaneous anomaly detection at multiple substations.

If a discrepancy exists between two different periods (rows), the vulnerability index $V_h^{\mathbf{\Omega}}$ is a number between 0 and 1. A value of 0 implies no discrepancy whereas 1 indicates the maximal discrepancy. A scalar index for temporal anomaly at time $t = t_i$ is defined as

$$V_{h(i)}^{\mathbf{\Omega}} = \frac{\sum_{j=1}^{n} \left| \mathbf{\Omega}_{(i,j)} - \mathbf{\Omega}_{(i+1,j)} \right|}{n}, \quad i = 1, \ldots, 6, \quad (1)$$

where n is the total number of anomaly indicators ($n = 4$ for this example). Based on (1) one can obtain a vector for temporal anomaly that provides irregularities of events during the selected time period, $t = t_1, \ldots, t_7$, from $\mathbf{\Omega}$ matrix, i.e.,

$$V_h^{\mathbf{\Omega}} = (0.25, 0.25, 0, 0, 0.5, 0). \quad (2)$$

The first element of (2) is the value from the calculation based on first and second row of $\mathbf{\Omega}$ in Table I, similarly for other elements. The anomaly of this substation is determined by the vector $V_h^{\mathbf{\Omega}}$. If $V_h^{\mathbf{\Omega}}$ is a zero vector, then there is no anomaly event on this substation. Otherwise, the substation will be included in the credible list to be evaluated further.

The proposed temporal anomaly detection is extended to detect simultaneous anomaly detection among multiple substations. The simultaneous anomaly detection is achieved in 3 steps, i.e., 1) Find the total number of types of attacks, 2) Find the same attack groups, and 3) Calculate the similarity between attacks in the same group. The total number of types of attack can be calculated by

$$\text{Total number of types of attack} = \sum_{k=1}^{n} \frac{n!}{k!(n-k)!} + 1, \quad (3)$$

where n is the total number of anomaly indicators. Eq. (3) is based on binomial coefficients. The total number of types of attacks for the specific example above is 15 since it has 4 anomaly indicators (number of columns). Let the event log matrix $\mathbf{\Omega}'$ be an indicator for a different substation, as shown in Table I. In comparison with $\mathbf{\Omega}$, it is assumed that the $\mathbf{\Omega}'$ matrix has identical values except for the 5th row which is [1, 1, 0, 1]. Then the attack patterns of $\mathbf{\Omega}$ and $\mathbf{\Omega}'$ are considered to be the same since they eventually have the same values in the last row, i.e., [1, 1, 1, 1]. It indicates that substations $\mathbf{\Omega}$ and $\mathbf{\Omega}'$ are under a simultaneous attack but the attack sequences are different. Once the same type of attack groups is found as described above, the similarity between attacks can be calculated by

$$Attack\ Similarity = 1 - \frac{\sum_{i=1}^{x} \sum_{j=1}^{y} \left| \mathbf{\Omega}_{(i,j)} - \mathbf{\Omega}'_{(i,j)} \right|}{x \cdot y}, \quad (4)$$

where x and y are total number of rows and columns of matrix, respectively ($x = 7$ and $y = 4$ for this example). Attack similarity value of 0 indicates no overlap and a value 1 indicates a complete overlap. Therefore, by (4), the similarity index between substation $\mathbf{\Omega}$ and $\mathbf{\Omega}'$ is 0.9643.

B. Network-Based Anomaly Detection

The proposed method also provides a network-based anomaly detection algorithm for multicast messages in the substation automation network. The multicast messages are based on IEC 61850 standard, e.g., GOOSE and SMV. The proposed Substation Multicast Message Anomaly Detection (SMMAD) model in Fig. 3 is divided into 3 process modules, i.e., packet filtering, anomaly detection, and evaluation. The packet filtering module consists of functions to identify GOOSE and SMV messages. The filter will only allow passing for GOOSE and SMV messages so the burden of processing can be reduced and the system performance will increase. The anomaly detection module is used to find violations based on predefined rules. The evaluation module will decide if the detected anomaly status is "abnormal" or "attack." Details will be explained in the next section.

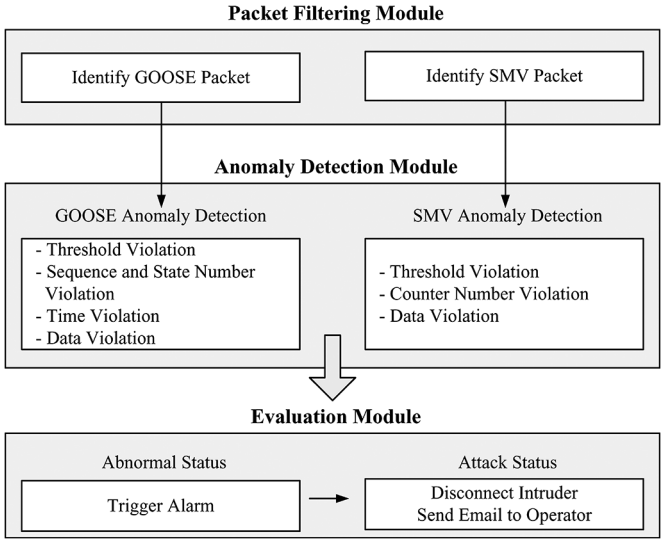

Fig. 3. SMMAD modeling for ADS.

IV. Substation Multicast Message Anomaly Detection

A. Multicast Messages in IEC 61850

Multicast messages in IEC 61850, e.g., GOOSE and SMV, are different from other protocols used in substation automation because they use three layers in Open Systems Interconnection (OSI) model stack, i.e., physical, data link, and application layer, in a real-time requirement. The multicast scheme uses the Media Access Control (MAC) address [29]. The GOOSE service uses a re-transmission scheme to enhance the communication speed and reliability, i.e., the same GOOSE message is re-transmitted at different time intervals but no response is sent from the receiver. The sequence number of GOOSE messages will be increased for each transmission and the state number will increase when the data status is changed. The sequence number will be set to 0 when the state number is changed. However, the specific time of re-transmission (interval) is not defined in the IEC 61850 standard so different vendors' GOOSE re-transmission times may vary [30].

SMV of voltage and current messages are published from the Merging Unit and subscribed by IEDs. The resolution amplitude of the Merging Unit in this project is 16 bits so it will send 960 SMV voltages and currents to IEDs in a second [31]. The message counter is incremented each time when a new sampled packet is published.

B. Detection Method

Unwanted multicast message packets can be identified by rules that match known signatures. Therefore, anomalies which match the predefined rules can be detected by the ADS. Each rule has been defined based on the IEC 61850 standard. First, an "anomaly" state has been identified as a result of violation of predefined rules. Second, an "attack" state is identified if the detected anomaly will adversely affect proper functioning of the substation control and measurement, e.g., open circuit breaker and change of voltage and current values. A binary status is used as indicators of the status, i.e., "0" means no anomaly and "1" indicates that an anomaly is detected.

Port mirroring is a function to copy all packets from port(s) to the specific port in order to monitor and analyze packets. General network-based ADS will need port mirroring to capture all communication packets in the network [17]. Note that the proposed ADS is able to capture the GOOSE and SMV without the port mirroring function as it is focused on multicast messages and not other packets.

C. Main Framework

After calculating the violation detection indicators in GOOSE and SMV anomaly detection modules, the anomaly detection module will determine if there is an anomaly using the rules in Appendix I. As shown in Appendix I, Line 7 is used for GOOSE anomaly detection, i.e., any detected anomaly in threshold violation α_{Th}^{G}, sequence and state number violation β_{S}^{G}, GOOSE time violation γ_{Ti}^{G}, and GOOSE data violation δ_{d}^{G} will change GOOSE network-based anomaly indicator ψ^{G} from $false$ to $true$. On the other hand, Line 12 is developed for SMV anomaly detection, i.e., any detected anomaly in SMV threshold violation ε_{Th}^{SV}, counter number violation θ_{cn}^{SV}, and SMV data violation μ_{d}^{SV} will set the SMV network-based anomaly indicator ψ^{SV} from $false$ to $true$.

After the anomaly detection task is completed, a network-based substation vulnerability index V_{n}^{GS} is defined as follows:

$$V_{n}^{GS} = \begin{cases} 1, & \text{If } \psi^{G} = true \\ 1, & \text{If } \psi^{SV} = true \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where ψ^{G} is the GOOSE network-based anomaly indicator and ψ^{SV} is the SMV network-based anomaly indicator. A result of $V_{n}^{GS} = 1$ indicates the existence of an intrusion based on GOOSE and SMV messages whereas $V_{n}^{GS} = 0$ indicates that there is no evidence of a multicast message based cyber intrusion.

The proposed SMMAD examines all GOOSE and SMV packets in the substation network, and then checks if there is a security violation, as shown in Appendix I. SMMAD has two phases: initialization and detection. Line 1 represents the initialization of the examination process. Line 2 captures all packets in a substation network. Lines 3 and 10 are to check whether this is a GOOSE or SMV message. Line 4 and 11 are used to analyze the captured packets. Lines 5 and 6 create anomaly detection threads if there is more than one type of GOOSE messages. Lines 7 and 12 are used to check if there is a security violation. Finally, Lines 8, 9 and 13, 14 show whether there is an intrusion.

D. GOOSE Anomaly Detection

The threshold of GOOSE packets G_{th} can be calculated by the pre-defined re-transmission rule. The proposed ADS can filter the GOOSE packets by checking recommended MAC address from 01-0C-CD-01-00-00. Then the count of GOOSE packet is maintained, and details of this packet are saved. When the captured number of GOOSE packets G_{cnp} within predefined time G_{th}^{T} is greater than the predefined threshold for GOOSE packets G_{th} within G_{th}^{T} or there is no captured GOOSE packet within G_{th}^{T}, an anomaly is deemed to be occurring and details are written to the log file. This process can also detect a GOOSE based denial-of-service (DoS) attack. Hence, the GOOSE violation indicator (GVI) α_{Th}^{G} is changed from 0 to 1. Line 1 in Appendix II is used for the detection of threshold violation α_{Th}^{G}.

The state number of GOOSE messages G_{st} will change and the sequence number of GOOSE G_{sq} will be set to 0 when the GOOSE state is changed. The sequence number of GOOSE will increase when GOOSE is published. Hence, if a captured GOOSE message's sequence number is not set to zero after the state is changed or sequence number is not matched as a sequence, it will detect the anomalies that are suspicious as attacker(s)'s

packet modification or injection to the substation network. The GVI β_S^G will be changed from 0 to 1. Line 2 in Appendix II is for the GOOSE sequence and state number violation β_S^G detection.

In general, the GOOSE clients and servers are synchronized within a few microseconds for the critical protection and control functions. The time stamp will be implemented in the GOOSE packet by the sender. So anomaly will be detected when the generated time stamp G_{ge}^T is greater than the receiver's time G_{re}^T. The recommended GOOSE transfer time G_{tr}^T is defined in the IEC 62351-1 standard, which is 4 ms. If the difference between the generated time and received time is greater than the transfer time, it will be considered an anomaly. The GVI γ_{Ti}^G will be changed from 0 to 1. Line 3 in Appendix II is for the GOOSE time violation γ_{Ti}^G detection.

When the GOOSE indicator that contains the binary control value is changed from $false$ to $true$ or vice versa, the state number of GOOSE will be changed to the next number and the sequence number will be set to 0. Therefore, if there is any violation of this rule, the GVI δ_d^G will be changed from 0 to 1. Line 4 in Appendix II is to perform the detection of GOOSE data violation δ_d^G.

E. Sampled Measured Values Message Anomaly Detection

The threshold for SMV packets S_{th} depends on the sampling rate. The proposed ADS will capture the SMV message by checking MAC address which starts from 01-0C-CD-04-00-00. Then it will count the number of SMV every second, and save the detailed information. When the captured number of SMV packets S_{cnp} within predefined time S_{th}^T is greater than the predefined threshold for SMV packets S_{th} within S_{th}^T or there is no captured SMV packet within S_{th}^T, an anomaly is deemed to be occurring and details are written to the log file. This can also detect a SMV based denial-of-service attack. A SMV violation indicator (SVI) ε_{Th}^{SV} will be changed from 0 to 1. Line 5 in Appendix II is used to perform the detection of SMV Threshold violation ε_{Th}^{SV}.

For the counter number violation detection, "SmpCnt" is a SMV protocol attribute and its attribute type is INT16U. This value will be incremented each time SMV is published. The count will be set to zero when sampling is synchronized by a clock signal [29]. The SMV message counter S_{mc} corresponds to SmpCnt so it will also increase after each transmission. If the SMV message counter is not increased or equal to the previous count when sampling is not synchronized, the SVI θ_{cn}^{SV} will be changed from 0 to 1. Line 6 in Appendix II is used to carry out the counter number violation θ_{cn}^{SV} detection.

Each group of SMV message has its own identification S_{id} and name of dataset S_{ds} [29]. They will not change unless the configuration of the Merging Unit is changed. Therefore the proposed algorithm will detect the anomalies when there is a modification of the name of identification and dataset, and they still contain the same source and destination MAC address. Then, the SVI μ_d^{SV} will be changed from 0 to 1. Line 7 in Appendix II is to detect the SMV data violation μ_d^{SV}.

V. Simulation Results

A testbed is developed at WSU to perform different types of cyber intrusions and analyze the effectiveness of the proposed detection and mitigation techniques in a realistic substation environment. Government agencies and other organizations have been using various testbeds for cyber security testing [32]–[34]. In this paper, several types of cyber attacks have been generated for validation of the proposed anomaly detection algorithms, e.g., replay, packet modification, injection, generation and DoS using the testbed. The results of the simulated attacks are shown in Tables III and IV and Fig. 5.

TABLE III
Consequence of Goose Based Malicious Behaviours Without Anomaly Detection System

Action	Result
Disconnect Ethernet cable from IED	Lost availability of IED
Send normal control	Open CB
Replay attack	Open CB
Modify sequence & state number	Warning occurred at CB
Modify transferred time	Warning occurred at CB
Modify GOOSE control data	Open CB
Denial of Service attack	Lost availability of CB
Generate GOOSE control data	Open CB

TABLE IV
Consequence of SMV Based Malicious Behaviours Without Anomaly Detection System

Action	Result
Disconnect Ethernet cable from MU	Lost availability of MU
Increase measured values	Open CB
Replay attack	Open CB
Modify counter number	Warning occurred at IED
Modify SMV dataset	Warning occurred at IED
Denial of Service attack	Lost availability of IED
Generate SMV data	Open CB

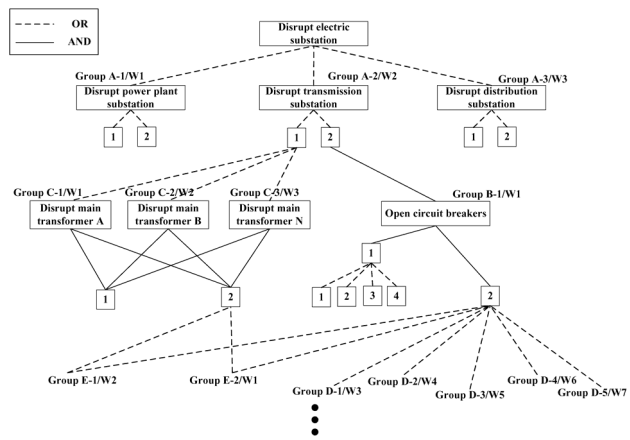

Fig. 4. Attack tree for the substations.

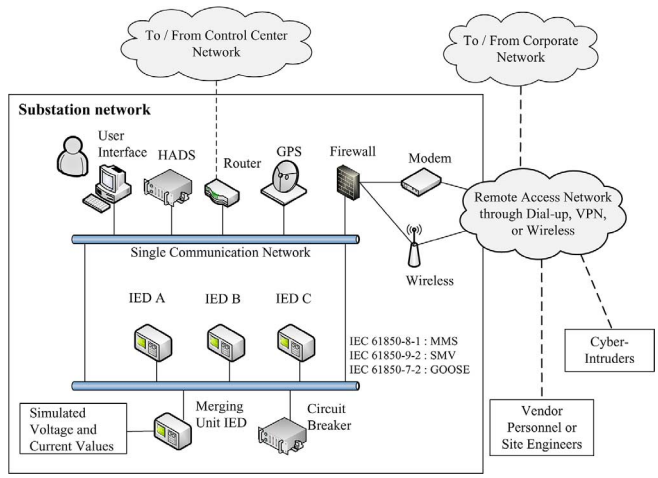

Fig. 5. WSU cyber security testbed for the substation.

Fig. 4 shows a portion of the attack tree for the substations that have been used in Case study I, II, and III. For instance, the

goal of group B-1/W1 is to open a circuit breaker. The preconditions of this attack are: an attacker can open a circuit breaker via IED, the control center user interface, and substation user interface. The goal of this attack is achieved with two AND conditions, i.e., 1) find target CB, and 2) open target CB, as shown in Fig. 4. Find target has four OR conditions: 1) send a GOOSE message to CB using IED, 2) use control center user interface, 3) use substation user interface, and 4) modify the protection setting (to low value) of IED. The post condition of this attack is that the attacker will open target CB. It is shown that some intrusions are able to execute a switching action on the circuit breaker. The C language based source code library was used for the proposed integrated ADS. The proposed anomaly detection algorithms are implemented in the C language. C++ has been used for ADS HMI in order to test the real-time anomaly detection and alarms to the substation operator. The circuit breaker is designed to subscribe GOOSE messages generated from the IEDs. IED A is designed to subscribe to SMV messages that are from the Merging Unit. Free available software tools are used for all intrusion processes, e.g., Wireshark, Colasoft Packet Builder, Nmap, etc.

The simulation results include 3 Study Cases. Case I shows the GOOSE cyber intrusions and detection on the substation communication network. Both single and simultaneous attacks are considered. The results demonstrate that the proposed method detects all intrusions and triggers the appropriate alarms. Case II is a simulation of SMV intrusions and detection. SMV packets are captured and retransferred to the substation network after they are falsified to include high current and voltage values. The results showed that the proposed anomaly detection can detect simulated intrusions and then trigger the alarms. Case III is concerned with simultaneous anomaly detection at multiple substations. Cyber intrusions are generated by attacker(s) and detected by the ADS up to 2000 substations. The results show that proposed algorithm is faster than others.

A. Case Study I: GOOSE Anomaly Detection

As shown in Table V, the threshold of GOOSE messages G_{th} has been set to 12, including a margin of error of 20%, since the peak number of GOOSE messages when normal control was issued is 10. The GOOSE anomaly detection results are given in Table V.

TABLE V
GOOSE ANOMALY DETECTION TEST RESULTS

Test case	Set packet threshold (per 1 sec)	Normal control issued	Disconnect Ethernet cable from IED	Detected anomalies	Alert issued
T1	12	No	No	-	No
T2	12	Yes	No	-	No
T3	12	No	No	, ,	Yes
T4	12	No	No		Yes
T5	12	No	No		Yes
T6	12	No	No		Yes
T7	12	No	No	, ,	Yes
T8	12	No	No		Yes
T9	12	No	Yes		Yes
T10	12	No	No	, , ,	Yes
- The peak number of normal GOOSE message when control was issued: 10 (per second) - Number of normal GOOSE message: 1 (per second)					

— T1, normal status: There was no alarm under a normal operating condition.
— T2, normal control issued: There was no alarm when normal control was issued to IED.
— T3, replay attack (20 packets/s): The normal control GOOSE packet was captured from T2 and retransferred to the substation network by the attacker without any modification.
— T4, sequence and state number modification attack (5 packets/s): Change sequence and state number of GOOSE packets and then transfer to substation network by the attacker.
— T5, transferred time modification attack (5 packets/s): Change time stamp of GOOSE packets and then transfer to substation network by the attacker.
— T6, GOOSE control data modification attack (5 packets/s): Change control data of GOOSE packets and then transfer to the substation network by the attacker.
— T7, Denial of Service attack (2000 packets/s): Execute GOOSE based DoS attack by the attacker.
— T8, generating GOOSE control data attack (5 packets/s): Generate GOOSE control messages and publish to the substation network by the attacker.
— T9, disconnect Ethernet cable: Disconnect Ethernet cable from IED by the attacker so there was no GOOSE message in the substation network.
— T10, simultaneous attack: Change sequence, state number, time stamp, and control data of GOOSE packets and then transfer to substation network by the attacker.

B. Case Study II: SMV Anomaly Detection

Table VI shows that the threshold of SMV messages S_{th} has been set to 1178, including a margin of error of 20%, since the peak number of SMV messages when normal control was issued is 982.

Once ADS detects an anomaly in a substation network, it will trigger an alarm and send a message to operators. Also ADS will send a disconnect control command to the firewall and block the intruder's connection as a mitigation action.

C. Case Study III: Multiple Substations

An anomaly detection system is intended to find malicious behaviors quickly so that system operators can disconnect the intruder(s) from the network and take other mitigation actions. If there are simultaneous intrusions from multiple attackers, however, it is difficult to mitigate the situation since different types of intrusions will require corresponding countermeasures. The ability to find the same type of attacks and their locations will reduce the mitigation time and effort. The total number of types of attack is 57 since the proposed ADS has 6 anomaly indicators [4 of host-based anomaly indicators from Section III-A and 2 of network-based anomaly indicators from (5)] as shown in Tables VII and VIII. Tables VII and VIII also report sample ADS logs of substations 1 and 2, respectively, where 0 indicates no anomaly and 1 indicates a detected anomaly. Table VII includes logs indicating an intrusion into substation 1, leading to a change of settings and GOOSE attack. This attack is shown to start from intrusion attempts ψ^a at t_2. Then logs indicate an unauthorized change of settings ψ^{cs} for a protective device at t_3. This type of attacks may happen when attackers know the password for the IED configuration tool. The intruder also attempts the GOOSE based attack at t_4. Table VIII provides logs from the ADS in substation 2. It shows the same attack as the one at substation 1 since the attack pattern of substation 1 is [1, 0, 1, 0, 1, 0] and substation 2 also has [1, 0, 1, 0, 1, 0] at t_7 but the attack time is different. Therefore, by (4), the attack similarity index between substations 1 and 2 is 0.9048.

TABLE VI
SMV ANOMALY DETECTION TEST RESULTS

Test case	Set packet threshold S_{th} (per 1 sec)	Disconnect Ethernet cable from MU	Detected anomalies	Alert issued
T11	1178	No	-	No
T12	1178	No	ε_{Th}^{SV}, θ_{cn}^{SV}	Yes
T13	1178	No	θ_{cn}^{SV}	Yes
T14	1178	No	μ_d^{SV}	Yes
T15	1178	No	ε_{Th}^{SV}, θ_{cn}^{SV}	Yes
T16	1178	No	ε_{Th}^{SV}, θ_{cn}^{SV}	Yes
T17	1178	Yes	ε_{Th}^{SV}	Yes
T18	1178	No	ε_{Th}^{SV}, θ_{cn}^{SV}, μ_d^{SV}	Yes
- The peak number of SMV message: 982 (per second)				

— T11, normal status: There was no alarm under a normal operating condition.
— T12, replay attack (200 packets/s): The normal SMV packet was captured and retransferred to the substation network without modification by the attacker.
— T13, counter number modification attack (20 packets/s): Change the counter number of SMV packets and then transfer to substation network by the attacker.
— T14, SMV dataset modification attack (20 packets/s): Change the dataset of SMV packets and then transfer to the substation network by the attacker.
— T15, Denial of Service attack (2000 packets/s): Execute SMV based DoS attack by the attacker.
— T16, generating SMV data attack (100 packets/s): Generate SMV messages that contain high current and voltage values, and publish to the substation network by the attacker.
— T17, disconnect Ethernet cable: Disconnect Ethernet cable from MU by the attacker so there was no SMV message in the substation network.
— T18, simultaneous attack: Change the counter number and dataset of SMV packets and then transfer to the substation network by the attacker.

TABLE VII
DETECTED ANOMALY LOG SUBSTATION 1

Time	Host-based				Network-based	
	ψ^a	ψ^{fs}	ψ^{cs}	ψ^o	ψ^G	ψ^{SV}
t_1	0	0	0	0	0	0
t_2	1	0	0	0	0	0
t_3	1	0	1	0	0	0
t_4	1	0	1	0	1	0
t_5	1	0	1	0	1	0
t_6	1	0	1	0	1	0
t_7	1	0	1	0	1	0

TABLE VIII
DETECTED ANOMALY LOG SUBSTATION 2

Time	Host-based				Network-based	
	ψ^a	ψ^{fs}	ψ^{cs}	ψ^o	ψ^G	ψ^{SV}
t_1	0	0	0	0	0	0
t_2	0	0	0	0	0	0
t_3	0	0	0	0	0	0
t_4	1	0	0	0	1	0
t_5	1	0	1	0	1	0
t_6	1	0	1	0	1	0
t_7	1	0	1	0	1	0

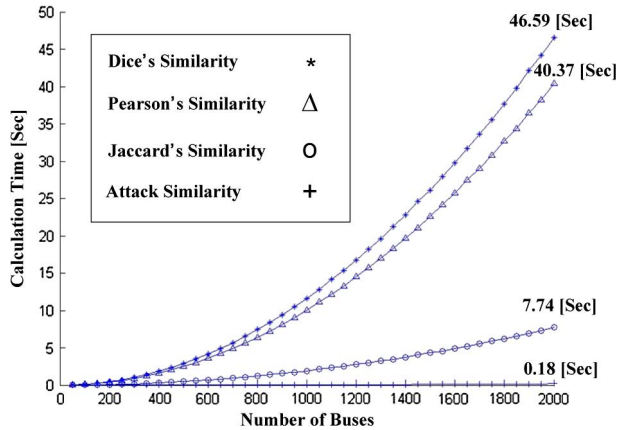

Fig. 6. Comparison of similarity coefficient algorithms.

The simulation steps are explained as follows. First, different types of attacks are randomly generated from multiple attackers. Second, all anomalies are captured and detected by the proposed ADS, and then the ADS generates logs at each substation. Third, simultaneous intrusion detection has been performed using generated logs. The proposed methodology for simultaneous anomaly detection at multiple substations is validated using the simulated data shown in Fig. 6. The proposed simultaneous anomaly detection method is able to identify the same type of attacks and its similarity within 0.18 seconds among 2000 substations. It also shows that the computational performance of the proposed host-based anomaly detection algorithm is faster than the previous algorithm developed by the authors that uses Pearson's Similarity and the other similarity coefficient algorithms [35].

D. ADS Evaluation

The false positive ratio (FPR) is defined as the number of misclassified normal packets divided by the total number of normal packets. The false negative ratio (FNR) is defined as the number of misclassified abnormal packets divided by the total number of abnormal packets. The FPR and FNR of the proposed host-based anomaly detection system depend on the accuracy of the event log matrix $\boldsymbol{\Omega}$ generated from the substation logs. They are 0.00013 and 0.0002, respectively. FPR and FNR of the proposed network-based anomaly detection system depend on the number of packets per second. This is due to the fact that ADS may lose packets when the number of packets exceeds 2000 per second. FPR and FNR are 0.00013 and 0.00016 for the case of 2100 packets, respectively. In order to compare the performance of the proposed network-based anomaly detection, a rule-based detection system using Tshark is used [36]. TShark is a network protocol analyzer. It allows users to capture packet data from a live network, or read packets from a previously saved capture file, either printing in a decoded form to the standard output or writing the packets to a file [37]. The resulting FPR and FNR of the rule-based detection system is 0.00142 and 0.0019, respectively. Therefore the proposed network-based anomaly detection shows a higher performance.

VI. CONCLUSION

This paper provides an integrated anomaly detection system which contains host- and network-based anomaly detection for a

single substation, and simultaneous anomaly detection for multiple substations. The host-based ADS uses logs that are extracted from malicious footprints of intrusion-based steps across substation facilities. The network-based ADS can detect malicious behaviors that are related to multicast messages in the substation network. The proposed simultaneous intrusion detection method is able to find the same type of attacks on multiple substations and their locations. The methods have been validated by testing with realistic intrusion scenarios using the testbed, e.g., replay, modification, man-in-the-middle, generation, and DoS. In order to enhance the detection rate, substation systems need to generate more system and security logs since the proposed host-based anomaly detection depends on the generated logs. The network-based anomaly detection algorithm should be updated periodically since it is not able to detect unknown attacks that are not defined in the algorithm. In future work, it will be useful to include other substation automation communication protocols, e.g., MMS, SNTP, DNP, Modbus, and IEC 60870-5 based anomalies.

APPENDIX I

SMMAD Algorithm

1. $\psi^G, \psi^{SV}, V_n^{GS} = 0$; // Initialize
2. capture C_{pkt}; // Capture all packets in the substation network
3. if (C_{pkt} is IEC GOOSE);
4. $G_{cp} = [G_{st}, G_{sq}, G_{ge}^T, G_{re}^T]$; // Parse packet
5. if ($G_{cp}[G_{sm}, G_{dm}] \neq G_{cp-1}[G_{sm}, G_{dm}]$); // Find different GOOSE
6. make $G_{at\ new}$; // Create new anomaly detection thread
7. $\alpha_{Th}^G \vee \beta_S^G \vee \gamma_{Ti}^G \vee \delta_d^G \rightarrow \psi^G$; // Calculate GOOSE intrusion
8. if ($\psi^G = true$), set $V_n^{GS} = 1$; // Detect GOOSE intrusion
9. else set $V_n^{GS} = 0$; // No intrusion
10. elseif (C_{pkt} is IEC SMV);
11. $S_{cp} = [S_{mc}, S_{ds}, S_{id}, S_{sm}, S_{dm}]$; // Parse packet
12. $\varepsilon_{Th}^{SV} \vee \theta_{cn}^{SV} \vee \mu_d^{SV} \rightarrow \psi^{SV}$ // Calculate SMV intrusion
13. if ($\psi^{SV} = true$), set $V_n^{GS} = 1$; // Detect SMV intrusion
14. else set $V_n^{GS} = 0$; // No intrusion
15. return V_n^{GS};

The GOOSE and SMV messages have its own recommended MAC address as defined in IEC 61850-8-1 standard. The first three octets are assigned by IEEE with 01-0C-CD. Then fourth octet shall be 01 for GOOSE and 04 for multicast sampled values. The last two octets shall be used as individual addresses assigned by the range defined in Table IX.

Therefore, the proposed ADS filters the GOOSE and SMV packets by checking the recommended MAC addresses, 01-0C-CD-01-00-00 and 01-0C-CD-04-00-00, respectively.

TABLE IX
RECOMMENDED ADDRESS RANGE ASSIGNMENTS

Service	Starting address (hexadecimal)	Ending address (hexadecimal)
GOOSE	01-0C-CD-01-00-00	01-0C-CD-01-01-FF
SMV	01-0C-CD-04-00-00	01-0C-CD-04-01-FF

The ADS can create anomaly detection threads if there is more than one type of GOOSE messages by checking the MAC address. For instance, the first GOOSE MAC address is 01-0C-CD-01-00-01 and, if there is another GOOSE packet that has a MAC address 01-0C-CD-01-00-02, ADS will create a new anomaly detection thread. The proposed ADS can handle up to two different types of GOOSE messages. If a captured packet is a GOOSE, ADS will analyze the captured packets. Then ADS detects malicious activities and abnormal behaviors that match predefined security rules described in Section IV-D. Finally ADS shows to the operator whether there is a GOOSE related intrusion or an anomaly. The creation of the new detecting thread is not applicable for SMV detection at this moment since ADS cannot handle too much data. The resolution (bits) amplitude of SV for protection and control is defined in IEC 61850-5, e.g., 8 bits (P1 class), 16 bits (P2 class), and 32 bits (P3 class). For example, SMV used in this research publishes approximately 960 packets in a second (using 16 bits). In the same manner, if a captured packet is a SMV, ADS will analyze the captured packets. Then ADS will detect malicious activities and abnormal behaviors that match predefined security rules described in Section IV-E. Finally, ADS provides an indication to the operator whether there is a SMV related intrusion or an anomaly.

APPENDIX II

GOOSE and SMV Violation Indicators

1. α_{Th}^G: [(G_{max} within $G_{Th}^T > G_{cnp}$ within G_{th}^T) $\vee$ ($G_{cnp} = 0$ within G_{th}^T)].
2. β_S^G : [$(G_{sq} \geq G_{sq-1}) \wedge (G_{st} \leq G_{st-1})$]$\vee$[$(G_{sq} \leq G_{sq-1}) \wedge (G_{st} \geq G_{st-1})$].
3. γ_{Ti}^G : $(G_{ge}^T \geq G_{re}^T) \vee [(G_{re}^T - G_{ge}^T) > G_{tr}^T]$.
4. δ_d^G : $(G_{cp} \neq G_{cp-1}) \wedge [(G_{sq} \leq G_{sq-1}) \wedge (G_{st} \leq G_{st-1})]$.
5. ε_{Th}^{SV}: (S_{max} within $S_{th}^T > S_{cnp}$ within S_{th}^T) $\vee$ ($S_{cnp} = 0$ within S_{th}^T).
6. θ_{cn}^{SV}: $(S_{mc} \leq S_{mc-1})$ when $(S_{si} = S_{si-1} = False)$.
7. μ_d^{SV} : [$(S_{sm} = S_{sm-1}) \vee (S_{dm} = S_{dm-1})$]$\wedge$[$(S_{ds} \neq S_{ds-1}) \vee (S_{id} \neq S_{id-1})$].

Examples are provided on how the proposed host- and network-based anomaly detection system can find the GOOSE and SMV related anomalies and intrusions.

Example I: An intruder gains access to the substation network via VPN. (S)he scans all IP address and opens ports using a scanning tool. After the information of protection IED is found, (s)he captures GOOSE packets of the target IED. Then the intruder modifies control data of GOOSE messages and retransfers to the substation network. Now ADS will detect the modified GOOSE

TABLE X
AN EXAMPLE OF NORMAL GOOSE OPERATION AND ANOMALY IN A SUBSTATION

	Normal operation			Anomaly		
Time	State number	Sequence number	Data	State number	Sequence number	Data
1	3	145	False	3	145	False
2	3	146	False	3	146	False
3	4	0	True	3	146	True
4	4	1	True	3	146	True
5	4	2	True	3	146	True

message since the intruder fails to synchronize the sequence number, state number, and time stamp of GOOSE.

Example II: An intruder gains access to the substation network via a dial-up connection. (S)he has a communication topology diagram and information. Intruder checks whether MU is live. After the information of the merging unit is found, (s)he captures SMV packets of the target merging unit. Then the intruder modifies the measured current values of the SMV message and retransfers to the substation network. Now ADS will detect the modified SMV messages since the counter number of injected SMV messages is not synchronized with the original SMV messages.

Example III: The left column of Table X shows a normal operation whereas the right column shows a GOOSE modification attack. When there is an open circuit breaker control event between time 2 and time 3, the state number is changed from 3 to 4 and the sequence number is set to 0. Then the sequence number is increased from 0 to 1, 1 to 2, etc. However, if an intruder captures, modifies data and retransfers GOOSE messages to the substation network, the state number and sequence number are not changed even though GOOSE data have changed.

Example IV: Suppose that there is a SMV packet insertion to the substation network using captured SMV packets. This action will trigger the SMV threshold violation $\varepsilon_{\mathrm{Th}}^{\mathrm{SV}}$ if the total numbers of SMV packets (inserted packets + normal SMV packet) are higher than the SMV threshold. This will trigger the counter number violation $\theta_{\mathrm{cn}}^{\mathrm{SV}}$ since the inserted SMV packets will violate "SmpCnt" as explained in Section IV-E. This may also trigger the data violation $\mu_{\mathrm{d}}^{\mathrm{SV}}$ if the intruder inserts packets after modification of the SMV messages. It will show an alarm to the operator, who can find more details from the alarm logs and event logs.

APPENDIX III
NOMENCLATURE

α_{Th}^{G}	GOOSE threshold violation indicator.
β_{S}^{G}	GOOSE sequence and state number violation indicator.
γ_{Ti}^{G}	GOOSE time violation indicator.
δ_{d}^{G}	GOOSE data violation indicator.
ε_{Th}^{SV}	SMV threshold violation indicator.
θ_{cn}^{SV}	SMV counter number violation indicator.
μ_{d}^{SV}	SMV data violation indicator.
ψ^{a}	Intrusion attempts upon user-interface or IEDs host-based anomaly indicator (HAI).
ψ^{cf}	Change of the file system HAI.
ψ^{cs}	Change of IED critical settings HAI.
ψ^{o}	Change of status on switches or transformer taps HAI.
ψ^{G}	GOOSE network-based anomaly indicator (NAI).
ψ^{SV}	SMV network-based anomaly indicator.
T	Predefined time for each anomaly detection indicator.
C_{pkt}	Captured packets in a substation network.
$V_{h}^{\mathbf{\Omega}}$	Substation vulnerability index for host-based anomaly.
V_{n}^{GS}	Substation vulnerability index for network-based anomaly.
G_{sm}	GOOSE source MAC address.
G_{dm}	GOOSE destination MAC address.
G_{at}	Anomaly detection thread for GOOSE.
G_{cnp}	Captured number of GOOSE packets.
G_{st}	State number of GOOSE packets.
G_{sq}	Sequence number of GOOSE packets.
G_{th}	Predefined threshold for GOOSE packets (depending on the re-transmission time).
G_{th}^{T}	Predefined time for GOOSE threshold violation detection.
G_{ge}^{T}	GOOSE packet, time at which it is generated.
G_{re}^{T}	GOOSE packet, time at which it is received.
G_{tr}^{T}	GOOSE transfer time (4 ms, defined in IEC 62351-1 [26]).
G_{cp}	Data of captured GOOSE packet.
S_{th}	Predefined threshold for Sampled Values packets (depending on the sampling rate).
S_{cnp}	Captured number of Sampled Values packets.
S_{cp}	Captured SMV packet.
S_{mc}	SMV message counter.
S_{ds}	Object reference of the data set (datSet).
S_{id}	Value of attributes MsvID of the MSVCB (smvID) [29].
S_{sm}	SMV source MAC address.
S_{dm}	SMV destination MAC address.
S_{si}	SMV synchronization indicator ($true$ = synchronized by a clock signal, $false$ = not synchronized).
S_{th}^{T}	Predefined time for SMV threshold violation detection.

REFERENCES

[1] A. R. Metke and R. L. Ekl, "Security technology for smart grid networks," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 99–107, Jun. 2010.

[2] F. Li, W. Qiao, H. Sun, H. Wan, J. Wang, Y. Xia, Z. Xu, and P. Zhang, "Smart transmission grid: Vision and framework," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 168–177, Sep. 2010.

[3] M. Kezunovic, "Smart fault location for smart grids," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 11–22, Mar. 2011.

[4] *NIST Framework and Roadmap for Smart Grid Interoperability Standards*, NIST 1108R2, National Institute for Standards and Technology, Feb. 2012 [Online]. Available: http://www.nist.gov/smartgrid/upload/NIST_Framework_Release_2-0_corr.pdf

[5] J. D. McDonald, *Electric Power Substations Engineering*. Boca Raton, FL, USA: CRC, 2012.

[6] J.-W. Wang and L.-L. Rong, "Cascade-based attack vulnerability on the US power grid," *Safety Sci.*, vol. 47, no. 10, pp. 1332–1336, Dec. 2009.

[7] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, Jan. 2012.

[8] V. Chandola, B. Arindam, and K. Vipin, "Anomaly detection: A survey," *ACM Comput. Surveys (CSUR)*, vol. 41, no. 3, pp. 15–58, Jul. 2009.

[9] J. P. Anderson, *Computer Security Threat Monitoring and Surveillance*. Washing, PA, USA: James P. Anderson Co., 1980.

[10] D. E. Denning, "An intrusion detection model," in *Proc. 7th IEEE Symp. Security Privacy*, May 1986, pp. 119–131.

[11] J. Hall, M. Barbeau, and E. Kranakis, "Anomaly-based intrusion detection using mobility profiles of public transportation users," in *Proc. IEEE Int. Conf. Wireless Mobile Comput. Netw. Commun. (WiMob'2005)*, Aug. 2005, vol. 2, pp. 17–24.

[12] B. Sun, F. Yu, K. Wu, Y. Xiao, and V. C. M. Leung, "Enhancing security using mobility-based anomaly detection in cellular mobile networks," *IEEE Trans. Veh. Technol.*, vol. 55, no. 4, pp. 1385–1396, Jul. 2006.

[13] N. Ye, S. M. Emran, Q. Chen, and S. Vilbert, "Multivariate statistical analysis of audit trails for host-based intrusion detection," *IEEE Trans. Comput.*, vol. 51, no. 7, pp. 810–820, Jul. 2002.

[14] J. Hu, X. Yu, D. Qiu, and H.-H. Chen, "A simple and efficient hidden Markov model scheme for host-based anomaly intrusion detection," *IEEE Netw.*, vol. 23, no. 1, pp. 42–47, Jan. 2009.

[15] M. J. Shevenell and R. F. Erbacher, "Design and implementation of an open network and host-based intrusion detection testbed with an emphasis on accuracy and repeatability," in *Proc. 9th Int. Conf. Inf. Technol.: New Generations (ITNG)*, Apr. 2012, pp. 409–416.

[16] S. Bijan and A. M. Kazemitabar, "HIDMN: A host and network-based intrusion detection for mobile networks," in *Proc. Int. Conf. Comput. Electr. Eng. (ICCEE)*, Dec. 2008, pp. 204–208.

[17] U.-K. Premaratne, J. Samarabandu, T.-S. Sidhu, R. Beresh, and J.-C. Tan, "An intrusion detection system for IEC 61850 automated substations," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2376–2383, Oct. 2010.

[18] C.-W. Ten, C.-C. Liu, and G. Manimaran, "Vulnerability assessment of cybersecurity for SCADA systems," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1836–1846, Nov. 2008.

[19] C.-W. Ten, J. Hong, and C.-C. Liu, "Anomaly detection for cybersecurity of the substations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 865–873, Dec. 2011.

[20] S.-M. Amin and A. M. Giacomoni, "Smart grid- safe, secure, self-healing: Challenges and opportunities in power system security, resiliency, and privacy," *IEEE Power Energy Mag.*, pp. 33–40, Jan. 2012.

[21] G. Dan, H. Sandberg, M. Ekstedt, and G. Bjorkman, "Challenges in power system information security," *IEEE Security Privacy*, vol. 10, no. 4, pp. 62–70, Jul. 2012.

[22] D. Kushner, "The real story of Stuxnet," *IEEE Spectrum*, vol. 50, no. 3, pp. 48–53, Mar. 2013.

[23] M. Govindarasu, A. Hann, and P. Sauer, "Cyber-physical systems security for smart grid," Future Grid Initiative White Paper, PSERC, Feb. 2012 [Online]. Available: http://www.pserc.wisc.edu/documents/publications/papers/fgwhitepapers/Govindarasu_Future_Grid_White_Paper_CPS_May_2012.pdf

[24] Government Accountability Office (GAO), "Electricity grid modernization: Progress being made on cyber security guidelines, but key challenges remain to be addressed," GAO-11-117, Jan. 2011 [Online]. Available: http://www.gao.gov/new.items/d11117.pdf

[25] *Guidelines for Smart Grid Cyber Security*, NISTIR 7628, National Institute for Standards and Technology, Aug. 2010 [Online]. Available: http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf

[26] *Standard: Part 1: Communication Network and System Security—Introduction to Security Issues*, IEC TS 62351-1, Power Systems Management and Associated Information Exchange—Data and Communications Security, May 2007, 1st Ed..

[27] P. Judge and M. Ammar, "Security issues and solutions in multicast content distribution: A survey," *IEEE Netw.*, vol. 17, no. 1, pp. 30–36., Jan. 2003.

[28] C.-C. Liu, A. Stefanov, J. Hong, and P. Panciatici, "Intruders in the grid," *IEEE Power Energy Mag.*, vol. 10, no. 1, pp. 58–66, Feb. 2012.

[29] *Standard: Sampled Values over ISO/IEC 8802-3*, IEC 61850 9-2, Specific Communication Service Mapping (SCSM), Apr. 2004, 1st Ed..

[30] *Standard: Mapping to MMS (ISO/IEC9506-1 and ISO/IEC 9506-2)*, IEC 61850 8-1, Specific Communication Service Mapping (SCSM), May 2004, 1st edition.

[31] *Standard: Communication Requirements for Functions and Device Models*, IEC 61850-5, Communication Networks and Systems in Substations, Jul. 2003, 1st Edition.

[32] U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability, National SCADA Test Bed Program Jan. 2008 [Online]. Available: http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/DOE_OE_NSTB_Multi-Year_Plan.pdf

[33] A. Hahn and M. Govindarasu, "Cyber attack exposure evaluation framework for the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 835–843, Jun. 2011.

[34] G. Dondossola, F. Garrone, G. Proserpio, and C. Tornelli, "Impact of DER integration on the cyber security of SCADA systems—The medium voltage regulation case study," in *Integration of Renewables into the Distribution Grid, CIRED Workshop*, May 2012, pp. 1–4.

[35] M. Hillenmeyer, *Machine Learning*. Stanford, CA, USA: Stanford Univ., Jul. 2005 [Online]. Available: http://www.stanford.edu/~maureenh/quals/pdf/ml.pdf

[36] A. Harvey, "Cybersecurity enhancement in a power substation," M.S. Thesis, School of Electrical, Electronic and Mechanical Engineering, University College Dublin, Dublin, Ireland, 2011.

[37] tshark [Online]. Available: http://www.wireshark.org/docs/man-pages/tshark.html

Junho Hong (S'08) received his B.S.E.E. and M.S.E.E. from Myongji University, Korea, in 2008 and 2010, respectively. He is pursuing his Ph.D. at Washington State University, Pullman, WA, USA. His research interests include cyber-physical security of EMS/DMS/AMI systems, substation automation, and power system restoration.

Chen-Ching Liu (F'94) serves as Boeing Distinguished Professor at Washington State University, Pullman, WA, USA, and Professor at University College Dublin, Ireland. He was Palmer Chair Professor at Iowa State University and a Professor at the University of Washington. Dr. Liu served as Chair of the IEEE PES Technical Committee on Power System Analysis, Computing and Economics (PSACE).

Manimaran Govindarasu (SM'10) is a Professor in Electrical and Computer Engineering at Iowa State University, Ames, IA, USA. His expertise is in the areas of real-time systems, cyber security, cyber-physical systems security of power grids. He co-authored *Resource Management in Real-Time Systems and Networks* (MIT Press, 2001). He is chairing the Cyber Security Task Force at IEEE PES CAMS Subcommittee.