A Synthesized Dataset for Cybersecurity Study of IEC 61850 based Substation

Partha P. Biswas*, Heng Chuan Tan*, Qingbo Zhu*, Yuan Li*, Daisuke Mashima*, Binbin Chen†*

*Advanced Digital Sciences Center †Singapore University of Technology and Design
{partha.b, hc.tan, qingbo.zhu, yuan.li, daisuke.m, binbin.chen}@adsc-create.edu.sg

***Abstract*—Cyber attacks pose a major threat to smart grid infrastructures where communication links bind physical devices to provide critical measurement, protection, and control functionalities. Substation is an integral part of a power system. Modern substations with intelligent electronic devices and remote access interface are more prone to cyber attacks. Hence, there is an urgent need to consider cybersecurity at the electrical substation level. This paper makes a systematic effort to develop a synthesized dataset focusing on IEC 61850 GOOSE communication that is essential for automation and protection in smart grid. The dataset is intended to facilitate the research community to study the cybersecurity of substations. We present the physical system of a typical distribution level substation and several of its critical electrical protection operation scenarios under different disturbances, followed by several cyber-attack scenarios. We have generated a dataset with multiple traces that correspond to these scenarios and demonstrated how the dataset can be used to support substation cybersecurity research.**

I. Introduction

The complexity of power system is growing day by day with the inclusion of smart devices and communication network. Maintaining network security and resilience is becoming more challenging as the smart grid is vulnerable to cyber attacks. A cyber attack can intrude the computer network system, thus compromising its authenticity, integrity, and availability [1]. As a consequence, the operation of a modern power grid, where large communication network and infrastructure is prevalent, can be jeopardised. In recent years, extensive research works have been carried out to study possible cyber threat scenarios and their mitigation strategies for power grids [2], [3], [4], [5].

Although most research efforts focus on security aspects of the whole power grid, not many publications concentrate on electrical substations. Substations are important entities in the power grid, both at transmission level and distribution level. The primary function of a substation is to convert one voltage level to another and host communication among devices in the station, bay and process levels. The advanced communication needs increase the attack surface of the substation, potentially enhancing the probability and quantum of cyber attacks. To detect these attacks, many cybersecurity solutions such as the intrusion detection systems (IDS) have been proposed. However, there is currently no realistic substation dataset that can be used to validate those.

At present, many researchers adopt heuristics reasoning as a mean of analyzing the security, which may lead to certain bias. Others have developed their own datasets which are non-standard and unpublished. The difficulties in developing standard datasets can be attributed to the gap in knowledge across various domains (e.g., the domain knowledge required to understand the various functionalities and operating modes of a substation and the cybersecurity knowledge is needed about the evolving threats and attack techniques), as well as the absence of standard communication protocols in electrical substations prior to the advent of IEC 61850 [6]. Consequently, key challenges faced by cybersecurity community and industrial practitioners are low reproducibility, comparability and peer validated research. These factors motivate us to develop a benchmark dataset for electrical substation cybersecurity study. Though there are intrusion datasets for generic computer network system [7] and smart grid [8], our benchmark dataset will facilitate study on cyber attacks specifically on substations. Opportunities that we leverage to conduct this work are enhanced cross-domain communication, systematization of industrial control system attack and standard communication protocol IEC 61850 that deals with heterogeneity of IEDs from different vendors. By specifying the workflow for creating a dataset, we aim at providing a realistic benchmark synthesized dataset for evaluation of cybersecurity solutions such as the IDS and false data detector.

II. Related Work

To enable the cybersecurity study of industrial networks, several testbeds have been developed to capture network traffic for experimentation and testing purposes. These testbeds include the "Geek Lounge Lab" deployed at 4SICS conference [9] and the Electric Power Intelligent Control (EPIC) testbed in Singapore University of Technology and Design [10]. Unfortunately, the 4SICS testbed does not support IEC 61850-based traffic, which is widely used in smart grid systems. Although the EPIC dataset [11] captures Manufacturing Message Specification (MMS) messages, it does not contain GOOSE messaging which is the focus in this paper as GOOSE communication is prevalent for automated protection and control in modernized substations. Furthermore, the dataset does not contain attack traces. Other popularly used datasets for evaluating network-based IDS include KDD-Cup99 [12], NSL-KDD [13], UNSW-NB15 [7], and CICIDS2017 [14]. However, these datasets are better suited to modeling intrusions in traditional computer networks. They

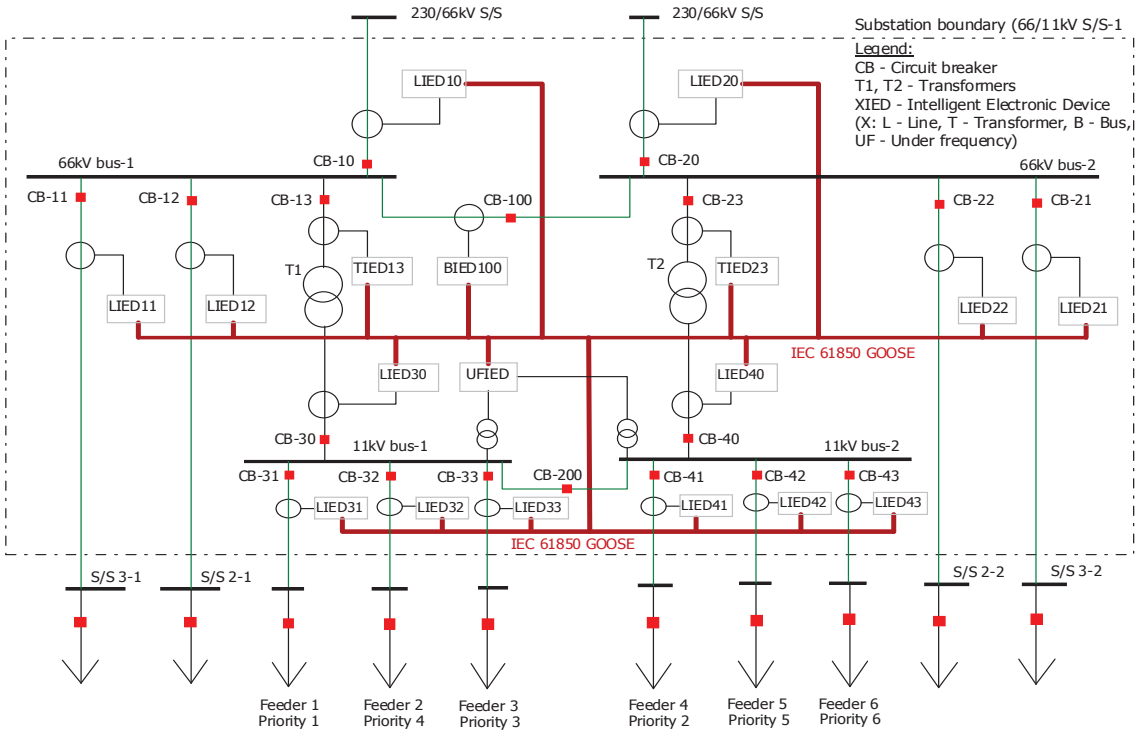

Fig. 1. Substation one-line diagram. The prefixes to IEDs denote - 'L' for line, 'T' for transformer and 'B' for bus.

lack traffic diversity that characterizes substation communication such as the network traffic based on IEC 61850 standard.

Some efforts have also been made to generate IEC 61850 Generic Object Oriented Substation Events (GOOSE) and Sampled Value (SV) traffic [15], [16], [17]. Hegazi et al. [15] described the modeling and generation of GOOSE communication traffic. However, since security application was out of scope, the attacker model was not considered in their modeling. Lopes et al. [16] proposed a GOOSE traffic generator for evaluation and testing of security mechanisms. While their approach is similar to ours, the authors did not discuss realistic substation models or threat models. Thus, the practical use of the tool requires further design and understanding about the substation systems. In the same vein, Blair et al. [17] proposed an open source platform for rapid prototyping of GOOSE and SV traffic without discussing the threat models. In contrast, we provide a framework to easily generate attack-free and attack traces for benchmarking security mechanisms developed by the research community.

III. Substation Model

In this section, we discuss a typical substation model and its operation under some selected attack-free, but disturbance scenarios which may occur in the power system due to faults.

A. Description of Physical Model

Fig. 1 shows one-line diagram of a 66/11kV substation. A transformer in each 66kV bus of the redundantly designed system steps down the voltage to 11kV level. The line feeders connect this substation to other nearby substations to provide a redundant configuration. 11kV voltage level is supposedly

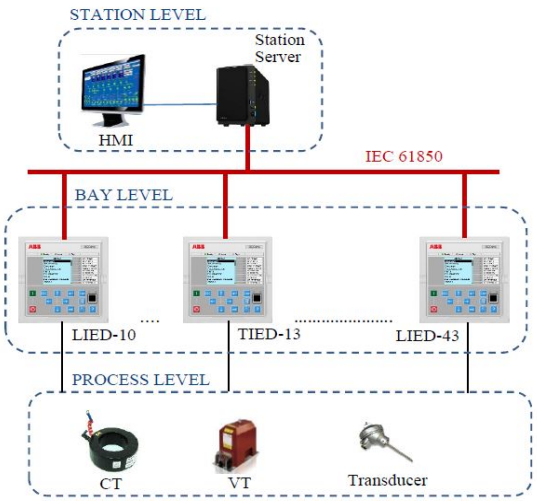

Fig. 2. Simplified network architecture

for distribution level. Underfrequency load shedding relay is connected at this voltage level. Bus-coupler breakers (CB-100 and CB-200) are normally in the open state under normal operating scenario. CB-200 goes into the closed state (closed by electrical interlocking) whenever there is no fault at 11kV bus and either of CB-30 or CB-40 is open.

B. Network Architecture

Fig. 2 shows the simplified network architecture for the substation in discussion. At process level, the interfacing equipment like current transformers (CT), voltage transformers (VT), transducers, circuit breakers, etc. are present. At

bay level, all the Intelligent Electronic Devices (IEDs) are interconnected via Ethernet LAN cables and communicate among themselves using IEC 61850 GOOSE protocol. It is worthwhile to note that IEC 61850 standard [6] defines communication protocols for IEDs in substations. A number of protocols (such as GOOSE, MMS and SV) can be mapped to the abstract data models defined in this standard. In IEC 61850 GOOSE communication, data (status, alarms, measurements) of any format is grouped into a data set and transmitted over the electrical substation network in a fast and reliable manner. The connection between the IEDs and Human Machine Interface (HMI) at station level is usually established by LAN via the station server. IEC 61850 utilizes MMS, an international standard (ISO 9506) dealing with messaging systems for transferring real time process data and supervisory control information between network devices (such as IEDs) and station server.

C. Attack-Free Scenarios in a Substation

Besides the communication under the normal operation, we consider 3 representative disturbance scenarios under which substation protection system operates. These are not attacks, but abnormal operation due to faults in the power system. Under normal operation, these scenarios do not exist. While describing the benchmark datasets (presented in Section VI), we present the attack-free normal operation dataset and attack-free selected abnormal scenario datasets separately.

1. Busbar protection
- Let us consider that a fault occurs at busbar 66kV bus-1. The incomer line LIED10 will pick up on overcurrent, the other IEDs wont.
- The incomer line LIED10 will know through GOOSE communication that the overcurrent elements of other IEDs have not picked up.
- The incomer line LIED10 will quickly realise of busbar fault and trigger a trip to its own breaker CB-10 first and subsequently to the breakers associated with the busbars i.e., CB-11, CB-12 and CB-13.
- The trip status of CB-10 is sent by LIED10 through GOOSE communication to LIED11, LIED12 and TIED13 to trigger trip for their respective breakers.

2. Breaker failure protection:
- Assume that a fault occurs in the feeder connecting substation S/S 3-1. The associated LIED11 overcurrent (O/C) element picks up, however, the breaker CB-11 does not trip due to mechanical failure.
- The GOOSE communication of breaker failure and O/C element pick-up is sent from LIED11 to the LIED10 (incomer), LIED12 and TIED13.
- The communication triggers tripping of circuit breakers CB-10, CB-12 and CB-13. Subsequently, the remote CB (in S/S 3-1) is tripped using proper communication media.

3. Underfrequency load-shedding:
- Under frequency IED (UFIED) can sense the under frequency in both 11kV buses when voltage transformer (VT) inputs are directly given to this IED.
- Alternately, incomer line IEDs (LIED30 and LIED40) can relay the information to the under frequency IED over GOOSE (when there is no direct VT input to UFIED).
- UFIED triggers a trip over GOOSE to the least priority (Priority 6) consumer first via LIED43. Then, CB-43 is tripped at first stage.
- After a time delay (usually 2-4 seconds) if the frequency is not stabilized at the desired value, further loads are shed. The sequence of tripping goes as: Priority 6 (CB-43) $\rightarrow$ Priority 5 (CB-42) $\rightarrow$ Priority 4 (CB-32) ...
- Between each two stages of tripping there is a time delay of 2-4 seconds. The trip is initiated via GOOSE communication to respective IEDs such as LIED43, LIED42, LIED32, etc.

IV. Cyber Attack Scenarios

In this section, we define the threat model, the GOOSE communication model and discuss some attacks encountered by the substation.

A. Threat Model and Assumptions

The engineering and operator workstations are HMIs deployed in a substation to provide users with a graphical user interface for monitoring and controlling devices. While access to the operator workstations is restricted to authorized personnel locally, the engineering workstations may be equipped with remote access capabilities that allow access from locations outside the substation network, i.e., corporate offices and control centers. As a result, they become soft targets and easy entry points for attackers to infiltrate into the substation. External attackers can use social engineering techniques or phishing emails to compromise the engineering workstation and gain a foothold in the substation network, just like the Ukraine case [18]. Once inside the network, the attackers can gather knowledge about the substation's topology and the IEDs' operations, including the IEDs' login credentials. With this information, the attackers can launch attacks on any IED of their choice, assuming that both the engineering workstation and the IEDs are connected to the same LAN and that no VLAN membership is configured on the switch.

B. Attack Model Against GOOSE Communication

In mentioning about communication models, of particular interest is the GOOSE protocol as it is the key driver for automated control and protection. Under the threat model discussed in the previous subsection, an attacker can compromise the GOOSE communication to negatively impact the protection scheme described in Section III-C. In fact, the GOOSE messaging is not encrypted for performance reasons. Thus, attackers can eavesdrop, analyze, and spoof GOOSE frames [19]. By spoofing, we mean that an attacker can masquerade as a legitimate IED to inject GOOSE frames. In doing so, the attacker can send malicious GOOSE frames to various IEDs to cause damage, including modifying the response messages sent from the IEDs to mislead the operators about the actual state of the substation. The following is a

TABLE I
COMMON CYBER ATTACKS

Attacks	Tactics
DoS.1	• Flood bogus frames
MS.1	• Inject GOOSE sequence number (SqNum)
MS.2	• Inject GOOSE status number (StNum)
DM.1	• Modify current measurements reported by the merging units
DM.2	• Inject modified Boolean value of circuit breaker
DM.3	• Replay a previously valid message
CM.1	• Inject modified Boolean value of circuit breaker (from HMI)

consolidated list of GOOSE-related attacks that may compromise the substation operations. Table I enumerates the different tactics that make these attacks possible.

- Denial of Service **(DoS)** - Block the flow of information to the intended IEDs by flooding etc. to lower the availability of service.
- Message Suppression **(MS)** - Prevent legitimate IEDs from receiving critical messages or updates by modifying the GOOSE header fields to hijack the communication channel.
- Data Manipulation **(DM)** - Spoof false information to the HMI or IEDs to mask unauthorized changes.
- Control Manipulation **(CM)** - Modify the payload to control field devices such as circuit breakers.

For MS.1 attack in Table I, the sqNum is a counter that increments each time a GOOSE frame is sent. An attacker can modify this value to cause the subsequent GOOSE frames to arrive out of sequence. As for MS.2, the stNum is a counter that increments each time a value change has been detected within the dataset. An attacker can inject a GOOSE frame with a stNum higher than the one previously captured. Once the modified GOOSE frame is processed by the subscribers, subsequent legitimate GOOSE frames, with status number equal to or less than the modified stNum will be dropped [20].

V. DATASET GENERATION FRAMEWORK

In this section, we discuss our framework for generating GOOSE dataset under attack-free and attack-induced scenarios. The proposed framework consists of two parts: an Attack-Free Trace Generator and an Attack-Induced Trace Generator. The former generates GOOSE traces that represent substation communication without attack, while the latter can inject multiple attacks at different locations in the attack-free trace using strategies discussed in Section IV to generate attack-induced trace. In the following, we describe the various components of the framework in detail and discuss how it can be used to generate the required trace data.

A. Generation of Attack-Free Traces

The top part of Fig. 3 shows the workflow for generating Attack-Free GOOSE traces. The first step is to parse the SCL (substation configuration language) file, which is part of IEC 61850 standard, through the SCL conversion tool and extract meaningful IEDs models for use in the Attack-Free Trace Generator. The SCL files are required because they contain data structures and values that define the IED data models.

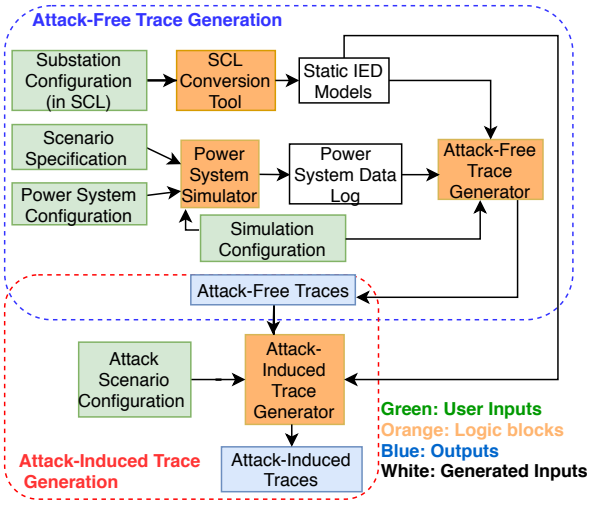

Fig. 3. Framework for generating attack-free and attack dataset

Besides the static IED models, the Attack-Free Trace Generator accepts a power system data log and a simulation configuration as inputs. The power system data log contains individual IED's nominal current and voltage measurements in time series order. These measurements can be generated by power system simulators (e.g. PowerWorld [21] or DIgSILENT PowerFactory [22]) whose inputs depend on the scenario configuration, the power system configuration, and the simulation configuration.

The scenario specification contains high-level descriptions for modeling not only attack-free scenarios but also disturbances whenever such a condition arises. For example, we can define a normal scenario between 0 and 10 sec, line fault at 11^{th} sec, line trip at 12^{th} sec, and so forth. The power system configuration defines the number of buses, generators, loads and lines for use in the power system simulator to generate the power system data log. The simulation configuration contains user inputs for specifying parameters, such as the simulation duration, the GOOSE sending interval and so on. Given these inputs, the Attack-Free Trace Generator will generate the relevant GOOSE traces as output.

B. Generation of Attack-Induced Traces

The bottom part of Fig. 3 illustrates the Attack-Induced Trace Generation workflow for creating GOOSE attack traces to simulate attacks on the substation. In addition to the static IED models, the Attack-Induced Trace Generator seeks two more inputs - a network trace and an attack scenario configuration. The input network trace can be an attack-free trace generated by the aforementioned Attack-Free Trace Generator, or a network trace captured in a real system, if available.

In order for the Attack-Induced Trace Generator to edit the traffic in the input network trace, a traffic replay tool (e.g., tcpreplay [23]) which is part of the Attack-Induced Trace Generator program is utilized to reproduce the traffic in the input network trace as a baseline. Then by specifying the attack type (as covered in Table I), the IED GOOSE identifier,

the time of attack, and the value to be modified in the attack scenario configuration, the Attack-Induced Trace generator can inject specific attack signatures into the input network trace to generate an attack-induced trace.

VI. An IEC-61850 Substation Dataset

Using the framework, we create a dataset to mimic three different scenarios: attack-free normal operation, attack-free with disturbances and attack trace with different attacking scenarios. The research community can download this dataset from our Github repository [24] and use them to validate their cybersecurity solutions, such as IDS, false data detectors, and more. In the following, we describe how each of these scenarios is generated.

Simulation Setup: We emulate 18 IEDs according to the one-line diagram in Fig. 1 and each IED is assigned a unique MAC address to simulate GOOSE communications. We assume that all 18 IEDs are located in the same multicast group, i.e., an IED can receive multicast frames sent by any IED in the group using 01-0C-CD-01-00-01 as the destination MAC address. To simplify the simulation, we manually create a power system data log for each IED in CSV format to describe the operating current, voltage, power and frequency measurements under normal and disturbance scenarios.

Attack-free Normal Scenario: We configure all IEDs to run normally from $t = 0\,sec$ to $t = 600\,sec$ in the scenario specification. Accordingly, we set the Attack-free Trace Generator program to run for $600\,sec$ in the simulation configuration and set the sending rate to 1 GOOSE frame/second. As can be seen from Fig. 4, whenever a GOOSE frame is sent, the sqNum is incremented by 1, while the stNum and timestamp values remain unchanged since there is no event change in the GOOSE dataset i.e., no disturbance or attack.

Attack-free with Disturbance Scenario: We use the breaker failure scenario outlined in Section III-C as an example to generate a trace file with disturbance. In the scenario specification, we simulate a line fault at $t = 11\,sec$ on the feeder connected to the substation (S/S 3-1) to invoke the breaker failure protection scheme. In other words, LIED11 multicasts a GOOSE frame to inform the neighboring IEDs of a surge in current and the malfunctioning of CB-11. Fig. 5 shows part of the resulting trace file where the StNum increases from 1 to 2 and the sqNum is reset to 0 when an overcurrent value of 6000 amps is detected. Concurrently, the timestamp t is updated and the attribute associated with the mechanical fault condition of the CB-11 is changed from FALSE to TRUE.

Attack Scenario: We use the attack-free trace generated under normal condition as a baseline in our Attack-Induced Trace Generator program to generate the attack-induced trace. We assume an attack scenario where an attacker combines multiple tactics to attack CB-11. The attacker hijacks the communication channel by injecting a GOOSE frame with high status number (MS.2) at around $t = 11\,sec$ and later at $t = 16\,sec$, trips the circuit breaker via LIED11 by injecting another GOOSE frame containing modified Boolean value

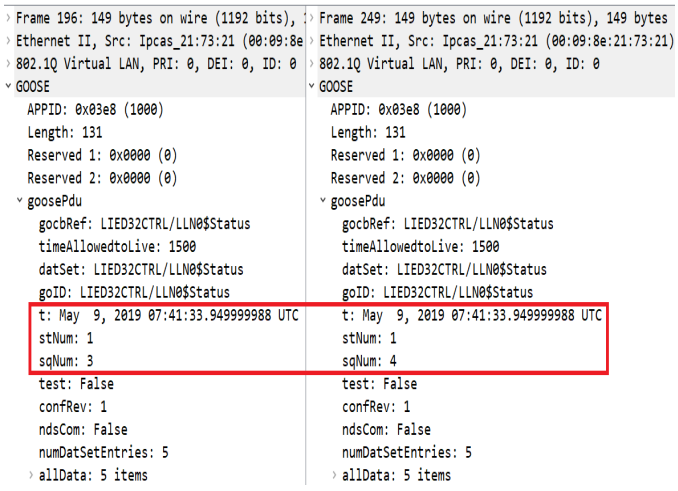

Fig. 4. Screenshot of attack-free network trace showing a change in sqNum while the stNum and timestamp t values remain unchanged under normal scenario

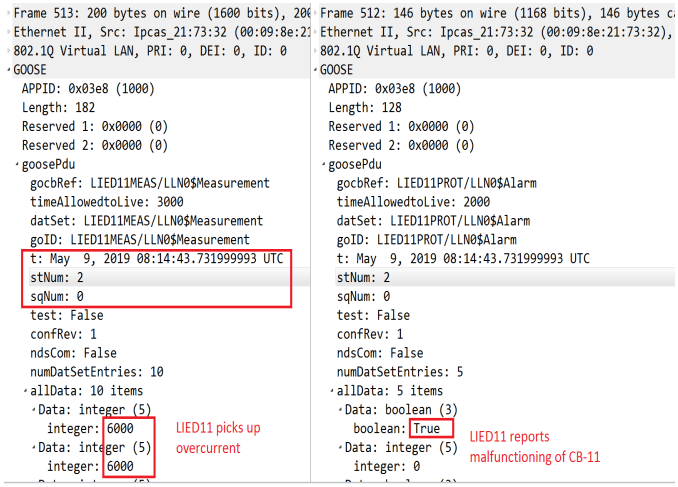

Fig. 5. Screenshot of attack-free but with disturbance network trace showing changes in stNum, sqNum and timestamp t

(DM.2). Fig. 6 and Fig. 7 are screenshots of the attack-induced trace, showing multiple attacks at different timings in the original normal attack-free trace. As shown in Fig. 6, a GOOSE frame containing a high status number of 9999 is successfully injected between $t = 11\,sec$ and $t = 12\,sec$ to simulate the communication channel hijacking. Fig. 7 shows that at around $t = 16\,sec$, a malicious GOOSE frame is injected to open CB-11, causing other IEDs to trip their CBs as a result. It may be noted that though only attack types MS.2 and DM.2 are discussed here, the other common cyber attacks (in Table I) can also be realized using our framework. Indeed, our GitHub repository includes attack traces for most types of attacks mentioned in Table I.

VII. Use of the Dataset for Security Study

In this section, we discuss the use of our dataset to support the design and evaluation of cybersecurity solutions

TABLE II

DESCRIPTION FOR EACH NETWORK TRACE IN OUR DATASET INCLUDING PRELIMINARY RESULTS OF SOME IDSES RESULTS AGAINST EACH NETWORK TRACE. – MEANS CANNOT DETECT; ✓ MEANS CAN DETECT

Network Trace	Description	[25]	[26], [27], [28]	[29]
MS.2/AS1.pcapng	Inject a high stNum value or slightly higher than previously recorded stNum and sqNum ≠ 0	–	✓	✓
MS.2/AS2.pcapng	Replay a previously valid frame containing high stNum, sqNum = 0 but stale timestamp	–	–	✓
MS.2/AS3.pcapng	More advanced attack; inject a high stNum frame, sqNum = 0 with a valid timestamp	–	–	–

IDS Detection rules:
[25] Traffic analysis/statistics, [26], [27], [28] Check stNum is increased and sqNum=0
[29] Check stNum is increased and sqNum=0 and timestamp must be coherent with the previous GOOSE frame

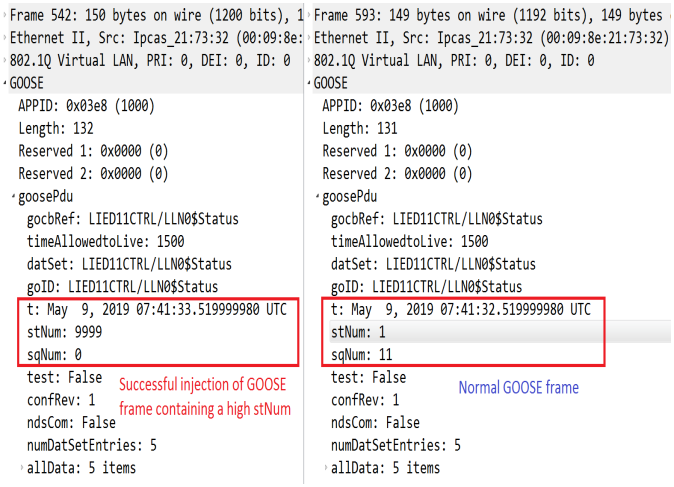

Fig. 6. Screenshot of attack-induced trace showing successful injection of GOOSE frame with a high status number

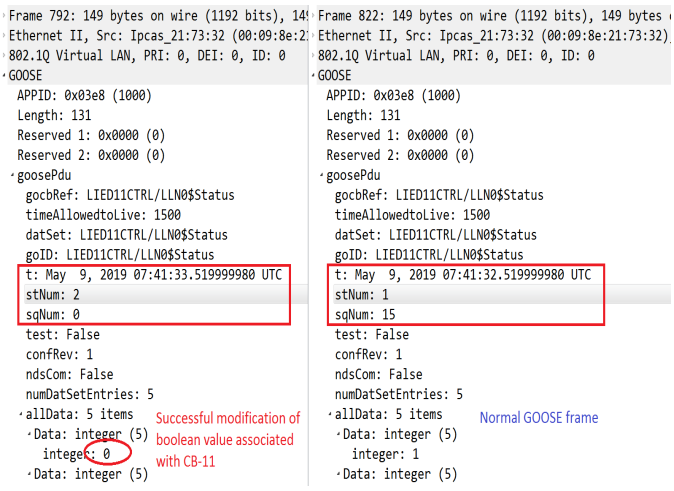

Fig. 7. Screenshot of attack-induced trace showing successful injection of GOOSE frame with modified Boolean values associated with the circuit breaker CB-11

for substations. Due to page limitation, we report only some preliminary findings here. For example, several papers have discussed IDS for handling stNum attacks. In particular, Ren et al. [25] proposed an Edge-Based Multi-Level Anomaly Detection (EDMAND) for detecting network anomalies associated with Modbus and DNP3 traffic. Their IDS is based on collecting traffic statistics (e.g., packet interarrival time, packet size) at the transport layer, operation layer, and control layer to do detection. We note that even if we adapt EDMAND to parse GOOSE traffic, the idea of using high-level traffic statistics may not be sufficient to detect stNum attacks, since only one or a small number of injected packets can cause the harm. Yang et al. [26] proposed a specification-based IDS to check if the sqNum is 0 when the stNum is incremented. Similar IDS rule is also used in [27], [28]. However, if the attacker knows how the GOOSE messaging protocol works, (s)he can bypass this simple IDS rule by resetting the sqNum after modifying the stNum value. Kabir-Querrec et al. [29] proposed a more elaborate IDS design by comparing the stNum, sqNum including timestamps from two consecutive messages. However, this approach can only detect stNum modification by replay, since the timestamp will be inconsistent with respect to the previously received GOOSE frames. If the attacker modifies the timestamp to be consistent with stNum modification, this approach will fail. To this end, we have created two separate network traces: one to capture the replay of a previously valid high stNum (refer to MS.2/AS2.pcapng) and another trace to inject a high stNum frame with a valid timestamp (refer to MS.2/AS3.pcapng). Table II summarizes the various scenarios of stNum modification attacks, their respective network traces and the capabilities of some IDSes we discussed above.

It is also worth mentioning that the stNum change may not be related to attacks, but due to some disturbances on the bus or feeder, as described in Section III-C. In this case, IDSes [25], [26], [27], [28], [29] that rely on monitoring the network traffic for suspicious activities without the knowledge of the physical behavior of a power grid may not be able to distinguish real attacks from disturbances. Thus, these IDSes may suffer from false negatives and/or false positives. Based on our findings, we conclude that in addition to checking whether the stNum has increased, the sqNum is 0 and the timestamp is valid, it could be necessary to detect attacks by looking beyond a single packet (e.g., considering the repeated appearance of inconsistent stNum from different packets as an attack indicator) and use the knowledge of physical system.

Simple anomaly detection rules that detect bad data injection (e.g., by looking at the change of power system measurement data) will cause false alarm in some attack-free disturbance traces we generated. As an example, if we consider the scenario of busbar protection or breaker failure discussed in Section III-C, the fault current realized by the relevant IEDs will be ten times or more than the nominal current depending on the system impedance. The situation will be wrongly flagged as an attack by a simple range-based anomaly detector

(e.g., the one described in [26]). A more advanced IDS which can validate the reading by cross-checking among IEDs will provide a better detection accuracy. Furthermore, variation in load is a common phenomenon in power system. Network loading can be anywhere between 0% to 100%. An IDS employing statistical approach or sometimes even knowledge-based approach may trigger false alarms [30] under different loading conditions. A comprehensive dataset incorporating multiple load levels and attack-free disturbance scenarios (e.g., as those in our BusbarProtection.pcapng, BreakFailure.pcapng, VariableLoad.pcapng, etc. pcap traces) would help evaluate the performance and threshold for the IDS.

VIII. Conclusion

In this paper, we discuss the design of IEC 61850 GOOSE network traffic traces that can be used to benchmark smart grid security solutions, such as intrusion detection systems. We elaborate a realistic substation system topology and operation models used as the basis for traffic generation. We further discuss attack models against GOOSE-based communication, which are used for injecting attacks into the trace. The generated traces are published online and in the future, we plan to test the generated attack traces on state-of-the-art IDS solutions available in the market as well as those developed in academia. Furthermore, we also intend to publish the toolchain we have implemented for generation of such traces, so that researchers can utilize it to generate attack-free and/or attack-induced traces that are of their interest.

Acknowledgement

This research is supported in part by the National Research Foundation, Prime Minister's Office, Singapore under the Energy Programme and administrated by the Energy Market Authority (EP Award No. NRF2017EWT-EP003-047), and in part by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

References

[1] R. Heady, G. Luger, A. Maccabe, and M. Servilla, "The architecture of a network level intrusion detection system," Los Alamos National Lab., NM (United States); New Mexico Univ., Albuquerque , Tech. Rep., 1990.

[2] A. Gupta, A. Anpalagan, G. H. Carvalho, L. Guan, and I. Woungang, "Prevailing and emerging cyber threats and security practices in iot-enabled smart grids: A survey," *Journal of Network and Computer Applications*, vol. 132, pp. 118–148, 2019.

[3] H. He and J. Yan, "Cyber-physical attacks and defences in the smart grid: a survey," *IET Cyber-Physical Systems: Theory & Applications*, vol. 1, no. 1, pp. 13–27, 2016.

[4] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Computer Networks*, vol. 57, no. 5, pp. 1344–1371, 2013.

[5] H. C. Tan, C. Cheh, B. Chen, and D. Mashima, "Tabulating cybersecurity solutions for electrical substations towards pragmatic design and planning," in *presented at the IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), Chengdu, China*. IEEE, 2019.

[6] "IEC 61850 - communication networks and systems in substations," May 2019. [Online]. Available: https://webstore.iec.ch/

[7] N. Moustafa and J. Slay, "Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set)," in *2015 military communications and information systems conference (MilCIS)*. IEEE, 2015, pp. 1–6.

[8] V. Babu, R. Kumar, H. H. Nguyen, D. M. Nicol, K. Palani, and E. Reed, "Melody: synthesized datasets for evaluating intrusion detection systems for the smart grid," in *2017 Winter Simulation Conference (WSC)*. IEEE, 2017, pp. 1061–1072.

[9] "Capture files from 4sics geek lounge," April 2019. [Online]. Available: https://www.netresec.com/?page=PCAP4SICS

[10] S. Adepu, N. K. Kandasamy, and A. Mathur, "Epic: An electric power testbed for research and training in cyber physical systems security," in *Computer Security*. Springer, 2018, pp. 37–52.

[11] "EPIC dataset," https://itrust.sutd.edu.sg/itrust-labs_datasets/dataset_info/epic, April 2019.

[12] "Kdd cup 1999 data," May 2019. [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[13] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*. IEEE, 2009, pp. 1–6.

[14] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization." in *ICISSP*, 2018, pp. 108–116.

[15] O. Hegazi, E. Hammad, A. Farraj, and D. Kundur, "IEC-61850 GOOSE traffic modeling and generation," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2017, pp. 1100–1104.

[16] Y. Lopes, D. C. Muchaluat-Saade, N. C. Fernandes, and M. Z. Fortes, "Geese: A traffic generator for performance and security evaluation of IEC 61850 networks," in *2015 IEEE 24th International Symposium on Industrial Electronics (ISIE)*. IEEE, 2015, pp. 687–692.

[17] S. M. Blair, F. Coffele, C. D. Booth, and G. M. Burt, "An open platform for rapid-prototyping protection and control schemes with IEC 61850," *IEEE Transactions on Power Delivery*, vol. 28, no. 2, pp. 1103–1110, 2013.

[18] Defence Use Case, "Analysis of the cyber attack on the ukrainian power grid," 2016.

[19] M. T. A. Rashid, S. Yussof, Y. Yusoff, and R. Ismail, "A review of security attacks on IEC61850 substation automation system network," in *Proceedings of the 6th International Conference on Information Technology and Multimedia*. IEEE, 2014, pp. 5–10.

[20] J. Hoyos, M. Dehus, and T. X. Brown, "Exploiting the GOOSE protocol: A practical attack on cyber-infrastructure," in *2012 IEEE Globecom Workshops*. IEEE, 2012, pp. 1508–1513.

[21] "PowerWorld Simulator Overview," April 2019. [Online]. Available: https://www.powerworld.com/products/simulator/overview

[22] "Power System Software Engineering," April 2019. [Online]. Available: https://www.digsilent.de/en/

[23] "Tcpreplay - Pcap editing and replaying utilities," April 2019. [Online]. Available: https://tcpreplay.appneta.com/

[24] "GOOSE dataset," May 2019. [Online]. Available: https://github.com/smartgridadsc/IEC61850SecurityDataset

[25] W. Ren, T. Yardley, and K. Nahrstedt, "Edmand: Edge-based multi-level anomaly detection for scada networks," in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, 2018, pp. 1–7.

[26] Y. Yang, H.-Q. Xu, L. Gao, Y.-B. Yuan, K. McLaughlin, and S. Sezer, "Multidimensional intrusion detection system for IEC 61850-based scada networks," *IEEE Transactions on Power Delivery*, vol. 32, no. 2, pp. 1068–1078, 2017.

[27] J. Hong, C.-C. Liu, and M. Govindarasu, "Detection of cyber intrusions using network-based multicast messages for substation automation," in *ISGT 2014*. IEEE, 2014, pp. 1–5.

[28] ——, "Integrated anomaly detection for cyber security of the substations," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1643–1653, 2014.

[29] M. Kabir-Querrec, S. Mocanu, P. Bellemain, J.-M. Thiriet, and E. Savary, "Corrupted GOOSE detectors: Anomaly detection in power utility real-time ethernet communications," in *GreHack 2015*, 2015.

[30] E. Biermann, E. Cloete, and L. M. Venter, "A comparison of intrusion detection systems," *Computers & Security*, vol. 20, no. 8, pp. 676–683, 2001.