energies

Article

GOOSE Secure: A Comprehensive Dataset for In-Depth Analysis of GOOSE Spoofing Attacks in Digital Substations

Oscar A. Tobar-Rosero [1,*,†], Omar A. Roa-Romero [1], Germán D. Rueda-Carvajal [1], Alexánder Leal-Piedrahita [2], Juan F. Botero-Vega [2], Sergio A. Gutierrez-Betancur [2], John W. Branch-Bedoya [1] and Germán D. Zapata-Madrigal [3]

[1] TyT Group, Universidad Nacional de Colombia, Medellín 050034, Colombia; oaroaro@unal.edu.co (O.A.R.-R.); gdruedac@unal.edu.co (G.D.R.-C.); jwbranch@unal.edu.co (J.W.B.-B.)
[2] GITALab, Universidad de Antioquia, Medellín 050010, Colombia; erwin.leal@udea.edu.co (A.L.-P.); juanf.botero@udea.edu.co (J.F.B.-V.); sergio.gutierrezb@udea.edu.co (S.A.G.-B.)
[3] GIDIA Group, Universidad Nacional de Colombia, Medellín 050034, Colombia; gdzapata@unal.edu.co
* Correspondence: oatobarr@unal.edu.co; Tel.: +57-317-750-6055
† Current address: Industrial Automation and Communications Laboratory, Universidad Nacional de Colombia, Sede Medellín, Medellín 050034, Colombia.

Abstract: Cybersecurity in Critical Infrastructures, especially Digital Substations, has garnered significant attention from both the industrial and academic sectors. A commonly adopted approach to support research in this area involves the use of datasets, which consist of network traffic samples gathered during the operation of an infrastructure. However, creating such datasets from real-world electrical systems presents some challenges: (i) These datasets are often generated under controlled or idealized conditions, potentially overlooking the complexities of real-world operations within a digital substation; (ii) the captured data frequently contain sensitive information, making it difficult to share openly within the research community. This paper presents the creation of a new dataset aimed at advancing cybersecurity research, specifically focused on GOOSE spoofing attacks, given the crucial role of the GOOSE protocol in managing operational and control tasks within Digital Substations. The dataset highlights the real-world impacts of these attacks, demonstrating the execution of unintended operations under different operational scenarios, including both stable conditions and situations involving system failures. The data were collected from a laboratory testbed that replicates the actual functioning of a real digital substation with two bays. The experiments provided insights into key characteristics of GOOSE protocol traffic and the vulnerability of DS infrastructure to Spoofing Attacks.

Keywords: cybersecurity; dataset; digital substation; GOOSE protocol; IEC 61850; smart substation; spoofing attack

check for updates

Citation: Tobar-Rosero, O.A.; Roa-Romero, O.A.; Rueda-Carvajal, G.D.; Leal-Piedrahita, A.; Botero-Vega, J.F.; Gutierrez-Betancur, S.A.; Branch-Bedoya, J.W.; Zapata-Madrigal, G.D. GOOSE Secure: A Comprehensive Dataset for In-Depth Analysis of GOOSE Spoofing Attacks in Digital Substations. *Energies* **2024**, *17*, 6098. https://doi.org/10.3390/en17236098

Academic Editor: Michael Negnevitsky

Received: 11 September 2024
Revised: 7 October 2024
Accepted: 23 October 2024
Published: 4 December 2024

1. Introduction

Digital Substations (DS) or Smart Substations are fundamental in adopting novel electrical systems monitoring, control, and supervision mechanisms. Various protocols and standards are employed within a DS environment to implement appropriate communication among multiple devices and technological tools [1]. However, although the process of substation digitalization brings benefits, it also introduces cyber security vulnerabilities that must be addressed to guarantee secure, reliable, and resilient systems [2,3]. An example of these vulnerabilities is the fact that by manipulating information in the Supervisory Control and Data Acquisition (SCADA) communication channels, attackers could inject false commands or tamper with communications among critical devices [4].

One of the communication protocols of high interest in the operation of DSs is GOOSE (Generic Object Oriented Substation Event). This protocol is used to transmit messages for protection and control operations and monitor the behavior of elements within DSs [5].

Considering the relevance of GOOSE in the operation of DSs, it turns out to be a relevant target for several types of cyberattacks. Several studies in the literature have demonstrated the devastating impacts and effects on Smart Grids infrastructures caused by cyberattacks such as False Data Injection (FDI), Denial of service (DoS), Distributed Denial of Service (DDoS), and Malware, among others [6,7]. Among these attacks, Spoofing stands out due to its potential impact [8–10]. This type of attack involves resembling legitimate devices by injecting malicious traffic to disrupt the normal operation of a DS possibly causing failures in the overall electrical system [11].

A valid approach present in the literature for the study of cyberattacks, threats and vulnerabilities is the analysis of datasets. In computer networks and cybersecurity research, a dataset typically consists of structured data that can be used to identify and characterize the flow of data and information exchanges among devices in a system [12,13]. In the context of DSs, the analysis of these data flows and information exchanges among devices enable recognizing behavioral patterns, identifying devices and messages, and determining other relevant attributes that can be leveraged to establish normal operation baselines, and therefore, detecting anomalies [3].

A desirable property expected in datasets used in the context of DSs is that they contain messages actually expressing the typical traffic patterns associated with actual operational events of DSs infrastructures [14]. However, to the best of our knowledge, datasets built for DS infrastructures are mostly derived from synthetic traffic generated through simulation tools [15,16]. For security reasons, companies are very secretive about the operational information of their electrical systems and, therefore, of the digital substations. Hence, the access to DSs is limited, and hampers to obtain real operation data. In general, these datasets are created under controlled or ideal conditions, thus ignoring the dynamics of real-world operations that might appear in actual DSs.

This paper presents the construction of a dataset for in-depth analysis of the GOOSE protocol in DSs in different conditions. The dataset has been obtained from a laboratory testbed deployed on a physical infrastructure that reproduces the operation of a DS with two bays in an electrical system. This physical infrastructure encompasses typical communication protocols of a DS and allows the reproduction of the operational dynamics of DSs. Hence, the dataset includes traffic samples associated with operations performed in a real system in a company. To complement the analysis, the testbed infrastructure was also attacked with a Spoofing attack, which evidently caused disruptions and, therefore, failures on the DS. Thus, this dataset contains traffic samples presenting the behavior of the DS infrastructure when facing actual attacks.

The dataset generated does not use data from real clients or customers, given the operational and ethical limitations established by the different companies in the energy sector. Under this premise, the information disclosed in our dataset does not have any information that requires the implementation of access or security measures for its use. This is especially important considering this dataset is intended to be public, in such a way that the scientific community can use it for further research.

The remainder of this paper is organized as follows. Section 2 provides a theoretical background related to DSs and the type of attack of interest considered in the dataset. Next, Section 3 presents previous work related to the proposed topic. It also establishes differences in the datasets generated, highlighting the need and importance of the dataset presented here. Section 4 presents a detailed description of test scenarios as well as the methodology employed to capture data for each of them. In Section 5, we analyze the dataset and explain the detailed behavior of GOOSE messages. Finally, Section 6 concludes this paper by condensing the results and presenting the section for future work.

2. Background

This section provides a brief introduction to several key concepts considered in this paper. First, we introduce a basic definition of the concept of Digital Substation (DS). Next, we present an overview of the infrastructure elements that are typically found in the architecture of DSs, with particular emphasis on the communication protocols used among these elements. From these protocols, GOOSE is presented with further detail in order to understand why the Spoofing attacks, finally discussed, are especially critical against this protocol.

2.1. Digital Substation in a Nutshell

The digitalization process of electrical substations is guided by the IEC 61850 standard [17], which proposes an architecture such as the one illustrated in Figure 1 [18]. This standard comprises different concepts including systems integration, interoperability, and specifications of communication technologies and protocols, aiming at simplifying supervision, control and monitoring activities in DSs [1,19]. According to the IEC 61850 standard, a DS includes three operation levels known as (i) Level 0: Process; (ii) Level 1: Bay; (iii) Level 2: Station. These operation levels are linked through two communication buses named station bus and process bus, which are used for information exchange among the devices of each level. For this information exchange, the DS leverages a set of communication protocols [20]. In addition to the mentioned levels, there is Level 3 which is considered external to the DS. Level 3 comprises the control center of the utility company operating the DS, where tasks associated with remote control and supervision are carried out. There are additional standards and protocols for Layel 3 which are out of the scope of IEC 61850 such as IEC 60870 [21] or IEEE 1815 [22]. The protocols associated with these standards are not considered in this work.

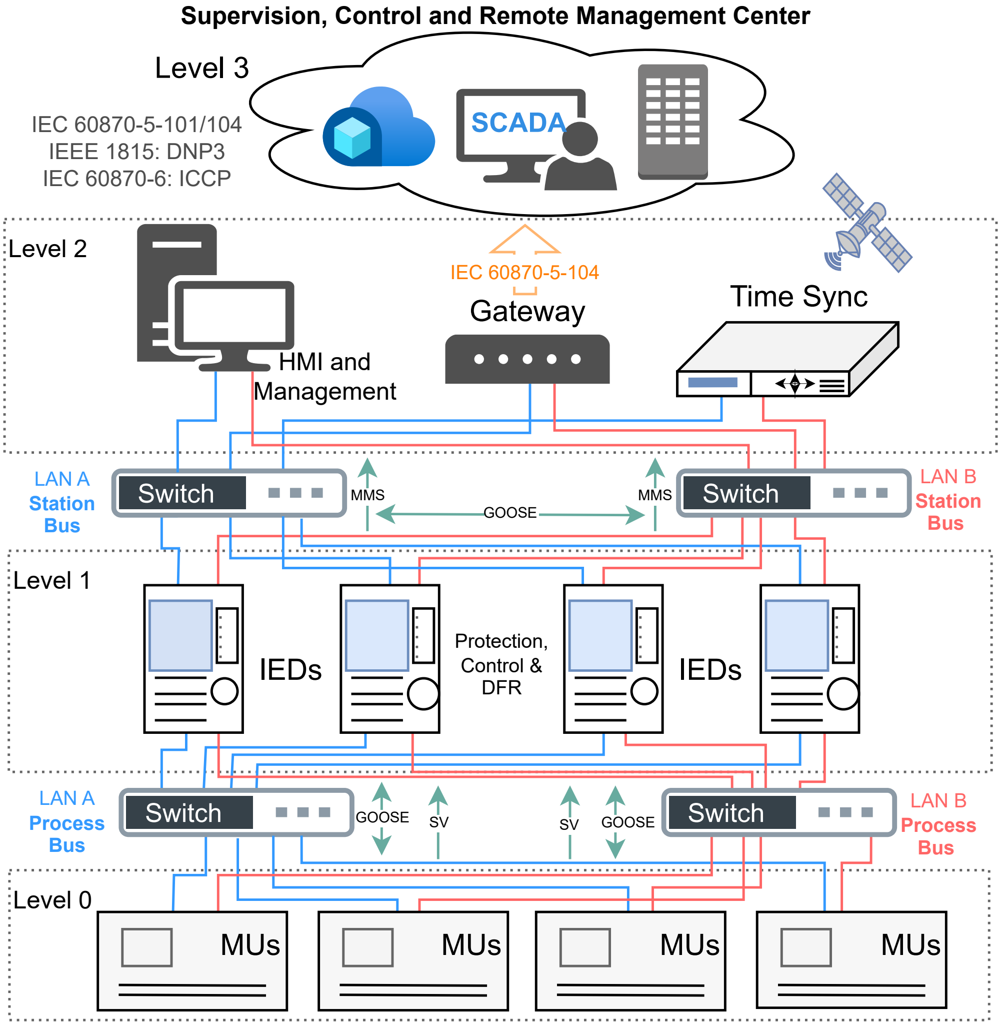


Figure 1. Digital Substation (DS)-Reference architecture.

2.1.1. DS Infrastructure

In this section, we describe the main devices that appear in the infrastructure of a DS according to the reference architecture (Figure 1). We highlight some of their technical functionalities and operational features [18,23]:

Protection or Control Intelligent Electronic Device (IED): It is a class of devices responsible for executing protection functions or performing control functions in DS. This device transmits messages through the communications network, associated with aspects such as general operating states, programmed protection functionalities, and signal states for supervision purposes.

Digital Fault Record (DFR) or Backup Protection IED: It is a subclass of IED devices that can perform two main functions. On the one hand, it is in charge of recording and storing in a standard format [24] electrical and digital signals associated with fault events. This stored information allows further fault analysis and identification of spontaneous electrical events within DSs. On the other hand, this kind of device can also function as a backup protection IED.

Gateway: It is a device that translates the monitoring and reporting protocols used within a DS onto protocols implemented by utility companies in control centers, comprising supervision systems external to the DS.

Merging Unit (MU): It is a class of device in charge of digitizing electrical signals associated with current and voltage in order to make possible their transmission as sampled values within the DS. An MU can send state reports and subscribe to messages from IEDs in order to execute switching actions within the electrical system.

Switch: This network device connects various devices within a Local Area Network (LAN). Due to their properties, switches can implement mechanisms (e.g., IEEE 802.1Q VLANs) to isolate traffic in such a way it is delivered only to the intended destination devices. It is important to remark that network switches used in DSs must ensure that the message transmission time falls within the limits defined by the IEC 61850 standard [17].

Precision Time Server: It is a device in charge of providing time synchronization for all the devices in the DS. This synchronization is particularly critical for devices such as IEDs and MUs which process digitized signals coming from electrical measurements of currents and voltages [25].

Management Workstations: These are computers used for the configuration, management and analysis of both, the devices within the DS and the information generated by these devices [1].

2.1.2. DS Communication Protocols

In this section, we will outline the communication protocols used in the operation of DSs, according to the IEC 61850 standard [18,23]. However, although the most important protocols used in DSs are presented to understand the flow of information within the communication system, it is worth mentioning that the focus of this research will be the GOOSE protocol. This emphasis on GOOSE is due to its nature and criticality, and the potential to exploit vulnerabilities associated with its operation, which will be highlighted later. The analysis of other protocols and the assessment of their vulnerabilities is out of the current scope of our work, and this will be addressed in future work.

Generic Object Oriented Substation Event (GOOSE): It is a messaging model for transmitting highly critical events in a DS. It is specified in part 8-1 of the IEC 61850 standard. GOOSE is a data link layer protocol and it is based on the publisher/subscriber model. It uses multicast Medium Access Control (MAC) addresses for the identification of devices involved in its communication [5,26].

Sampled Measured Values, or Sampled Values (SMV/SV): It is a protocol defining essential messages used to transmit current and voltage digitized signals coming from current and potential transformers in the IEDs within the DS. The structure of these messages is defined in Part 9-2 of the IEC 61850 standard. Similar to GOOSE, this protocol is a data link layer protocol, and it is based on the publisher/subscriber model [1,27].

Precision Time Protocol (PTP): It is a communication protocol used to synchronize time information of devices in a network. Part 9-3 of the IEC 61850 standard defines its adoption for DSs. Similar to GOOSE and SMV/SV, it is a data link layer protocol and it is based on the publisher/subscriber model [25,28].

Manufacturing Message Specification (MMS): It is a messaging specification for industrial applications, adopted in part 8-1 of the IEC 61850 standard. MMS operates in a Client/Server Model. It uses TCP as a transport protocol (MMS servers are used to listen at port 102). It is used in the station bus for the transmission of messages from IEDs up to Supervisory Control and Data Acquisition (SCADA) systems, Human-Machine Interfaces (HMI), or Gateway devices [26,29].

Next, we present in detail elements of the operation of the GOOSE protocol, with particular emphasis on those elements of the protocol that introduce vulnerabilities that can be exploited leading to compromising the DS infrastructures.

2.2. GOOSE Operation

GOOSE messaging plays a pivotal role in the operational infrastructure of DSs. This communication protocol enables reliable real-time exchange of data and control commands among IEDs [5,26]. GOOSE is crucial to ensure operational integrity and safe and resilient operation of the overall electrical system [30].

In a DS, GOOSE messaging is mainly used for fast and reliable transmission of information associated with critical events and control signaling. In its specification, it defines requirements in order to provide deterministic data transmission [20]. IEDs broadcast GOOSE messages over the network (i.e., publishers) which are processed by receiving IEDs (i.e., subscribers). Hence, no intermediary devices are involved in the processing of GOOSE messages, which reduces response systems while increasing DS reliability [5,30].

Figure 2 shows a schematic diagram of GOOSE operation. In this figure, it is displayed how the GOOSE messaging is generated whenever a state change occurs in signals (represented as Boolean fields in protocol messages). The IEC 61850 standard defines that upon these changes, a burst of messages is generated as a confirmation of the data sent. These bursts are generated due to the occurrence of *trips* (i.e., operation of circuit breakers in the DS infrastructure) associated with protection signals, or changes on control supervision signals [31].

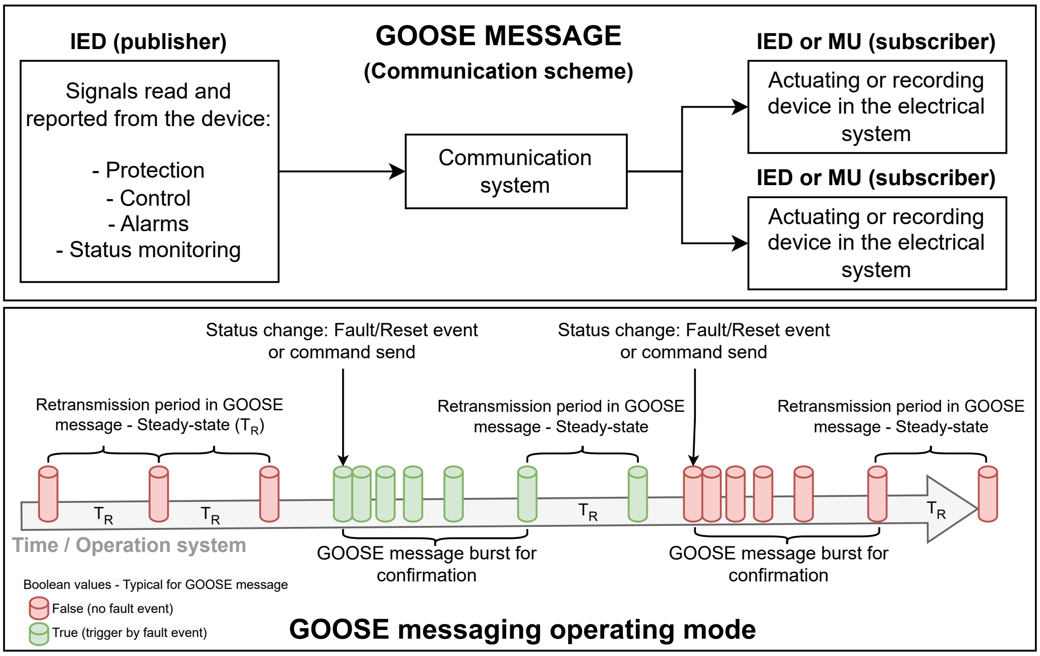

IED (publisher)
GOOSE MESSAGE
(Communication scheme)
IED or MU (subscriber)
Signals read and reported from the device:
- Protection
- Control
- Alarms
- Status monitoring
Communication system
Actuating or recording device in the electrical system
IED or MU (subscriber)
Actuating or recording device in the electrical system
Status change: Fault/Reset event or command send
Status change: Fault/Reset event or command send
Retransmission period in GOOSE message - Steady-state (T_R)
Retransmission period in GOOSE message - Steady-state
Retransmission period in GOOSE message - Steady-state
T_R
T_R
T_R
T_R
Time / Operation system
GOOSE message burst for confirmation
GOOSE message burst for confirmation
Boolean values - Typical for GOOSE message
False (no fault event)
True (trigger by fault event)
GOOSE messaging operating mode

Figure 2. GOOSE Operation scheme and mode.

2.3. Spoofing Attack

Communication systems are exposed to multiple cybersecurity threats, such as Spoofing, Distributed Denial of Service (DDoS), Man-in-the-Middle (MitM) attacks, and Ransomware, among others. A preliminary literature review allows us to identify that the most prominent attacks that affect DSs are False Data Injection and Spoofing. Therefore, we take as the focus of our research the Spoofing Attacks [3,7,10]. Spoofing attacks are those attacks where an offending device resembles a legitimate device. This resemblance is performed to inject malicious traffic, as generated from a legitimate device, to disrupt the normal operation of a DS. This action might cause failures in the overall electrical system [11].

In a DS (see Figure 3), the spoofing of messages or devices represents a critical vulnerability that might lead to various failure scenarios. These failure scenarios represent a direct impact on the operation of the electrical infrastructure [8]. For instance, impersonating a protection device or Spoofing a specific *trip* message (change in the state of a circuit breaker) might result in the omission of an activation command directed to a substation switchgear. Spoofing could lead to ignoring an operation command or fault event, potentially escalating to a large-scale event with significant effects on the safety of the electrical infrastructure. Also, due to spoofing, a subscriber device might be manipulated to execute protection actions or controlled operations that do not align with the actual state of the system. This might result in triggering failure events, often leading to cascading failures with significant impacts on the electrical system [32]. Spoofing attacks typically target IEDs, but they can also be directed against other equipment within the DSs [33]. For instance, Spoofing attacks against Precision Time Servers might cause failures in time synchronization, which is critical for IEDs and MUs [34]. In this work, we focus on analyzing the impact of a Spoofing attack against an IED.

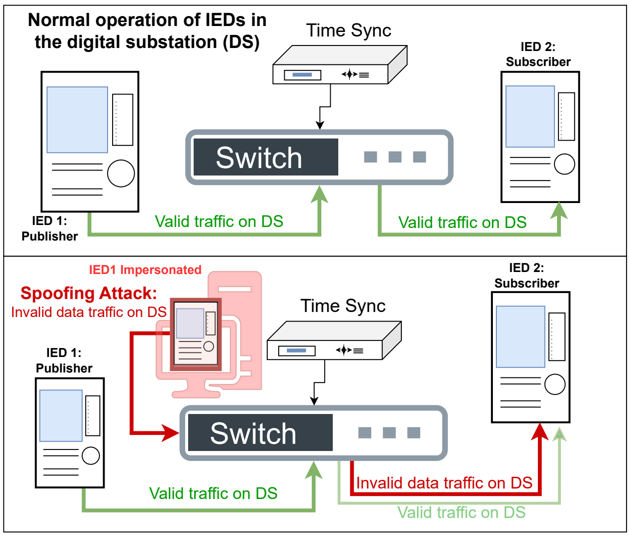


Figure 3. Spoofing Attack on Digital Substation.

3. Related Work

Vulnerability assessment is a critical task in cybersecurity research [32,35], especially for the identification in advance of possible threats against networks and infrastructures. For instance, in [36], the authors provide a conceptual framework for understanding how vulnerabilities in data and learning models can affect smart grid cybersecurity, and how cyberattacks can exploit these weaknesses. However, a restriction for vulnerability assessment is the fact that, in many cases, it is not possible to have access to production infrastructures in order to execute experiments. Thus, an approach commonly used by the scientific community is leveraging datasets to model the infrastructures under study, understand their dynamics, operations and interactions and propose models and solutions in different fields [37,38].

For years, researchers in the field of cybersecurity have been using datasets as input for their studies, and in many cases, they have also produced datasets as contributions to the community [38]. One of the first attempts at a dataset for cybersecurity research that is referenced in the literature is the *KDD-Cup99* [39]. However, due to the evolution in traffic dynamics and the emergence of new and different services and applications, *KDD-Cup99* has been rendered obsolete. New datasets such as *NSL-KDD and CICIDS 2017* have been presented to the community, containing large volumes of both, samples associated with both, benign and malign traffic, associated with a wide set of cyber attacks [40–42].

Despite the relevance of these recent datasets and their usage to produce high-relevance research [43], they are intended to be general, containing samples of diverse traffic patterns and attacks. More tailored datasets are required for the assessment and analysis of particular infrastructures such as DSs [3,44]. However, dealing with such infrastructures introduces several challenges. First of all, in the context of critical infrastructures such as DSs, there are important concerns associated with information privacy. Both, client and vendor information exchanged by the devices in these infrastructures have strict requirements imposed regarding the management of this information. Also, corporate policies enforce restrictions on the disclosure of topology and descriptive information about the equipment involved in particular deployments. Hence, the generation of datasets as contributions to the academic community is more complicated than in other research fields [3,45].

Some examples can be found in the literature presenting datasets for cybersecurity study in the context of critical infrastructures. The *HIL-based augmented ICS (HAI) of 2020* presents a dataset developed on a testbed based on a steam turbine and a hydroelectric system. This dataset contains different traffic samples of normal traffic and some synthetically generated attacks [46]. *Secure Water Treatment (SWaT)* dataset is another example of a dataset generated on critical infrastructures, in this case, a water treatment facility. This dataset contains also samples of normal traffic and induced synthetic cyber attacks [47].

There are some efforts of datasets for electrical substations. *ELECTRA* focuses on Modbus communication register [48]. *EPIC* addresses MMS but leaves other protocols out of the scope of the study [49]. *IEC61850SecurityDataset* includes GOOSE protocol traffic, but the traffic is captured by the emulation of IEDs [50]. In [51], authors present a dataset containing traffic samples of a DS in a steady state. Hence, it can be used as a baseline for the development of anomaly detection solutions. A simulated dataset features a substation operating in a steady state, which implies that the flow dynamics are constant. In this case, it is analyzed using a fractional auto-regressive integrated moving average *(FARIMA)* from which the characterization of traffic and anomaly detection is performed.

Quincozes et al. [3] conducted a survey of intrusion detection and prevention systems in DS, highlighting the importance of accurately characterizing the attributes of the communication protocols under analysis. This is a task that can be performed by leveraging adequate datasets. Additionally, Wang et al. [52] present an analysis of anomaly detection and attacks against IEDs on smart substations, considering different operation scenarios and working over simulated testbeds. This paper discusses the importance of having real datasets and the limitations associated with the availability of operational infrastructure datasets, access to electrical substation systems, and the inclusion of cyber attacks in the data samples.

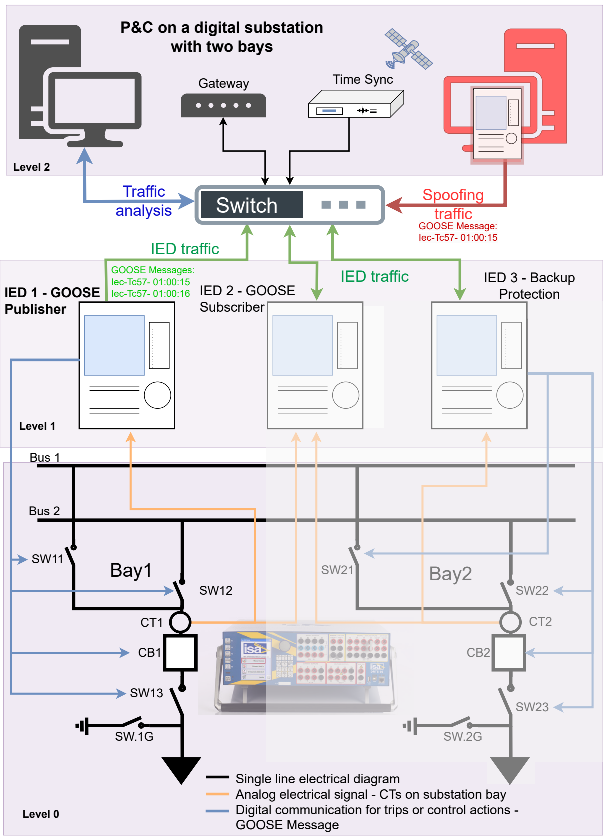
The main contribution of our paper is presenting a public dataset for cybersecurity research in Digital Substations. This dataset is generated on a real testbed, aligned with the IEC 61850 standard, rather than on a simulated infrastructure. This dataset contains both samples of normal traffic in different operational conditions, and attacks exploiting vulnerabilities of the GOOSE protocol. Hence, we aim to address two limitations of existing works. First, the dataset is based on a sampling of traffic generated with real hardware. Second, the vulnerabilities of GOOSE, especially those allowing spoofing, are extensively leveraged for the generation of actual attacks, with real impact on the infrastructure.

4. Testbed and Dataset Description

In this section, we describe the testbed used in the construction of the dataset developed in this work. We detail the infrastructure composing this testbed and we outline the scenarios considered, which fit real operational conditions of an actual infrastructure. Finally, we introduce some relevant properties of the dataset generated.

4.1. Infrastructure Description

As previously explained, DSs comprise different devices, which perform specific functions. The testbed built for this work aims at reproducing as precisely as possible the infrastructure of a small DS. For this reproduction, we include actual physical equipment that can reproduce all the operational procedures and conditions of a real-world infrastructure. Figure 4 presents a scheme of the testbed. This is a DS with two bays, a station bus level and three devices performing Protection and Control functions. The testbed also contains a Precision Time Server and a Gateway. Given its design, the testbed incorporates the communication protocols defined in the IEC 61850 standard besides complementary protocols associated with the monitoring and management of the equipment.



P&C on a digital substation with two bays
Gateway
Time Sync
Level 2
Traffic analysis
Switch
Spoofing traffic
GOOSE Message: Iec-Tc57- 01:00:15
IED traffic
GOOSE Messages: Iec-Tc57- 01:00:15 Iec-Tc57- 01:00:16
IED 1 - GOOSE Publisher
IED 2 - GOOSE Subscriber
IED traffic
IED 3 - Backup Protection
Level 1
Bus 1
Bus 2
SW11
Bay1
SW12
SW21
Bay2
SW22
CT1
CT2
CB1
CB2
SW13
SW23
SW.1G
SW.2G
Level 0
Single line electrical diagram
Analog electrical signal - CTs on substation bay
Digital communication for trips or control actions - GOOSE Message

Figure 4. Testbed used.

Next, we present some of the technical specifications of the equipment deployed in the testbed.

4.1.1. Level 2

Gateway: It is a Kalkitech brand device. It is connected to the three IEDs in the DS, for monitoring purposes

Workstations: Two high-performance workstations used for monitoring, data analysis and generation of attack traffic

Precision Time Server: It is a Meimberg GPS device. It can operate with both PTPv2 and NTP/SNTP protocols

4.1.2. Level 1

IEDs: These are equipments of ABB and GE brands. These devices have the following general characteristics:

- Network Interfaces: Ethernet 100BASE-FX ports (Optical Ports)
- Protection and control functions: Instantaneous overcurrent
- Configuration of GOOSE retransmission time set: 5 s.

For visualization purposes, the IEDs were configured with a local HMI. This interface included a flag to indicate whenever a GOOSE *trip* (change in state) occurs in the electrical system. Also, the HMI contained an indication to report whenever a device subscribes to a GOOSE *trip* message.

4.1.3. Level 0

This level includes an ISA DRTS 66 device. This device is a signal injector, responsible for emulating the electrical network and validating switch actuation upon the occurrence of a GOOSE *trip* event, which as mentioned, is associated with changes in states in electrical variables.

4.1.4. Station Bus

This bus includes a General Electric S2024 switch. This is a Layer 2 switch, supporting LAN protocols such as IEEE 802.1Q (VLANs). It provides connectivity among devices in the DS while enabling traffic isolation according to the required communication patterns (e.g., the different multicast groups for communication among IEDs).

4.2. Dataset Description

In this section, we describe the general aspects of the dataset generated in this work. This dataset contains different traffic samples representing normal operations, failure events associated with electrical conditions, and the behavior of the infrastructure when facing cyber-attacks. We focus on spoofing attacks, which are highly devastating and highly prevalent in the context of critical infrastructure, such as DSs [10,53].

4.2.1. Experiment Scenarios

In the development of this dataset, we considered four scenarios in order to cover different operational conditions, both in normal operation and when infrastructure is under attack. Next, we describe these scenarios.

Scenario 1: Stable operation of the electrical system

In this scenario, the devices are configured for joint and synchronous operation. In this stable state, devices transmit operational messages. That is, the traffic reflects the normal operation of the electrical system associated with the DS.

Scenario 2: Electrical system failure events

In this scenario, we considered the operation of the DS upon the occurrence of failure events. In order to reproduce this behavior, we configured general triggers associated with electrical variables monitored in the infrastructure. In this scenario, traffic varies according to the nature of the failure events. The goal of this scenario is to make evident the difference in traffic dynamics during failure events in comparison to the stable operation.

Scenario 3: Spoofing attack during stable operation of the electrical system

In this scenario, while the DS is transmitting information related to the stable operation of the electrical system, a spoofing attack is executed. This attack simulates the generation of a false report of a failure event reported by an attacker device. This spoofing causes the system to respond incorrectly due to the false event.

Scenario 4: Spoofing attack in the presence of electrical system failure events

In this scenario, a spoofing attack is induced while the system is facing a failure event. Given the behavior of the GOOSE protocol in failure events (i.e., generating a burst of messages), this situation of spoofing upon the occurrence of attacks becomes more devastating due to the cascade effect it might introduce.

Scenarios 3 and 4 are considered analogous to Scenarios 1 and 2, respectively. However, Scenarios 3 and 4 include spoofing attacks. Through the design of these scenarios, we want to illustrate the behavior of the DS Infrastructure both in normal stable state conditions, and when reacting to failure events that might arise during the operation of the infrastructure.

4.2.2. Time Window for Data Collection

In the four experimental scenarios, we have considered a time window of 180 s. To the best of our knowledge, this time window allows us to capture the traffic dynamics during operational conditions (normal, stable state and upon attacks), and the effect of spoofing attacks. This is aligned with the dynamics of the GOOSE protocol as described in the literature [5,23]. The time window has been defined considering the estimated duration of an electrical failure, typically less than 1 s. As can be seen in Scenarios 2 and 4, multiple failure events occur with different durations in each one, even in a short period of time, such as 180 s.

4.2.3. Dataset Structure

In this section, we describe the organization of the dataset. This includes capturing traffic with typical DS communication protocols, as mentioned above, and considering the previously defined test scenarios. The data flow contained in the dataset corresponds to the total traffic exchanged among the testbed devices. The focus of analysis of this research will be the data packets transmitted by IED 1. This device will be subject to spoofing in one of its GOOSE messages (Iec-Tc57_01:00:15). However, with the dataset data packets, transfer times, publication characteristics, and other attributes contained in the dataset, the analysis can be extrapolated to other devices or messages that may be present in the scenario under study.

For the sake of simplicity, traffic samples were stored in the libpcap-ng format, commonly used for traffic captures. Traffic samples were also exported to CSV format. Next, we present a brief description of the variables stored in the CSV files.

No: This is a sequence number assigned by the traffic capture tool. This is merely an informational and descriptive field of the record.

Time: This is a time mark, relative to the start of the capture process. It is also a merely transformational and descriptive field of the record.

Source: This is the source MAC address of the frame associated with the given record.

Destination: This is the destination MAC address of the frame associated with the given record. It is worth mentioning that for protocols such as GOOSE, this address represents a layer 2 multicast group. According to the IEC 61850 standard, the range of MAC Addresses for GOOSE protocols is Iec- Tc57_01:00:00 to Iec-Tc57_01:01:FF (Iec-Tc57_ represents the initial three octets 01:0C:CD). For this dataset, we used the following GOOSE identification for MAC Multicast destination addresses: IED 1: Iec-Tc57_01:00:15 (Protection), IED 1: Iec-Tc57_01:00:16 (Control), IED 2: Iec-Tc57_01:00:40 (Protection), IED 2: Iec-Tc57_01:00:41 (Control), IED 3: Iec-Tc57_01:00:90 (Protection), and IED 3: Iec-Tc57_01:00:91 (Control).

Protocol: This is the EtherType field of the frame associated with the corresponding record.

Boolean: This is a variable exclusive for GOOSE and MMS protocols which indicates whether the message contained in the frame reports state changes.

Info: This is an informational attribute containing a description of the corresponding frame associated with the record. It might contain information about devices, type of GOOSE message and error notifications. It is important to remark that this is not an actual protocol field.

Length: This value corresponds to the size of the corresponding frame in bytes.

4.2.4. Overview of Communication Protocols

The dataset built in this work contains traffic samples of the typical communication protocols present in a DS. That is, the dataset contains synchronization messages, scout messages, and redundancy protocols in addition to GOOSE, and PTP messages.

Table 1 presents a general overview of the distribution of packet counts for the communication protocols captured in the dataset during the experimental scenarios. It also contains information about other protocols associated with specific functions such as ARP (Address Resolution Protocol), PRP (Parallel Redundancy Protocol), HSR (High-availability Seamless Redundancy), NTP (Network Time Protocol) and PTPv2. Regarding GOOSE messaging, we discriminate GOOSE messages sent by each of the three IEDs of the testbed. Therefore, the table shows six sets of messages (three pairs), associated with signals for the protection and control in each device.

The table details the number of messages (CM: Count Message) observed in each analysis scenario. Particularly, the GOOSE messages have been separated according to their defined identification multicast MAC address to highlight the distinctive behavior of the protection and control messages. In this overview, the information associated with traffic attacks is highlighted in red.

Table 1. Communication protocols in the dataset.

	Message Type	**Scenario 1**	**Scenario 2**	**Scenario 3**	**Scenario 4**
		CM 1	**CM 2**	**CM 3**	**CM 4**
Digital Substation Protocols	GOOSE	365	401	391	487
	IED 1: Iec-Tc57_01:00:15	61	98	88	184
	IED 1: Iec-Tc57_01:00:16	62	61	61	61
	IED 2: Iec-Tc57_01:00:40	61	61	61	61
	IED 2: Iec-Tc57_01:00:41	61	61	61	61
	IED 3: Iec-Tc57_01:00:90	60	60	60	60
	IED 3: Iec-Tc57_01:00:91	60	60	60	60
	HSR/PRP	295	294	294	295
	NTP	14	14	15	13
	PTPv2	900	900	900	900
Other Network Protocols	ARP	73	74	76	75
	BROWSER	0	1	0	1
	LLDP	10	11	10	10
	LLMNR	4	12	0	0
	MDNS	24	44	0	0
	NBNS	0	3	0	0
	PRES	10	10	10	10
	SSDP	10	12	8	11
	STP	150	150	151	150
	TCP	34	34	37	34
	TPKT	1	1	0	1
	UDP	42	14	0	42
	TOTAL	1932	1975	1892	2029

CM: Count Message-GOOSE Message Iec-Tc57- 01:00:15 is spoofing attack object.

5. Dataset Analysis

In this section, we present a detailed description of the GOOSE messaging present in the dataset, in the different scenarios previously described. The dataset generated in this project is available at https://dx.doi.org/10.21227/jjv5-qg20, accessed on 9 May 2024.

5.1. Behavior of GOOSE Messages in the Dataset

In this part, we describe the behavior of the GOOSE messaging within each scenario. In order to make clearer the understanding of each scenario, we present a figure complementing the explanation of each one. The labels used in each figure are defined as follows:

Trip Count: The number of messages associated with state changes. These events of state changes are associated with the corresponding triggers configured for each scenario.

Attack GOOSE count: The number of GOOSE messages generated as part of the spoofing attack.

GOOSE_Ctrl: It marks the generation of control messages.

GOOSE_Prot: It marks the generation of protection messages.

GOOSE_Attack: It marks the start of the spoofing attack.

Stable state: It marks the time window of stable state operation of the DS. During this period of time, there are not either faults or anomalies. GOOSE messages associated with protection operations will indicate false since no protection has been triggered.

Trip (Yellow shading): It marks the time window where *trip* messages are generated in response to a fault event. GOOSE messages associated with protection operations will have a true value.

T (enclosed within a green circle): It indicates the time mark when a *trip* action is generated from the protection IED in response to an electrical fault event.

R (enclosed within a red circle): It indicates the time mark when a reset action is generated from the protection IED in response to a recovery in the electrical system.

Next, we present a detailed description of each scenario and the results obtained with the experiments performed in each case.

5.1.1. Scenario 1: Stable Operation of the Electrical System

Figure 5 illustrates the flow of GOOSE messaging during the stable state or stable operation. The figure shows the behavior in the stable state of the GOOSE messages: those associated with control functions (in grey) and the protection function (in yellow), as published by one of the IED devices in the DS (e.g., the IED 1). The system starts from an operating state with no-fault events. This is indicated through a boolean value contained in GOOSE messages which is set to False.

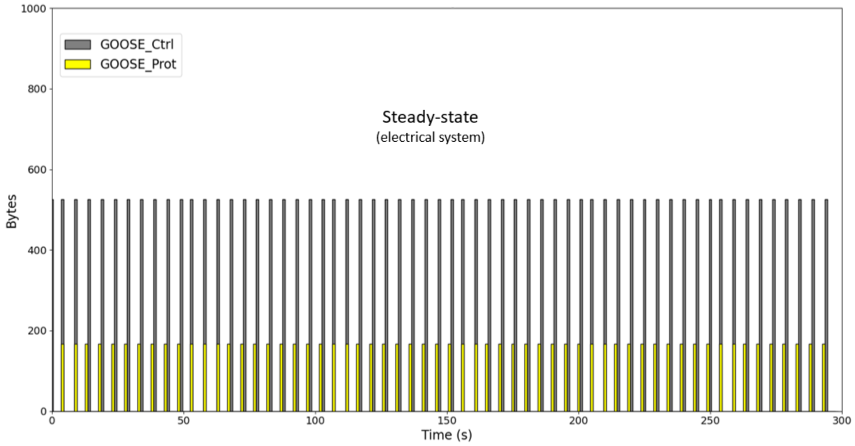


Figure 5. Scenario 1: GOOSE traffic in Stable Operation.

The figure shows that frames associated with protection packets have a size of 167 bytes whereas frames associated with control packets have a size of 526 bytes. This difference lies in the amount of data contained in the GOOSE packets, which tends to be larger in control packets. The IEC 61850 standard defines protection messages as time-critical. Hence, they contain only the information required for the particular operation. On the other hand, control messages have a variable length field containing information for the monitoring of the associated device, or control commands required to execute a particular operation. Variations in the size of these messages or in the frequency of their transmission might indicate changes in the behavior of the system.

5.1.2. Scenario 2: Electrical System Failure Events

In this scenario, several protection operations were engaged in the system, in response to induced failure events. Figure 6 shows different states of the electrical system, starting with a stable operation, followed by a fault event which starts with the activation of the protection function, and the resulting generation of GOOSE *trip* messages. This can be observed at time marks T at 55 s, 115 s, 195 s, and 280 s. Upon system recovery, new changes in state occur returning the system to normal condition. This can be observed at time marks at R at 76 s, 143 s, 222 s, and 283 s. Times for the generation of trigger and reset events are random. The figure shows how state changes are actually associated with protection messages. Hence, GOOSE control messages are transmitted periodically, as described in the previous scenario.

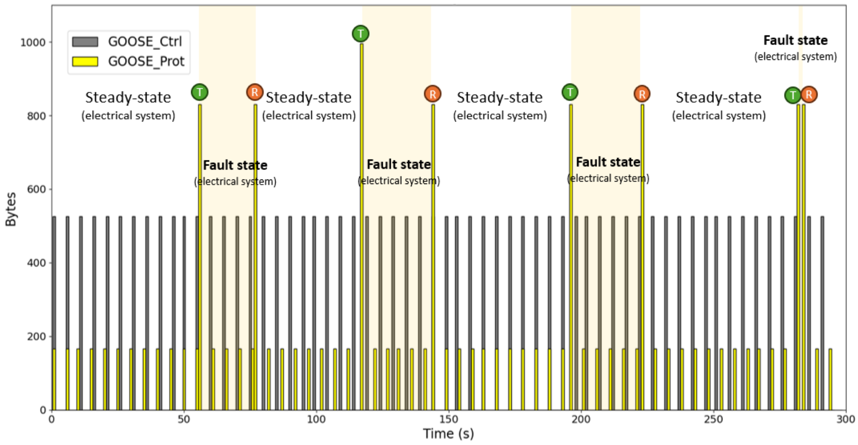


Figure 6. Scenario 2: GOOSE traffic during stable state, and upon the occurrence of some real failures in the electrical system.

The figure shows eight state changes, associated with the occurrence of four fault events (T) and their corresponding resets (R). Each state change in GOOSE messaging is characterized by an increment in the number of messages retransmitted within time windows of 500 ms. This is expressed with an increment in the size of the GOOSE protection messages due to the inclusion of the corresponding fields. This behavior is defined by the standard, in order to guarantee subscription, even during failures, since packet losses might occur due to congestion.

In real DS infrastructures, fault events do not have a fixed duration. We reproduced this behavior in our testbed by inducing failures with random duration. It is important to remark that this scenario does not contain attacks but failures, which can be conceived as events that might occur during the normal operation of the DS. This scenario is useful to determine a baseline of the behavior of the DS when dealing with failures not attributable to cyberattacks.

5.1.3. Scenario 3: Spoofing Attack During Stable Operation of the Electrical System

In this scenario, spoofing attacks are generated while the DS is in stable operation. These spoofing attacks introduce artificial failure events due to the false messages injected. Figure 7 illustrates the behavior of the GOOSE messaging during stable operation in a particular IED, which ultimately resembled the spoofing attack. The figure displays with red bars the spoofed GOOSE messages. There are 25 messages associated with the attack, which are separated into two groups. The first group is generated after 120 s and a second group is generated after 210 s. The duration of the attack is random. Hence, the change state triggering varies in each case.

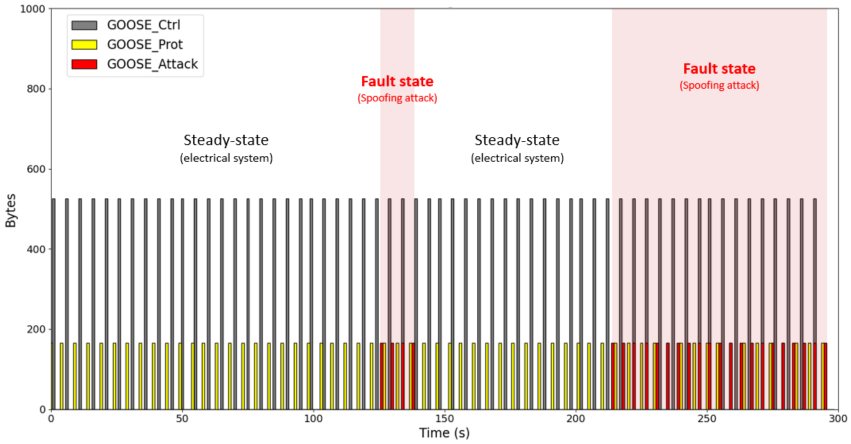


Figure 7. Scenario 3: GOOSE traffic during stable operation, and upon the occurrence of failure events due to Spoofing attacks.

In this scenario, there are no real failure events, understood as those attributable to operational electrical events. The GOOSE spoofing attacks do cause the observed failure events. Hence, it can be observed that spoofing attacks might compromise the real stability of a DS.

5.1.4. Scenario 4: Spoofing Attack in the Presence of Electrical System Failure Events

This scenario shows the combination of actual failure events with spoofing attacks. Figure 8 displays this scenario which contains 14 state changes associated with real failure events associated with operational conditions. That is seven fault events with their corresponding GOOSE *trip* messages (T) and their restoration events (R). In this scenario, 59 spoofing attacks were generated.

Similarly to the previous scenario, attacks were separated into two groups with specific characteristics. The first set of attacks introduces a change in state at a time mark of 94 s. This change in state induces the triggering of protection mechanisms, which remained active during 90 s, until the time mark at 185 s. During this time window, there were six actual state changes (*trip* and restore) not identified in the monitoring of the DS. This means that the protection system was not actually disengaged, which is a behavior consequence of the spoofing attack. The second attack scenario started around the time mark at 240 s. The attack starts shortly after the occurrence of a GOOSE *trip* message associated with a real operational event. Hence, the attack causes the system to be restored to a fake stable state operation which is kept even upon the generation of six state changes during the attack.

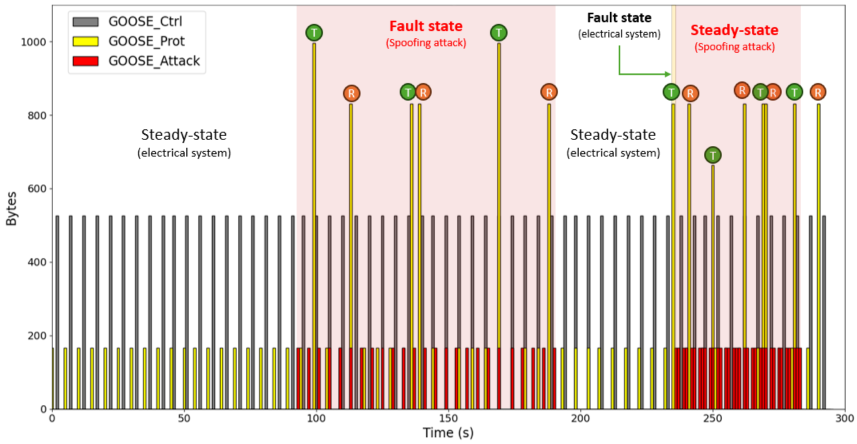


Figure 8. Scenario 4: Traffic of GOOSE protocol, Steady-State, Real Failure events, and Spoofing Failure events.

In this scenario, the Spoofing attack indeed degenerates towards a denial of service attack. The figure shows how the spoofing messages prevent the system from adequately processing the messages generated by the real events. This causes the protection and control mechanisms of the DS to become isolated.

5.2. General Remarks

From the development of the dataset presented in this paper, there are several features that became evident associated with the traffic of the protocols in a DS. The identification of these features can be considered the initial step and reference point for the development of cybersecurity solutions to cope with the security threats that might affect critical infrastructures such as DSs. First, it is evident that a prominent property of the traffic, especially during stable operation periods is its regularity. GOOSE messages flow through the network in regular intervals, with an almost constant Inter-Arrival Time. Secondly, in addition to direct features that can be extracted from traffic captures, such as the values of protocol fields indicating the type of message, source and destination addresses, sequence numbers, and packet sizes, among others, there are other useful features that can be derived. For instance, the packet Inter-Arrival Time previously mentioned can be used to determine the regularity of the packet transmission or the standard response when facing failure events. Behaviors that deviate from these dynamics observed in regular operational conditions might trigger alarms as a possible indication of anomalies or events associated with threat exploitation.

The dataset developed in this work, collected from a real testbed, with real DS equipment and reproducing actual operational conditions, is an important avenue to understand the behavior of DS infrastructures from the study of their traffic. It allows also the understanding of protocols such as GOOSE and SV. This understanding will allow the development of different solutions, not only in the cybersecurity field but also in the areas of traffic engineering and quality of service. This understanding will clearly contribute to the research of integration of novel network technologies, which will increase the resilience of performance of critical infrastructures, in general, and DSs, in particular.

6. Conclusions and Future Work

In this paper, we have presented the development of a dataset in the context of a Digital Substation. In contrast to some of the contributions available in the literature, this dataset has been developed on a testbed with actual real devices, which reproduce the operational dynamics that can be observed in real-world substations. Hence, the traffic samples captured and included in the dataset really represent the behavior and dynamics of the communication protocols such as GOOSE and SMV/SV that are used in Digital Substations.

Initial observations allowed us to understand the behavior of protocols in the stable operation of the infrastructure, and upon the occurrence of failure events due to operational conditions. Through these observations, it was evident that the traffic in these conditions presents an evident regularity. Properties such as the packet Inter-Arrival Time tend to be constant, and according to the IEC 61850 standard, there are some traffic bursts that are generated upon failure events, whose properties are predictable.

By creating different experimental scenarios, including the generation of attacks in combination with failure events due to operational conditions, we detected two important facts. First, Digital Substations are prone to Spoofing and False Data Injection attacks. Hence, and given their criticality, these are infrastructures that must be protected, especially from the point of view of their communication networks, in order to avoid further consequences. Second, these two types of attacks are not only dangerous by themselves, but they can cause Denials of Service and cascade failures which are even more catastrophic and cause incalculable effects to the users and services relying on Digital Substations.

Our literature survey has shown us that there is some related work on datasets in the context of critical infrastructures, including some examples in the field of electrical facilities.

Despite the relevance and value of these works, a noticeable drawback is that they have been developed in simulated environments, in ideal conditions and with emulated devices and infrastructures. Hence, in our work, we present a contribution based on experiments conducted on real equipment, with real operational conditions and the reproduction of actual cyberattacks, which reproduces faithfully the operation of a real-world infrastructure. We consider this dataset will be an important asset for further research on the topic of network traffic in the context of Digital Substations, and the development of novel solutions in fields such as cybersecurity, quality of service and traffic engineering.

The dataset presented in this work focused mainly on the GOOSE protocol, given its relevance and criticality in the operation of Digital Substations. But GOOSE is clearly one among other relevant protocols within these infrastructures. In future work, we plan to extend this dataset with the analysis of other protocols such as SMV/SV, MMS and PTP, which according to the literature in the area, are other important sources of vulnerabilities for Digital Substations. It is more than evident the importance of protecting DSs, since they are clearly the most critical among the critical infrastructures, due to existing interdependency. This research we performed is, therefore, a very first step, which we hope becomes extended by the collaborative work of the scientific community interested in these areas.

Author Contributions: In the paper, the various authors played an essential role in the culmination of the research. Although a collaborative effort was sought, each author undertook specific tasks. The principal author, O.A.T.-R., was responsible for consolidating a significant portion of the research and article writing. O.A.R.-R. was in charge of developing the tool implemented for the generation of attacks. G.D.R.-C. conducted corrections and made contributions to the introduction and background section. A.L.-P. conducted general revisions and contributed to enhancements in the graphics and styles used in the document. Finally, J.F.B.-V., S.A.G.-B., J.W.B.-B. and G.D.Z.-M. served as investigative leaders, overseeing, correcting, and providing necessary resources for the development of the associated research. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the General System of Royalties—Colombia (BPIN 2020000100381) and has been performed under the project "Fortalecimiento de las capacidades de investigación y desarrollo en gestión de riesgos cibernéticos en la infraestructura crítica del sector eléctrico desde las instituciones de educación superior públicas de Medellín". The APC was funded by the same project.

Data Availability Statement: The availability of data generated during the research process is consolidated on the GitHub platform, which, being widely recognized in the community, provides an infrastructure enabling anyone interested to access the data remotely and without geographical restrictions. This accessibility extends the reach of the research by allowing other researchers, students, and professionals interested in the topic to access and utilize the data in their projects. To access the repository containing all the information generated in this study, click the following link: https://dx.doi.org/10.21227/jjv5-qg20, accessed on 9 May 2024.

Conflicts of Interest: The authors of this study have no conflicts of interest in the research results. The funding comes from public resources, which guarantees impartiality in the process. Therefore, the only interest of the participants in this article is to deepen, broaden, and facilitate the study of future research related to the topic discussed here.

Abbreviations

The following abbreviations are used in this manuscript:

CI	Critical Infrastructure
DDoS	Distributed Denial of Service
DoS	Denial of Service
DS	Digital Substation
GOOSE	Generic Object Oriented Substation Event
HMI	Human Machine Interface
ICT	Information and Communication Technologies

IEC	International Electrotechnical Commission
IED	Intelligent Electronical Device
IT	Information Technologies
LAN	Local Area Network
MITM	Man-in-the-middle
MMS	Manufacturing Message Specification
MU	Merging Unit
NFV	Network Function Virtualization
OT	Operation Technologies
P&C	Protection and Control
PTP	Precision Time Protocol
QoS	Quality of Service
SCADA	Supervisory Control and Data Acquisition
SDN	Software-Defined Networking
SV/SMV	Sampled Values / Sampled Measured values

References

1. Aftab, M.A.; Hussain, S.S.; Ali, I.; Ustun, T.S. IEC 61850 based substation automation system: A survey. *Int. J. Electr. Power Energy Syst.* **2020**, *120*, 106008. [CrossRef]
2. Nair, M.M.; Tyagi, A.K.; Sreenath, N. The Future with Industry 4.0 at the Core of Society 5.0: Open Issues, Future Opportunities and Challenges. In Proceedings of the 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 27–29 January 2021; pp. 1–7. [CrossRef]
3. Quincozes, S.E.; Albuquerque, C.; Passos, D.; Mossé, D. A survey on intrusion detection and prevention systems in digital substations. *Comput. Netw.* **2021**, *184*, 107679. [CrossRef]
4. Zhang, Z.; Deng, R.; Tian, Y.; Cheng, P.; Ma, J. SPMA: Stealthy Physics-Manipulated Attack and Countermeasures in Cyber-Physical Smart Grid. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 581–596. [CrossRef]
5. Wang, S.; Yang, F.; Yan, X.; Liu, T. Analysis of GOOSE message and the engineering application for GOOSE message in the intelligent substation. *J. Eng.* **2020**, *2020*, 207–212. [CrossRef]
6. Liang, G.; Zhao, J.; Luo, F.; Weller, S.R.; Dong, Z.Y. A Review of False Data Injection Attacks Against Modern Power Systems. *IEEE Trans. Smart Grid* **2017**, *8*, 1630–1638. [CrossRef]
7. Achaal, B.; Adda, M.; Berger, M.; Ibrahim, H.; Awde, A. Study of smart grid cyber-security, examining architectures, communication networks, cyber-attacks, countermeasure techniques, and challenges. *Cybersecurity* **2024**, *7*, 10. [CrossRef]
8. Hoyos, J.; Dehus, M.; Brown, T.X. Exploiting the GOOSE protocol: A practical attack on cyber-infrastructure. In Proceedings of the 2012 IEEE Globecom Workshops, Anaheim, CA, USA, 3–7 December 2012; pp. 1508–1513. [CrossRef]
9. Roa, O.; Botero, J.F.; Gutierrez-Betancur, S.A.; Tobar-Rosero, O.A. GOOSEAttacker: Synthetic Attack Generation Tool for IEC61850. In Proceedings of the 2023 IEEE Latin-American Conference on Communications (LATINCOM), Panama City, Panama, 15–17 November 2023; pp. 1–6. [CrossRef]
10. Huseinović, A.; Mrdović, S.; Bicakci, K.; Uludag, S. A Survey of Denial-of-Service Attacks and Solutions in the Smart Grid. *IEEE Access* **2020**, *8*, 177447–177470. [CrossRef]
11. Jokar, P.; Arianpoo, N.; Leung, V.C. Spoofing detection in IEEE 802.15.4 networks based on received signal strength. *Ad Hoc Netw.* **2013**, *11*, 2648–2660. [CrossRef]
12. Oliveira, A.d.S.; Santos, H. Continuous Industrial Sector Cybersecurity Assessment Paradigm: Proposed Model of Cybersecurity Certification. In Proceedings of the 2022 18th International Conference on the Design of Reliable Communication Networks (DRCN), Vilanova i la Geltrú, Spain, 28–31 March 2022; pp. 1–6. [CrossRef]
13. Burgetová, I.; Matoušek, P.; Ryšavý, O. Anomaly Detection of ICS Communication Using Statistical Models. In Proceedings of the 2021 17th International Conference on Network and Service Management (CNSM), Izmir, Turkey, 25–29 October 2021; pp. 166–172. [CrossRef]
14. Malik, H.; Alotaibi, M.A.; Almutairi, A. Cyberattacks identification in IEC 61850 based substation using proximal support vector machine. *J. Intell. Fuzzy Syst.* **2022**, *42*, 1213–1222. [CrossRef]
15. Elmasry, A.; Albaseer, A.; Abdallah, M. OpenPLC and lib61850 Smart Grid Testbed: Performance Evaluation and Analysis of GOOSE Communication. In Proceedings of the 2023 International Symposium on Networks, Computers and Communications (ISNCC), Doha, Qatar, 23–26 October 2023; pp. 1–6. [CrossRef]
16. Ring, M.; Wunderlich, S.; Scheuring, D.; Landes, D.; Hotho, A. A survey of network-based intrusion detection data sets. *Comput. Secur.* **2019**, *86*, 147–167. [CrossRef]
17. *IEC 61850*; Communication Networks and Systems for Power Utility Automation. International Electrotechnical Commission: Geneva, Switzerland, 2013.
18. Tobar Rosero, O.A.; Pérez González, E.; Botero Vega, J.F.; Zapata Madrigal, G.; Roa, O.; Candelo-Becerra, J.E.; García Sierra, R. Digital Substations and Cybersecurity in the Transformation of the Electricity Sector. In Proceedings of the 2023 IEEE Colombian Caribbean Conference (C3), Barranquilla, Colombia, 22–25 November 2023; pp. 1–6. [CrossRef]

19. Mesmaeker, I.D. Trends in protection and substation automation systems and feed-backs from CIGRE activities. In Proceedings of the 2008 IET 9th International Conference on Developments in Power System Protection (DPSP 2008), Glasgow, UK, 17–20 March 2008; pp. 1–8. [CrossRef]
20. Apostolov, A. Impact of IEC 61850 on the interoperability and reliability of protection schemes. In Proceedings of the 2013 IEEE Power & Energy Society General Meeting, Vancouver, BC, Canada, 21–25 July 2013; pp. 1–5. [CrossRef]
21. Musil, P.; Mlynek, P. Overview of communication scenarios for IEC 60870-5-104 substation model. In Proceedings of the 2020 21st International Scientific Conference on Electric Power Engineering (EPE), Prague, Czech Republic, 19–21 October 2020; pp. 1–4.
22. Song, E.Y.; FitzPatrick, G.J.; Lee, K.B. Smart sensors and standard-based interoperability in smart grids. *IEEE Sens. J.* **2017**, *17*, 7723–7730. [CrossRef] [PubMed]
23. León, H.; Montez, C.; Valle, O.; Vasques, F. Real-Time Analysis of Time-Critical Messages in IEC 61850 Electrical Substation Communication Systems. *Energies* **2019**, *12*, 2272. [CrossRef]
24. Vahidi, S.; Ghafouri, M.; Au, M.; Kassouf, M.; Mohammadi, A.; Debbabi, M. Security of wide-area monitoring, protection, and control (WAMPAC) systems of the smart grid: A survey on challenges and opportunities. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 1294–1335. [CrossRef]
25. Hunt, R.; Dalmeny, C.; Geor, M. Time Synchronisation for IEC 61850 Systems. In *IEC 61850 Principles and Applications to Electric Power Systems*; Springer: Cham, Switzerland, 2023; pp. 95–130.
26. Lozano, J.C.; Koneru, K.; Ortiz, N.; Cardenas, A.A. Digital substations and iec 61850: A primer. *IEEE Commun. Mag.* **2023**, *61*, 28–34. [CrossRef]
27. Zakonjšek, J. CT/VT Sampled Value Acquisition Applied to IEC 61850. In *IEC 61850 Principles and Applications to Electric Power Systems*; Springer: Cham, Switzerland, 2023; pp. 301–319.
28. Balakrishnan, K.; Dhanalakshmi, R.; Sinha, B.B.; Gopalakrishnan, R. Clock synchronization in industrial Internet of Things and potential works in precision time protocol: Review, challenges and future directions. *Int. J. Cogn. Comput. Eng.* **2023**, *4*, 205–219.
29. Tightiz, L.; Yang, H. A comprehensive review on IoT protocols features in smart grid communication. *Energies* **2020**, *13*, 2762. [CrossRef]
30. Hou, J.; Hu, C.; Lei, S.; Hou, Y. Cyber resilience of power electronics-enabled power systems: A review. *Renew. Sustain. Energy Rev.* **2024**, *189*, 114036. [CrossRef]
31. Ağın, A.; Demirören, A.; Usta, Ö. A Novel Approach for Power System Protection Simulation via the IEC 61850 Protocol. *IEEE Access* **2024**, *12*, 107656–107669. [CrossRef]
32. Rajkumar, V.S.; Tealane, M.; Ştefanov, A.; Palensky, P. Cyber Attacks on Protective Relays in Digital Substations and Impact Analysis. In Proceedings of the 2020 8th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems, Sydney, Australia, 21 April 2020; pp. 1–6. [CrossRef]
33. Maziku, H.; Shetty, S.; Nicol, D.M. Security risk assessment for SDN-enabled smart grids. *Comput. Commun.* **2019**, *133*, 1–11. [CrossRef]
34. Akbarzadeh, A.; Erdodi, L.; Houmb, S.H.; Soltvedt, T.G.; Muggerud, H.K. Attacking IEC 61850 Substations by Targeting the PTP Protocol. *Electronics* **2023**, *12*, 2596. [CrossRef]
35. Rashid, M.T.A.; Yussof, S.; Yusoff, Y.; Ismail, R. A review of security attacks on IEC61850 substation automation system network. In Proceedings of the 6th International Conference on Information Technology and Multimedia at UNITEN: Cultivating Creativity and Enabling Technology Through the Internet of Things, ICIMU 2014, Putrajaya, Malaysia, 18–20 November 2014; pp. 5–10. [CrossRef]
36. Zhang, Z.; Liu, M.; Sun, M.; Deng, R.; Cheng, P.; Niyato, D.; Chow, M.Y.; Chen, J. Vulnerability of Machine Learning Approaches Applied in IoT-Based Smart Grid: A Review. *IEEE Internet Things J.* **2024**, *11*, 18951–18975. [CrossRef]
37. Alshaibi, A.; Al-Ani, M.; Al-Azzawi, A.; Konev, A.; Shelupanov, A. The comparison of cybersecurity datasets. *Data* **2022**, *7*, 22. [CrossRef]
38. Zheng, M.; Robbins, H.; Chai, Z.; Thapa, P.; Moore, T. Cybersecurity research datasets: Taxonomy and empirical analysis. In Proceedings of the 11th USENIX Workshop on Cyber Security Experimentation and Test (CSET 18), Baltimore, MD, USA, 18 August 2018.
39. Tavallaee, M.; Bagheri, E.; Lu, W.; Ghorbani, A.A. A detailed analysis of the KDD CUP 99 data set. In Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 8–10 July 2009; pp. 1–6. [CrossRef]
40. Thakkar, A.; Lohiya, R. A review of the advancement in intrusion detection datasets. *Procedia Comput. Sci.* **2020**, *167*, 636–645. [CrossRef]
41. Abdulraheem, M.H.; Ibraheem, N.B. A Detailed Analysis of New Intrusion Detection Dataset. *J. Theor. Appl. Inf. Technol.* **2019**, *97*, 4519–4537.
42. Panigrahi, R.; Borah, S. A detailed analysis of CICIDS2017 dataset for designing Intrusion Detection Systems. *Int. J. Eng. Technol.* **2018**, *7*, 479–482. [CrossRef]
43. Mittal, M.; Kumar, K.; Behal, S. Deep learning approaches for detecting DDoS attacks: A systematic review. *Soft Comput.* **2023**, *27*, 13039–13075. [CrossRef] [PubMed]
44. Quincozes, S.E.; Albuquerque, C.; Passos, D.; Mossé, D. ERENO: A Framework for Generating Realistic IEC–61850 Intrusion Detection Datasets for Smart Grids. *IEEE Trans. Dependable Secur. Comput.* **2023**, *21*, 3851–3865. [CrossRef]

45. Li, W.; Meng, W.; Kwok, L.F. Surveying trust-based collaborative intrusion detection: state-of-the-art, challenges and future directions. *IEEE Commun. Surv. Tutor.* **2021**, *24*, 280–305. [CrossRef]
46. Shin, H.K.; Lee, W.; Yun, J.H.; Min, B.G. Two ICS Security Datasets and Anomaly Detection Contest on the HIL-based Augmented ICS Testbed. In Proceedings of the 14th Cyber Security Experimentation and Test Workshop, New York, NY, USA, 9 August 2021; CSET'21; pp. 36–40. [CrossRef]
47. Goh, J.; Adepu, S.; Junejo, K.N.; Mathur, A. A Dataset to Support Research in the Design of Secure Water Treatment Systems. In *Critical Information Infrastructures Security*; Havarneanu, G., Setola, R., Nassopoulos, H., Wolthusen, S., Eds.; Springer: Cham, Switzerland, 2017; pp. 88–99.
48. Perales Gomez, A.L.; Fernandez Maimo, L.; Huertas Celdran, A.; Garcia Clemente, F.J.; Cadenas Sarmiento, C.; Del Canto Masa, C.J.; Mendez Nistal, R. On the Generation of Anomaly Detection Datasets in Industrial Control Systems. *IEEE Access* **2019**, *7*, 177460–177473. [CrossRef]
49. Adepu, S.; Kandasamy, N.K.; Mathur, A. EPIC: An Electric Power Testbed for Research and Training in Cyber Physical Systems Security. In *Proceedings of the Computer Security*; Katsikas, S.K., Cuppens, F., Cuppens, N., Lambrinoudakis, C., Antón, A., Gritzalis, S., Mylopoulos, J., Kalloniatis, C., Eds.; Springer: Cham, Switzerland, 2019; pp. 37–52.
50. Biswas, P.P.; Tan, H.C.; Zhu, Q.; Li, Y.; Mashima, D.; Chen, B. A Synthesized Dataset for Cybersecurity Study of IEC 61850 based Substation. In Proceedings of the 2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Beijing, China, 21–23 October 2019; pp. 1–7. [CrossRef]
51. Yang, Q.; Hao, W.; Ge, L.; Ruan, W.; Chi, F. FARIMA model-based communication traffic anomaly detection in intelligent electric power substations. *IET Cyber-Phys. Syst. Theory Appl.* **2019**, *4*, 22–29. [CrossRef]
52. Wang, X.; Fidge, C.; Nourbakhsh, G.; Foo, E.; Jadidi, Z.; Li, C. Anomaly Detection for Insider Attacks From Untrusted Intelligent Electronic Devices in Substation Automation Systems. *IEEE Access* **2022**, *10*, 6629–6649. [CrossRef]
53. Aoufi, S.; Derhab, A.; Guerroumi, M. Survey of false data injection in smart power grid: Attacks, countermeasures and challenges. *J. Inf. Secur. Appl.* **2020**, *54*, 102518. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.