Multimed Tools Appl (2015) 74:303–318
DOI 10.1007/s11042-014-1870-0

Novel Approach for Detecting Network Anomalies for Substation Automation based on IEC 61850

Hyunguk Yoo · Taeshik Shon

Published online: 2 March 2014

Abstract An SA (Substation Automation) system based on IEC 61850 is an intelligent substation; it has been receiving considerable attention as a core component of a smart grid. The explosive increase of threats to cyber security has been expanded to critical national infrastructures including the power grid. Substation Automation has also become a main target of cyber-attacks. Currently, various countermeasures such as firewalls, IDS (Intrusion Detection System)s, and anti-virus solutions have been developed, but to date, these have not sufficiently reflected the inherent features of Substation Automation based on IEC 61850. This study suggests a method of anomaly detection for MMS (Manufacturing Message Specification) and GOOSE (Generic Object Oriented Substation Events) packets, the main communication protocols of IEC 61850 Substation Automation. 3-Phase preprocessing, EM (Expect Maximization), and one-class SVM (Support Vector Machine) techniques are applied. The effectiveness of the suggested method is evaluated through experiments.

Keywords IEC 61850 · Substation automation · Smartgrid · Anomaly detection · Machine learning · EM · SVM

1 Introduction

Substation Automation is a core component representing transmission and transformation of electricity and complies with the IEC 61850 communication standards. Using SA, it is possible to maximize operation efficiency and compatibility with more extensive information. However, as the substation system has changed from a closed structure in the past to one that is partially open (for example, increased connection points, use of commercial software, and use of publicized communication protocols), the threat of cyber-attack has increased. According to recent research by the United States Congress, electric utility companies in the US are being cyber-attacked up to 10,000 times a month [9], some of which are assumed to be APT (Advanced Persistent Threat) type attacks, such as Stuxnet.

H. Yoo · T. Shon (✉)
Ajou University, Suwon, Gyeonggi-do, Republic of Korea
e-mail: tsshon@ajou.ac.kr

H. Yoo
e-mail: cielo1025@ajou.ac.kr

There are anti-virus, firewall, and intrusion detection systems designed as security solutions for the general IT environment. However, when applying these existing IT security technologies to substation automation networks based on IEC 61850, there are many difficulties. For example, with the existing technologies, it is difficult to interpret the SA network packets because they use industrial-oriented protocols. Furthermore, most of the existing security solutions detect an attack by referring to a blacklist database with signature information on each attack. With this method, zero-day attacks that use unknown vulnerabilities cannot be detected. Furthermore, databases must be updated whenever a new attack pattern occurs. In an environment such as a power grid, where availability is critical, problems such as delayed service due to database updates could be a serious issue.

As a result, some security vendors are now developing security solutions specialized for industrial environments. Tofino Modbus TCP [16] Enforcer, installed in a firewall, can conduct a check for a Modbus protocol packet. Application Control [10] controls applications operating in control equipment using an application-whitelisting technique. There are, however, few studies addressing intrusion detection on SA networks based on the IEC 61850 standard.

In this paper, we present an anomaly-detection model for IEC 61850 SA networks through normal-behavior profiling of the MMS and GOOSE packets that are used for an SA network. An industrial control environment such as an SA network has a feature that makes anomaly detection through normal-behavior profiling easy. The occurrence of network traffic is regular and only restricted protocols are used. Based on this feature, this study proposes a power grid-based security technique to manage unknown attacks including an APT attack.

The remainder of this paper is organized as follows. In Section 2, we briefly discuss the related work. In the following section, the network configuration of the IEC 61850 SA network and the features of its traffic are described. Section 4 presents our anomaly-detection mode. We evaluate the proposed model in Section 5, and conclude the paper in Section 6.

2 Related works

The techniques used in intrusion-detection systems (IDS) can be largely divided into misuse detection and anomaly detection, depending on the method of intrusion identification [7]. The misuse detection technique detects attacks using a database in which attack signatures are saved. This has the advantage of high accuracy and speed. It also has some limitations for use in control systems because it cannot detect novel (zero-day) attacks and requires frequent database updating.

An anomaly-detection technique is a method that first establishes the normal behavior of a network or system and then considers behaviors that deviate from the established standard of normal behavior. Because it is based on normal behavior, it can detect a new type of attack and does not require frequent updating. The established normal behaviors of a control system have a limited range and the domain changes only slightly. The biggest disadvantages of an anomaly-detection technique are: 1) it generally has a high false-positive rate; and 2) because there is no established principle, at the beginning it demands considerable effort and time to find an optimized normal-behavior model. Because the environment of a control system has limited services and its behavior patterns are regular compared to general IT environments, it is an excellent candidate for an anomalydetection technique [2, 11, 19, 20].

Steven Cheung et al. [4] detected anomaly symptoms in a Modbus TCP network using three model-based techniques (protocol level, communication pattern, and learning-based). A protocol-level model establishes supportable function codes based on the Modbus protocol specification and considers Modbus packets violating these as anomaly packets. A communication pattern-based model establishes allowable packet groups based on the IP and TCP port

and identifies packets not included in such groups as abnormal. A learning-based model constructs a Bayesian Network using the Modbus function codes and detects abnormal symptoms using the relation of conditional probability.

Partrick Dussel et al. [6] suggested a payload-based real-time anomaly-detection system. This system presents a feature space by dividing the TCP payload data by n-byte and detects abnormal packets by comparing similarity with a normal byte sequence. This technique can be applied to the upper layer protocol because the TCP payload is divided by n-gram in a lump.

Dayu Yang et al. [21] used a pattern-matching technique for anomaly detection. First, traffic profiles of a normal network and abnormal network are made using an AAKR (Auto-Associative Kernel Regression) model. When the objects to be detected arrive, they are judged as normal or abnormal by comparing patterns using a SPRT (Sequential Probability Ratio Test). Upeka Premaratne et al. [12], for detection of intrusions to an IEC 61850 network, suggested a system that evaluates whether it is normal or abnormal using an SVM algorithm based on the traffic data. In this technique, normal or abnormal is classified using the SVM algorithm assuming the situation of a rapid increase of traffic as an abnormal behavior.

An anomaly-detection algorithm proposed by Chee-Wooi Ten et al. [15] establishes four abnormal events (attempt of intrusion, falsification of file system, change of system settings, and change of system states) that have different weighted values and constructs its matrix by calculating them at regular time intervals. It is regarded as abnormal when the deviation of the values measured at a neighboring time exceeds a specific standard value. Barbosa, R.R.R et al. [1] suggested a methodology that identifies an abnormal symptom when a traffic pattern is different from the normal traffic that was modeled through all the flows occurring in the SCADA system.

As we can see from the above works, there have been some studies on anomaly detection for control systems. There have not been, though, any published studies on anomaly detections suitable for the features of the IEC 61850 protocols (MMS and GOOSE). Though the system proposed by Upeka Premaratne et al. addressed the IEC 61850 network, it differs from what we are proposing in this study. It was modeled by simply differentiating a situation of rapid download from a normal situation, rather than using a modeling method suitable for the features of the GOOSE and MMS protocols.

Our study uses the widely accepted machine learning algorithms including EM (Expectation Maximization), LOF (Local Outlier Factor), and one-class SVM (Support Vector Machine). EM is a clustering algorithm based on probability. LOF is an algorithm to find outliers by comparing the density of the data [3, 5]. In this study, the outliers that could adversely affect the learning process were excluded by evaluating them using EM or LOC, selectively. One-class SVM is an unsupervised classification algorithm with only one class that can generate a pattern model [13]. A one-class SVM algorithm attempts to find the hyper-plane that has the longest distance from the origin point, after moving the input data as one class to a different dimensional space using various kernel functions [14]. This hyper-plane is the criterion to differentiate the learning class from the other classes. In the technique proposed in this study, it is used to determine whether test packets are normal or abnormal, after learning normal behaviors using the one-class SVM algorithm.

3 IEC 61850 substation automation

The serial-based IEC 60870-5 series and DNP3 were developed in the 1980s and have been used as a data interchange protocol in substations for many years. The IEC 60870-5 series, however, has many disadvantages. It has low compatibility between different vendors due to its high dependence on hardware and its functionality is basic, having only relevant point

information and status information. To overcome the deficiency problems of expandability and flexibility of the communication protocol in the existing substation systems, the IEC TC57 WG10 developed IEC 61850 as a new communication standard suitable for the next-generation of substation automation systems.

MMS and GOOSE are typical protocols used in IEC 61850. MMS is a TCP/IP-based protocol used for communication and ordinary control-order transmissions between the server and client. GOOSE is an Ethernet protocol used for peer-to-peer communication for transmitting IED (Intelligent Electronic Device) state information. Fig. 1 shows the IEC 61850 SA architecture and the range of uses of MMS and GOOSE protocols.

Cases or methods of the attacks against the MMS and GOOSE protocols have not yet been reported. However, the implementation vulnerability of the TPKT layer in the MMS protocol stack has been identified [17, 18]. There is a high possibility that attackers can exploit this because the basic MMS and GOOSE protocols do not provide encryption or authentication.

4 Proposed approach

In this study, we propose a method of normal-behavior profiling for the MMS and GOOSE packets to detect abnormal symptoms in the network of an IEC 61850 substation. As discussed in the related works section, the studies on anomaly detection in existing control systems are not suitable to an IEC 61850 environment. Most of them are based on the Modbus or DNP3 protocols, or do not consider the TCP upper-layer protocols. This study suggests a newly developed 3-phase preprocessing technique for the GOOSE/MMS packets, a normal-behavior grouping method to which we apply an unsupervised-learning-based EM algorithm, and methods of learning normal behaviors and detecting abnormal behaviors. The suggested IEC 61850 anomaly-detection system is a network-level detection system that collects

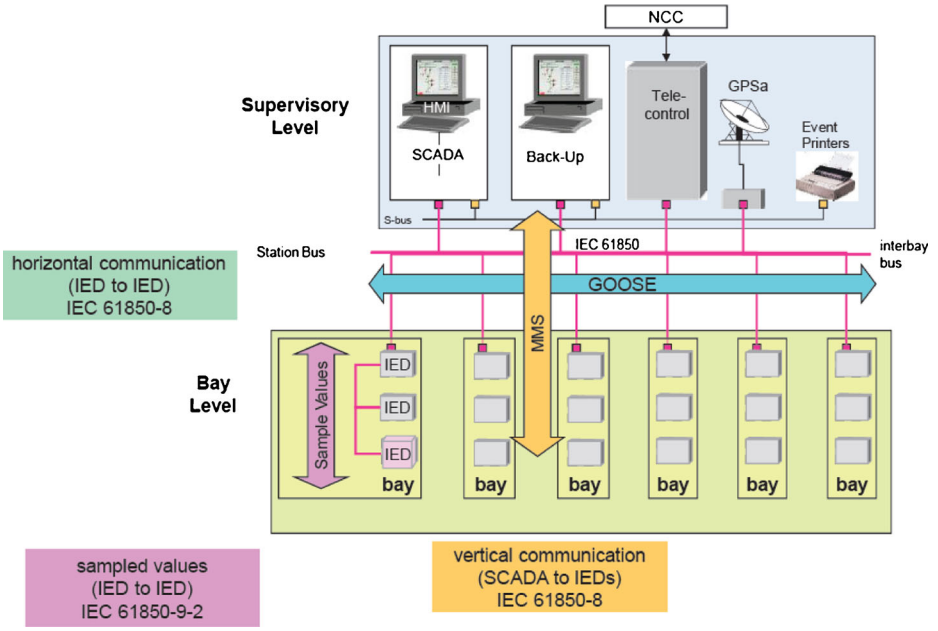

Fig. 1 IEC 61850 substation automation architecture [8]

communication packets from the equipment at the IEC 61850 network substation and bay levels, and detects anomaly symptoms from this data. Fig. 2 shows the locations of the packet collection and detection in the network of an IEC 61850 substation.

An IEC 61850 anomaly-detection system typically consists of pre-processing, normal-behavior learning, and anomaly detection. The detailed architecture is shown in the following Fig. 3.

In the pre-processing, MMS and GOOSE packets are extracted by performing packet filtering with the collected substation-network packet data. Then, the MMS and GOOSE packets are grouped into different sets of data through 3-phase preprocessing (single packet process, sequence packet process, and packet traffic process). In the single packet process, each datum represents one packet unit. In the sequence packet process, consecutive packets in the same flow are grouped into a datum. In the packet traffic process, a datum is made by calculating the traffic rate. In order to remove the outliers from the sets made through these processes, EM or LOF is applied. Using normal-behavior learning, the normal-behavior models for three sets of data are created by applying a one-class SVM algorithm to the data sets from which outliers have been eliminated. This is then installed in the anomaly-detection engine. In the anomaly-detection engine, pre-processing for real-time packets is performed by sending them to the pre-processing module. The pre-processed packets are compared with the normal-behavior model already installed. The packets are determined to be normal or abnormal based on this comparison and an alarm and log are updated.

4.1 3-phase preprocessing

- Packet Filtering
  In the packet-filtering module, using the fingerprint of the MMS and GOOSE packets, only MMS and GOOSE packets are extracted from the entire set of packets.
- Single Packet Process
  The packet fields selected to create the data sets in the single-packet process section are shown in Tables 1 and 2. The factor considered important in selecting the fields is the value change between packets. We excluded the fields that were expected to have no value change between packets. Meanwhile, for the MMS packet, the ngram method was selected

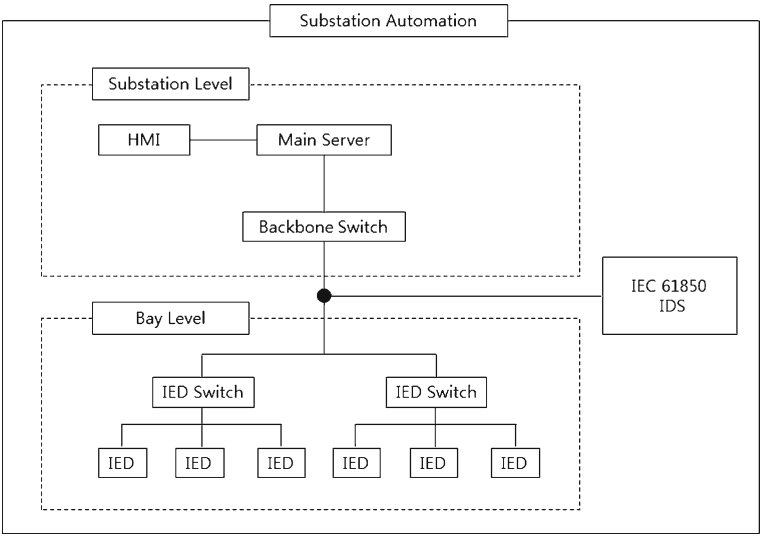

Fig. 2 Location of IDS in the IEC 61850 substation network

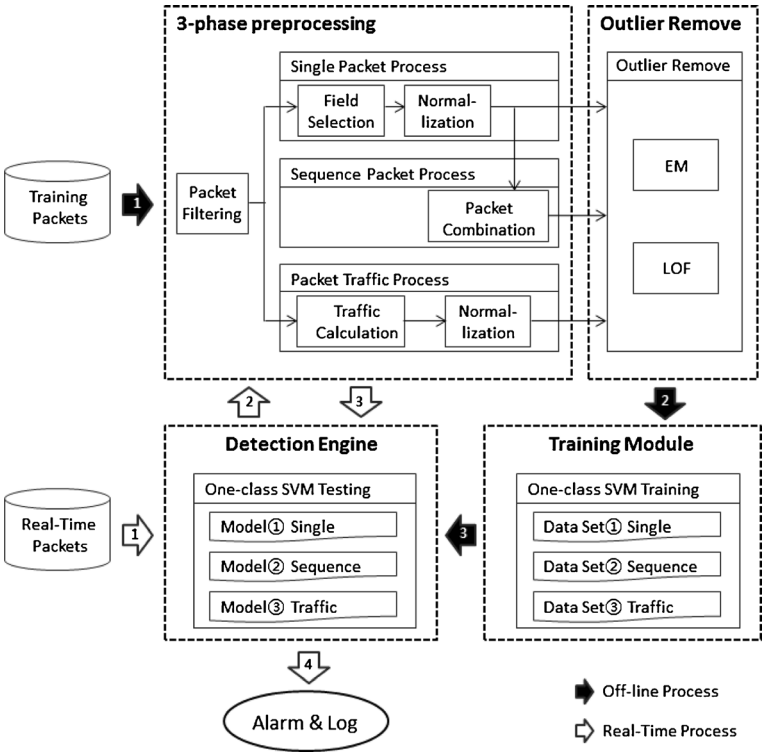

Fig. 3 IEC 61850 anomaly detection system architecture

to bring the MMS payload to the feature space. The reason for selecting this method for the MMS message only is that the fields of the MMS message section are variable, depending on the kind of message. It is difficult to get significant results, as the entire data set becomes a sparse matrix when all of the items are brought to the feature space.

Because the ranges of the values of the selected data fields are different, it is necessary to perform a normalization process. The normalization was conducted using the mean and standard deviation, which are commonly used.

$$\text{Mean} \qquad \overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\text{Variance} \qquad \hat{\sigma}^2 = \frac{\sum_{i}^{n} (x_i - \overline{x})^2}{n = 1}$$

$$\text{Standard deviation} \qquad \hat{\sigma} = \sqrt{\frac{\sum_{i}^{n} (x_i = \overline{x})^2}{n-1}}$$

$$x = \frac{x_i = \overline{x}}{\sigma}$$

Table 1 Features of single MMS packet

Field		Description
Ethernet	Destination MAC	Destination MAC Address
	Source MAC	Source MAC Address
IP	Total Length	Total Length of IP Packet
	Identification	Datagram Identification Number
	IP Flags	Fragment Flags
	TTL	Time To Live
	Source IP	Source IP Address
	Destination IP	Destination IP Address
TCP	Source Port	Source Port Number
	Destination Port	Destination Port Number
	Sequence Number	Sequence Number of Segment
	Ack Number	Acknowledgement Number
	TCP Flags	TCP Session Control
	Window	Buffer Size for Flow Control
TPKT	Length	Length of TPKT
COTP	EOT	End of TSDU mark
MMS	1–20 Bytes	MMS Message

- Sequence Packet Process

  In the sequence-packet process, data sets are created by grouping consecutive packets into one in the communication between same communication nodes to a set, based on the results from the single-packet process section. At this time, overlapping fields such as the MAC/IP address and port number are selected from the first packet only. The non-overlapping fields are selected from the succeeding packets. The reason for grouping consecutive packets into one is that messages

Table 2 Features of single GOOSE packet

Field		Description
Ethernet	Destination MAC	Destination MAC Address
	Source MAC	Source MAC Address
GOOSE	APPID	Application Identification
	Length	GOOSE message Length
	gocbRef	GOOSE Control Block Reference
	timeAllowedtoLive	Maximum Wait Time for Message
	datset	Object Reference of Control Block
	goID	GOOSE Message Identification
	time	Time to stNum Increase
	stNum	Status Number
	sqNum	Sequence Number
	confRev	Configuration Revision
	numDatSetEntries	Number of Data Set Entries

Table 3 Features of MMS packet traffic

Field		Description
IP	Source IP	Source IP Address
	Destination IP	Destination IP Address
Traffic	PPS	Packets per Second
	BPS	Bytes per Second

containing attack information may be cut into the IP datagram unit and may also contain those not considered during the learning one-grouped data of consecutive packets as a normal behavior.

- Packet Traffic Process

  The traffic-based process creates data sets by calculating the packet transfer rate and transfer byte size per unit of time. The normal-behavior model learned through the traffic-based data sets can detect an attack such as DoS (Denial of Service). Tables 3 and 4 show the fields contained in the MMS and GOOSE data sets, respectively.

4.2 Outlier removal

The data sets, before normalization, require outlier processing. This is needed to assure that all the data collected to model the normal-behavior are, in fact, normal. It is highly possible that outlier data can be included due to administrators' error, equipment defects, and noise, even when the substation network is operating normally. Therefore, outlier processing is essential. In order to remove the outliers, we applied EM clustering and LOF techniques, selectively. When we applied EM, we considered relatively small clusters as outliers and removed them from the learning data. Tables 5 and 6 are the results of clustering the MMS and GOOSE packets by EM, respectively. For clustering by EM, we used WEKA (http://www.cs.waikato.ac.nz/ml/weka/). Fig. 4 and Fig. 5 show the results of EM algorithm for MMS and GOOSE by WEKA tool.

When clustering the MMS data using the EM algorithm, the value of the log likelihood was highest when the size of the cluster was 3. We also determined that compared to clusters 2 and 3 in the clustering results, the size of cluster 1 was very

Table 4 Features of GOOSE packet traffic

Field		Description
Ethernet	Destination MAC	Destination MAC Address
	Source MAC	Source MAC Address
Traffic	PPS	Packets per Second
	BPS	Bytes per Second

Table 5 Single MMS packet clusters

Number of input data	100,000	
Clustering result	Cluster number	Number of data
	1	9,493 (9 %)
	2	58,013 (58 %)
	3	32,494 (32 %)
Log likelihood	165.5724	

small. Therefore, only the data contained in clusters 2 and 3 was used for the normalbehavior learning.

Using the EM clustering results of the GOOSE packets as shown in Table 6, we performed normal-behavior learning with clusters 2, 4, and 6. Clusters 1, 3, and 5 were considered too small.

4.3 Normal-behavior training

The method used to develop the normal-behavior model used in this study is the one-class SVM algorithm. SVM is widely accepted as a binary classification algorithm with excellent performance. However, in a control system environment where learning of attack packets is difficult, supervised learning techniques including binary class SVM are difficult to use. In this study, therefore, we used a one-class SVM that can derive a learning model with only a normal data class, similar to an unsupervised learning technique. There are kernel functions such as linear, polynomial, sigmoid, and RBF (Radial Basis Function) that can be used in a one-class SVM. Experimental results with the Libsvm 3.14 library in this study showed that the linear kernel and sigmoid kernel had the best performance (see Table 5). The results of these two were almost identical. Other experiments showed that the accuracy of the sigmoid kernel was slightly higher than that of the linear kernel.

Table 6 Single GOOSE packet clusters

Number of input data	171,011	
Clustering result	Cluster number	Number of data
	1	10,191 (6 %)
	2	58,570 (34 %)
	3	7,204 (4 %)
	4	54,832 (32 %)
	5	10,919 (6 %)
	6	29,295 (17 %)
Log likelihood	55.92068	

	Cluster		
Attribute	0	1	2
	(0.09)	(0.58)	(0.32)
att_1			
mean	0.0003	0.018	0.0157
std. dev.	0.0002	0.1254	0.1164
att_2			
mean	0.9496	0.9171	0.7489
std. dev.	0.0105	0.1389	0.353
att_3			
mean	0.1868	0.2072	0.026
std. dev.	0.3454	0.3405	0.1499
att_4			
mean	0.6098	0.3257	0.7105
std. dev.	0.41	0.4198	0.3688
att_5			
mean	0.144	0.5437	0.0479
std. dev.	0.0296	0.1877	0.0305

Fig. 4 Single MMS packet clustering result using EM algorithm

Hence, we used the sigmoid kernel for the learning in our suggested detection system.

4.4 Anomaly detection

Anomaly detection judges whether the packets collected in real time are normal or abnormal using the oneclass SVM learning model generated in Section 4.3. The detection results of each of the MMS and GOOSE packets are saved as log records. These are then added to the learning model to improve the detection performance. Generally, the detection performance when using the machine-learning algorithm is highly dependent on the pre-processing module. Accordingly, in order to improve the detection performance, it is necessary to learn and test the pre-processing technique and the features of the applied domains sufficiently.

Attribute	Cluster 0 (0.06)	1 (0.34)	2 (0.04)	3 (0.32)	4 (0.06)	5 (0.17)
att_1						
mean	0.0041	0.0041	0.0041	0.0041	0.0064	0.0041
std. dev.	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006
att_2						
mean	0.0039	0.0039	0.0039	0.0039	0.0166	0.004
std. dev.	0	0	0	0	0	0.0001
att_3						
mean	0.774	0.0008	0.0001	0.0001	0.0025	0.0015
std. dev.	0.0258	0.0009	0.1832	0	0.1832	0.0015
att_4						
mean	0.0241	0.0167	0.9942	0.7526	0.0166	0.0045
std. dev.	0.1259	0.0076	0	0.2465	0	0.0039
att_5						
mean	0.0086	0.0107	0	0.0075	0	0.0122
std. dev.	0.0016	0.0036	0.0048	0.0029	0.0048	0.0048

Fig. 5 Single GOOSE packet clustering result using EM algorithm

5 Evaluation

In order to verify the validity of the technique proposed in this study, we conducted experiments using the packets collected from an actual operating IEC 61850 substation. The developed IEC 61850 anomaly-detection system was run on Linux (Ubuntu 12.04) and used functions from the Libsvm v3.14 library to apply the oneclass SVM. The grouping of the normal behavior with EM was performed off-line using WEKA v3.6.8 on Windows [22, 26]. For the models based on sequence packet and packet flow, because memory space and time are spent in pre-processing, the performance evaluation was first conducted on the single-packet model only, to evaluate the field applicability. Furthermore, because there is no known attacking technique or packet based on the IEC 61850 protocol, the FPR (False positive rate) for normal packets was first evaluated. In the experiments, the value of the error tolerance used for the one-class SVM algorithm was

Table 7 One-class SVM kernel function accuracy comparison

Kernel Name	f (x, y)	Detection Accuracy	
		MMS	GOOSE
Polynomial	(g*x*y+r)d	97.6991 % (63734/65235)	94.1239 % (12334/13104)
Radial Basis	exp (−g *‖x−y‖2)	91.7851 % (59876/65235)	81.9368 % (10737/13104)
Linear	x * y	97.8294 % (63819/65235)	94.1239 % (12334/13104)
Sigmoid	tanh (g*u*y+r)	97.8294 % (63819/65235)	94.1239 % (12334/13104)

Parameters

d degree in kernel function (default : 3)

g gamma in kernel function (default : 1/number of features)

r coef0 in kernel function (default : 0)

nu nu in kernel function (default : 0.01)

fixed at 0.01, and the FPR was evaluated by a 10-fold cross-validation method using learning-packet data and a method using new test packets. Fig. 6 show the results of one-class SVM classification using WEKA tool. Fig. 7 and Table 8 represents the detection results and accuracy of our intrusion detection system, respectively.

Because we could not conduct the detection function in the above experiment, we could not measure the FNR (False Negative Rate). We did identify that the value of FPR is in the range of 1–6 %. In order to be used in an actual operating environment, the FPR must be further reduced. In addition, an evaluation of the FNR through an attack simulation is needed. Considering that this is a prototype program addressing the features of the IEC 61850 protocol, there is sufficient room for improvement.

6 Conclusion

In digital substations, standard protocols defined in IEC 61850 are used. They enable the interchange of information including control signals between vendors of different products using common standards. However, there are no algorithms or methods that are able to detect anomaly symptoms in the IEC 61850 communication packets. As the threats of cyber-attack to industrial infrastructures such as the power grid are increasing, it is urgent that this vulnerability be addressed. In this study, we proposed an anomaly-detection method for MMS and GOOSE packets. These are the main

- MMS -

Time taken to build model: 4.89 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	25099		98.9825 %
Incorrectly Classified Instances	0		0 %
Kappa statistic	1		
Mean absolute error	0		
Root mean squared error	0		
Relative absolute error	NaN	%	
Root relative squared error	NaN	%	
UnClassified Instances	258		1.0175 %
Total Number of Instances	25357		

- GOOSE -

Time taken to build model: 1.98 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	22249		98.5778 %
Incorrectly Classified Instances	0		0 %
Kappa statistic	1		
Mean absolute error	0		
Root mean squared error	0		
Relative absolute error	NaN	%	
Root relative squared error	NaN	%	
UnClassified Instances	321		1.4222 %
Total Number of Instances	22570		

Fig. 6 One-class SVM classification result using WEKA tool

communication protocols in the IEC 61850 substation automation system. In the proposed technique, we extracted the principal fields from the MMS and GOOSE packets using 3-phase preprocessing, generated three data sets for each protocol, and removed outliers by applying an EM technique. For normal-behavior learning and detection, we applied a one-class SVM, one of the most popular learning algorithms. In order to verify the validity of the proposed technique, we evaluated the detection accuracy with data sets collected from actual field data using a program developed in a Linux environment.

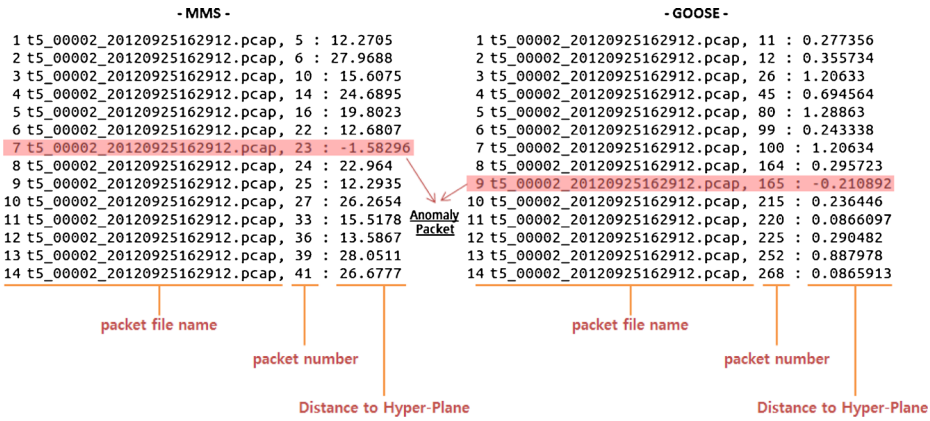

Fig. 7 Detection results in our proposed intrusion detection system

There have been few studies of anomaly-detection techniques reflecting the features of the IEC 61850 protocol. The anomaly-detection technique proposed in this study should operate efficiently in a IEC 61850-based digital substation environment using the features of the IEC 61850 protocol. However, in order to apply it in the field, its detection accuracy must be improved and an evaluation of its performance for attacking packets must be conducted. One method to improve its detection accuracy is to apply different normal-behavior models to the flow between nodes. To evaluate the ability to detect attack packets, a simulation should be developed using attack packets for IEC 61850.

Table 8 Accuracy of our proposed intrusion detection model

	MMS	GOOSE
Error tolerance	0.01	
The number of training packets	25,357	22,570
FPR with 10-fold validation (Mis-classified Packets/Total Validation Packets)	1.0175 % (258/25357)	1.4222 % (321/22570)
FPR with Test Packets (Mis-classified Packets/Total Test Packets)	2.1706 % (1416/65235)	5.8761 % (770/13104)

Acknowledgements This work was supported by the Power Generation & Electricity Delivery Core Technology Program of the Korea Institute of Energy Technology Evaluation and Planning(KETEP) granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (No. 20131020402090).

References

1. Barbosa RRR, Pras A (2010) Intrusion detection in SCADA networks. Mechanisms for autonomous management of networks and services. Springer, Berlin
2. Barbosa RRR, Sadre R, Pras A (2012) Towards periodicity based anomaly detection in SCADA networks. Emerging Technologies & Factory Automation (ETFA), 2012 I.E. 17th Conference on. IEEE
3. Breunig MM et al (2000) LOF: identifying density-based local outliers. ACM Sigmod Rec 29(2), ACM
4. Cheung S, Dutertre B, Fong M, Lindqvist U, Skinner K, Valdes A (2007) Using model-based intrusion detection for SCADA networks. SCADA Security Scientific Symposium
5. Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B Methodol 39:1–38
6. Dussel P, Gehl C, Laskov P, Buber J-U, Stormann C, Kastner J (2010) Cyber-critical infrastructure protection using real-time payload-based anomaly detection. Critical Information Infrastructures Security
7. Garitano I, Uribeetxeberria R, Zurutuza U (2010) A review of SCADA anomaly detection systems. Intelligent and Soft Computing
8. Kirrmann H (2012) Introduction to the IEC 61850 electrical utility communication standard. ABB
9. Markey EJ, Waxman HA (2013) Electric grid vulnerability: industry responses reveal security gaps.
10. Mcafee. Application control. http://www.mcafee.com/us/products/application-control.aspx
11. Pleijsier E (2013) Towards anomaly detection in SCADA networks using connection patterns
12. Premaratne U, Samarabandu J, Sidhu T, Beresh B, Tan J-C (2008) Evidence theory based decision fusion for masquerade detection in IEC 61850 automated substations. Information and Automation for Sustainability, 2008. ICIAFS 2008. 4th International Conference on. IEEE
13. Schölkopf B et al (2001) Estimating the support of a high-dimensional distribution. Neural Comput 13.7: 1443–1471
14. Shon T, Moon J (2007) A hybrid machine learning approach to network anomaly detection. Information Sciences
15. Ten C-W, Hong J, Liu C-C (2011) Anomaly detection for cybersecurity of the substations. IEEE Transactions on Smart Grid
16. Torfino. Torfino Modbus TCP enforcer. http://www.tofinosecurity.com/products/Tofino-Modbus-TCP-Enforcer-LSM
17. US-CERT, Vulnerability note VU#468798
18. US-CERT, Vulnerability note VU#372878
19. Valdes A, Cheung S (2009) Communication pattern anomaly detection in process control systems. Technologies for Homeland Security, 2009. HST'09. IEEE Conference on. IEEE
20. Valdes A, Cheung S (2009) Intrusion monitoring in process control systems. System sciences, 2009. HICSS'09. 42nd Hawaii International Conference on. IEEE
21. Yang D, Usynin A, Hines JW (2006) Anomaly-based intrusion detection for SCADA systems. 5th Intl. Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technologies (NPIC&HMIT 05)
22. Zhu B, Sastry S (2010) SCADA-specific intrusion detection/prevention systems: a survey and taxonomy. Proceedings of the 1st Workshop on Secure Control Systems

Hyunguk Yoo received B.E. degree in Computer Engineering from Ajou University, Korea in 2011. He is studying in Ajou University in Computer Engineering Integrated Course and working in ICS (Information Communication Security) lab. His research interests include Smartgrid Security, Anomaly Detection, and Digital Forensics.

Taeshik Shon received his Ph.D. degree in Information Security from Korea University, Seoul, Korea and his M.S. and B.S. degree in computer engineering from Ajou University, Suwon, Korea. While he was working toward his Ph.D. degree, he was awarded a KOSEF scholarship to be a research scholar in the Digital Technology Center, University of Minnesota, Minneapolis, USA, from February 2004 to February 2005. From Aug. 2005 to Feb. 2011, Dr. Shon had been a senior engineer in the Convergence S/W Lab, DMC R&D Center of Samsung Electronics Co., Ltd. He is currently a professor at the Division of Information and Computer Engineering, College of Information Technology, Ajou University, Suwon, Korea. His research interests include Convergence Platform Security, Mobile Cloud Computing Security, Mobile/Wireless Network Security, WPAN/WSN Security, anomaly detection algorithms, and machine learning applications.