This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2023.3272749

# Early Attack Detection for Securing GOOSE Network Traffic

Ghada Elbez*, Klara Nahrstedt†, Veit Hagenmeyer*
*Institute of Automation and Applied Informatics (IAI), KASTEL Security Research Labs, Karlsruhe Institute of Technology (KIT), Eggenstein-Leopoldshafen, Germany
{ghada.elbez, veit.hagenmeyer}@kit.edu
†Information Trust Institute (ITI), University of Illinois at Urbana-Champaign (UIUC), Urbana, Illinois, USA
{klara}@illinois.edu

***Abstract*—The requirements for the security of the network communication in critical infrastructures have been more focused on the availability of the data rather than the integrity and the confidentiality. The availability of communication in IEC 61850 substations can be hindered by Generic Object Oriented Substation Event (GOOSE) poisoning attacks that might result in threats such as Denial of Service (DoS) or flooding attacks. In order to accurately detect similar attacks, a novel method for the Early Detection of Attacks for GOOSE Network Traffic (EDA4GNeT) is developed in the present work. The EDA4GNeT method considers the dynamic behavior of network traffic in electrical substations. A mathematical modeling of GOOSE network traffic is adopted for the anomaly detection based on statistical hypothesis testing. The developed mathematical model of the communication traffic can also support the management of the network architecture in IEC 61850 substations based on appropriate performance studies. To test the novel anomaly detection method and compare the obtained results with related works found in the literature, a simulation of a DoS attack against a $66/11$ kV substation with several experiments is used as a case study.**

***Index Terms*—anomaly detection, communication network, cyber-security, electrical substations, GOOSE, IDS, IEC 61850, IEC 62351.**

## ACRONYMS

**AD** Anomaly Detection.
**ARFIMA** Auto-Regressive Fractionally Integrated Moving Average.
**ARIMA** Auto-Regressive Integrated Moving Average.
**CB** Circuit Breaker.
**CPS** Cyber-Physical Security.
**CUSUM** cumulative sum.
**DoS** Denial of Service.
**DR** Detection Rate.
**EM** Expectation Maximization.
**FA** False Alarm.
**FNR** False Negative Rate.
**FPR** False Positive Rate.
**GOOSE** Generic Object Oriented Substation Event.
**HMI** Human-Machine Interface.
**ICS** Industrial Control System.
**ICT** Information and Communication Technology.
**IDS** Intrusion Detection System.
**IEC** International Electrotechnical Commission.
**IED** Intelligent Electronic Device.
**IT** Information Technology.
**KF** Kalman Filter.
**LRD** Long-Range Dependency.
**MITM** Man-In-The-Middle.
**MLE** Maximum Likelihood Estimator.
**MMS** Manufacturing Message Specification.
**MU** Merging Unit.
**NRMSE** Normalized Root Mean Square Error.
**OT** Operational Technologies.
**PDF** Probability Density Function.
**ROC** Receiver Operating Characteristic.
**SCADA** Supervisory Control And Data Acquisition.
**SG** Smart Grid.
**SS-AR** State-Space Autoregressive model.
**SS** State-Space.
**SV** Sampled Values.
**TNR** True Negative Rate.
**TPR** True Positive Rate.
**TR** Trace of a matrix.
**VLAN** Virtual Local Area Network.
**VM** Virtual Machine.
**WGN** White Gaussian Noise.

## LIST OF SYMBOLS

$\boldsymbol{\xi}[k]$ The measurement disturbance vector.
$a_i$ An element of the $\mathbf{A}$ matrix.
$\mathbf{A}$ The state transition matrix $\in \mathbb{R}^{n \times n}$.
$\boldsymbol{B}$ The base rate representing the probability that there is an intrusion in the observed data set.
$C$ The ratio of the cost of an IDS failing to detect an intrusion and its cost when it generates a false alarm.
$\mathbf{C}$ The measurement matrix $\in \mathbb{R}^{g \times n}$.
$C_{exp}$ The expected cost metric.
$C_{ID}$ The intrusion detection capability metric.
$d$ The difference coefficient.
$\mathrm{D}$ A selection term for the measurement equation.
$\mathbf{E}$ Matrix used in the iterative computation of the parameter $\mathbf{A}$.
$\mathsf{e}[k]$ Value of the sequence $\{\mathsf{e}[k]\}$ at discrete-time $k$.
$\mathbf{F}$ Matrix used in the iterative computation of the parameter $\mathbf{Q}$.
$g$ The CUSUM decision function.

$G(\Theta)$ The estimated log-likelihood function.
$\mathbf{H}$ A selection matrix for the state equation.
$H$ Hurst parameter.
$\mathbf{J}$ Gain matrix computed in the Kalman smoother.
$k$ Discrete time index.
$k_a$ Time at which an attack occurs.
$\mathbf{K}$ The Kalman filter gain.
$\mathbf{P}$ The variance of the state vector.
$N$ Size of a time-series.
$\mathbf{Q}$ A $n \times n$ covariance matrix of the states or process noise.
$\mathrm{R}$ The variance of the measurement or signal noise.
$s[k]$ Log-likelihood ratio increment.
$S_i$ The standard deviation of a subset x calculated over the interval $[i, u]$.
$S[k]$ The novel score function.
$W_{i,u}$ The partial sum of a subset x calculated over the interval $[i, u]$.
$x$ A stochastic time-series.
$\boldsymbol{\alpha}[k]$ The state vector at sample $k$.
$\beta$ The signature of additive change on the estimates.
$\Gamma(.)$ The gamma (generalized factorial) function.
$\gamma$ The threshold for the statistical detector.
$\boldsymbol{\eta}[k]$ A state disturbance vector $\in \mathbb{R}^{g\times 1}$.
$\boldsymbol{\epsilon}$ Model residuals.
$\boldsymbol{\Theta}$ The parameter vector.
$\hat{\mu}$ Sample mean.
$\boldsymbol{\mu}$ The conditional mean of the state vector.
$\sigma_k$ Variance matrix of the innovations.
$\boldsymbol{\sigma}$ The conditional variance of the state vector.
$\sigma_e^2$ Variance of a white Gaussian noise process.
$\hat{\sigma}$ Sample variance.

## I. Introduction

To ensure an optimal operation of the electrical grid, security of the communication of this critical infrastructure is of a first concern. Work in the field of security in modern smart grids has been getting increasing interest within the research community. However, enhancing the security of smart grids requires improvement in different parts such as, for instance, the transmission and the distribution substations as their communication structure was not developed with security being a primary concern [1].

In fact, the increased interconnection of Information and Communication Technology (ICT) in transmission and the distribution substations increases their exposure to cyber-attacks. Different works in the literature e.g. [2]–[4] have shown the several vulnerabilities of smart grids.

Requirements including the time-critical operation of the power grid as well as the high availability of the communication network shall be considered when designing defense mechanisms against the aforementioned threats. Thus, several aspects such as hardware with multiple performance requirements, a reliable and safe software for control systems [5], and a secure communication network traffic are necessary to take up the challenge of securing next-generation energy systems.

Smart grids are composed of a heterogeneous structure with a high-level of integration between the physical and the IT system. Thus, for combining Operational Technologies and Information Technology cyber-security is necessary to ensure a secure operation of SGs. In [6], a detailed description of the network architecture in smart grids is presented including the communication infrastructure within electrical substations. A security analysis of the different attacks that target the smart grid are also detailed in the referred publication.

When compared with integrity and confidentiality, the availability of the data transmitted within the communication network of IEC 61850 substations is of a major concern [3] as it was also highlighted in [6] where Distributed Denial of Service (DDoS) attacks are thoroughly investigated.

Denial of Service (DoS) attacks resulting from GOOSE poisoning attacks are a considerable threat to the availability of the data [7]. To counter it, use of Intrusion Detection System (IDS) is suggested in the IEC 62351 standard where different recommendations to enhance the security of smart grids including electrical substations, are presented.

Although extensive research work [8]–[14] has been reported to develop IDS in Industrial Control System (ICS) and SCADA systems, we will only focus on anomaly-based IDSs for electrical systems. Contrarily to rule and specification-based IDSs, anomaly-based methods are able to detect zero-day attacks which make them more adapted to energy systems where scarcity of available data makes it hard to establish a satisfying set of rules.

Few of the available anomaly detection methods combine a good detection performance together with accounting for the specific features of IEC 61850 substations. A survey of learning-based detection methods for IoT systems including critical infrastructures such as electrical grids is presented in [15]. A list of widely used datasets for attack detection is also reported. However, the proposed list shows that available datasets from the attacks in the energy domain are limited. Thus, accurate simulation of the attacks in testbeds, as presented in this work, can overcome these limitations. The scalability of the detection solutions is an additional requirement lacking in available methods that is raised by the authors in [15]. Again, the scalability problem can be overcome by using an extensible model for the anomaly detection as developed in EDA4GNeT.

In Table I, a comparison between the closest works to the method developed in the present work is established. In fact, the considered characteristics are focused on the detection of DoS attacks, resulting from GOOSE poisoning attacks, while accounting for the specific characteristics of the IEC 61850 network traffic and the variations in the communication traffic.

In the following, we present a summary of the major limitations of available approaches. First, specific characteristics of the network in IEC 61850 substations including the different types of communications is not considered in available anomaly detection methods based on network telemetrics as in [16], [17]. Second, most of the models of the network traffic in the substations presented in the literature [12], [16]–[18] and considered for the anomaly detection rely on simplified assumptions for the representation of the substation network. Third, to the best of our knowledge, none of the existing work proposes an early anomaly detection approach of GOOSE attacks in IEC 61850 substations. In fact, the model in [7] corresponds to an accurate representation of the network in IEC 61850. However it cannot provide an early detection of advanced DoS attacks since $j$-step ahead predictions cannot be computed from the proposed model. Consequently, the early detection feature is not supported in any of the previous works as shown in Table I.

TABLE I
CHARACTERISTICS OF THE CLOSEST AVAILABLE IDS FOR IEC 61850 SUBSTATIONS

| IDS | Adaption for 61850 | Detection of DOS | No need for specifications | Accounting for characteristics | Robustness to variations | Early detection |
|---|---|---|---|---|---|---|
| [19] | ✓ | ✓ | ✓ | - | ✓ | - |
| [13] | ✓ | - | - | - | - | - |
| [17] | ✓ | ✓ | ✓ | ✓ | ✓ | - |
| [14] | ✓ | - | - | - | - | - |
| [7] | ✓ | ✓ | ✓ | ✓ | - | - |
| [20] | ✓ | ✓ | - | - | ✓ | - |
| [21] | ✓ | ✓ | - | - | ✓ | - |

In order to overcome those limitations and efficiently tackle threats such as DoS or flooding attacks caused by GOOSE poisoning attacks, we have developed a well-adapted Early Detection of Attacks for GOOSE Network (EDA4GNeT) method.

The developed EDA4GNeT method includes the specific characteristics of the network traffic in IEC 61850 substations in the anomaly detection approach. In fact, one of the main challenges tackled within the present work is to analyze the communication besides its added complexity due to the focus on the physical process as well as the use of several protocols for the network in electrical substations. In the following, we list the main contributions of the presented work:

- We have developed a mathematical model based on a State-Space representation of an ARFIMA process, the structured analysis of the communication network of GOOSE traffic in IEC 61850 substations. Although the primary use of the developed mathematical model is the anomaly detection in EDA4GNeT, it can support design of the network architecture of electrical substations and performance studies of the communication.
- We have developed an early anomaly detection method EDA4GNeT in order to detect DoS attacks coming from the infamous GOOSE poisoning attacks. We have designed and implemented an accurate early detection, using the multi-step ahead prediction for the State-Space (SS) model. The novel method is thoroughly explained using a block diagram, enumerating the different steps and a pseudo-code of the EDA4GNeT algorithm. Early detection of attacks is essential to allow implementation of corrective actions through response systems and enhance the overall security of IEC 61850 substations.
- The novel EDA4GNeT method considers dynamic changes in the traffic and our results show considerable rate decrease of false alarms. EDA4GNeT also allows the detection of multiple anomalies at unknown change times.
- We perform the evaluation of the performance of the novel detection method EDA4GNeT through performance metrics such as detection rate and False Alarm (FA) . The considered simulation includes DoS attack, resulting from a GOOSE poisoning attack on a $66/11\,\text{kV}$ substation. We perform several experiments under different conditions.

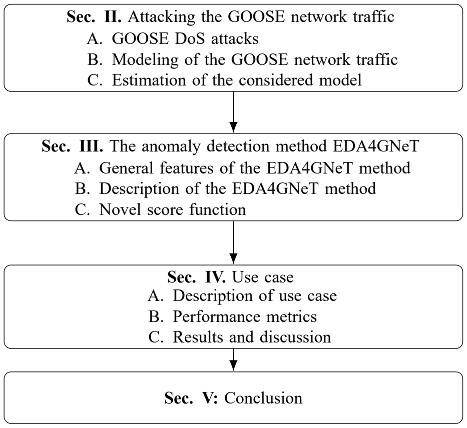
Sec. II. Attacking the GOOSE network traffic
A. GOOSE DoS attacks
B. Modeling of the GOOSE network traffic
C. Estimation of the considered model
Sec. III. The anomaly detection method EDA4GNeT
A. General features of the EDA4GNeT method
B. Description of the EDA4GNeT method
C. Novel score function
Sec. IV. Use case
A. Description of use case
B. Performance metrics
C. Results and discussion
Sec. V: Conclusion

Fig. 1. Organization of the remainder of the paper

The remainder of the paper is organized as shown in Fig. 1. In Section II, we present the description of the substation network traffic as an approximated ARFIMA state space model. Furthermore, we explain necessary foundations for the understanding of the modeling procedure as well the estimation of the different parameters. Section III describes the developed anomaly detection method EDA4GNeT including its general features and details about the detection technique. In Section IV, we evaluate EDA4GNeT using a case study and compare its results with related works available in the literature. Our method shows a strong performance using different performance criteria. Finally, we conclude and describe future work in Section V.

## II. Attacking the GOOSE network traffic

### A. GOOSE DoS attacks

Normal operation in electrical substations might be affected by disturbances or malicious actions, thus in the following the considered threat model is described. In modern electrical substations, HMIs are equipped with monitoring and control interfaces that are remotely accessible. Attackers are assumed to be able to compromise entry points in HMIs through monitoring and control interfaces that are remotely accessible, using for instance social engineering in order to connect to IEDs.

The exploitation of protocols used in IEC 61850 substations is described in [22]. The resulting DoS attacks are thoroughly explained with a focus on their impact on the substation. The different developed case studies are simulated using an OPNET model and the resulting consequences of the implemented attacks are presented. Accurate network telemetrics including Ethernet delay and link utilization are analyzed in order to investigate the system performance under attacks.

In fact, impact of loosing availability in communication networks of critical infrastructures can be much more severe than, for instance, in commercial systems. An attack script and a traffic attack replay tool such as Tcpreplay [23] are used to synthetically generate an attack from a normal operation scenario by injecting malicious GOOSE packets.

In the present work, we assume implementations of security recommendations such as demilitarized zone (DMZ) to isolate internal networks, reduce and control access to the substations LANs. However, we consider that several security breaches to DMZ can be exploited by attackers to access the substation network and inject malicious packets. Most common security breaches in critical infrastructures include phishing, compromising a domain controller or session hijacking or MITM.

Assuming that the attacker is able to exploit security breaches in the network, the attack model is based on compromising the GOOSE communication by spoofing the transmitted messages and masquerading a legitimate IED to inject maliciously crafted GOOSE messages. One of the ways to perpetrate a GOOSE DOS attack is to flood the network with bogus frames. More details about the generation framework of the different attack and attack-free scenarios can be found in [24].

By masquerading a legitimate IED, maliciously crafted GOOSE messages with a higher *StNum* than the legitimate packets, would result in a DoS attack. The injection of a maliciously crafted packet is only possible when the attack flooding rate is higher than the legitimate transmission rate. When the sending advantage of the attacker of packets with a higher StNum over the legitimate GOOSE publisher is reached, the poisoning attack can start. In the present case study, it is assumed that the attacker has acquired knowledge about the legitimate GOOSE frame rate in previous reconnaissance steps. In fact, from a defender perspective, the worst case scenario is when an attacker able to launch a successful poisoning attack with a small injection rate which would be hardly distinguished from the normal traffic. In the present threat model, this worst case scenario is assumed in order to demonstrate the capabilities of the EDA4GNET detection method. It is worth noting that further analysis of the success rate of poisoning attacks, required to cause a DoS attack shall be conducted, which is however, out of the scope of the present work. In summary, the DoS resulting from the GOOSE poisoning attack refers to a service that can not be correctly executed. In fact, the DoS is not a volumetric attack that creates flooding messages, but rather disables the service from being correctly executed since the state sequence number is changed. Interested reader can find further details in [2], [25], [26].

### *B. Modeling of the GOOSE network traffic*

In order to detect DoS attacks resulting from GOOSE poisoning attacks described in Section II-A, the use of an adapted anomaly detection system can address the problem as discussed in Section I. In fact, anomaly-based detection is a class of IDSs that is based on characterizing the normal behavior of a system, in our case, the GOOSE network traffic. An anomaly is, thus, referred to as a deviation from the normal behavior.

To thoroughly describe the characteristics of the GOOSE network traffic, an ARFIMA model is presented in [7]. The ARFIMA model is a generalization of the integer order in an autoregressive integral moving average (ARIMA) model. The use of fractional difference operator rather than an integer one as in ARIMA models, was suggested in [27] in the context of hydrology in order to represent the Long-Range Dependency (LRD). Using a state-space representation of the ARFIMA model offers several advantages including an efficient computational implementation of the model estimates as well as a general expression for a multivariate data. Additionally, for the early detection feature of EDA4GNeT, the $j$-step ahead prediction can be obtained using the Kalman Filter (KF) as explained in Section III-C.

Use of State-Space models for description of processes with LRD has been presented in several works [28]–[30]. As introduced by Chan and Palma in [30], ARFIMA models with long-memory determined by the parameter $d$ can be approximated using State-Space representation.

The general representation of a state-space model includes two equations. The first expression, namely the transition equation and the second one being measurement equation which describes the relation between the time series $\mathbf{x}[k] \in \mathbb{R}^g$ and the state vector $\boldsymbol{\alpha}[k]$. The first equation, namely the transition equation, defines the evolution of the state vector $\boldsymbol{\alpha}[k]$ and is described by Eq. (1):

$$\boldsymbol{\alpha}[k+1] = \mathbf{A}\boldsymbol{\alpha}[k] + \mathbf{H}\boldsymbol{\eta}[k],\ \boldsymbol{\eta}[k] \sim \mathcal{N}(0, \mathbf{Q}) \qquad (1)$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the state transition matrix and $\mathbf{H} \in \mathbb{R}^{n\times g}$ is the selection matrix and $\boldsymbol{\eta}[k]$ is a $g \times 1$ disturbance vector. $\mathbf{Q}$ is the $g \times g$ covariance matrix.

The second equation representing the time series $\mathbf{x}[k]$ is defined as follows according to [28]:

$$\mathbf{x}[k] = \mathbf{C}\boldsymbol{\alpha}[k] + \mathbf{D}\boldsymbol{\xi}[k],\ \boldsymbol{\xi}[k] \sim \mathcal{N}(0, \mathbf{R}) \qquad (2)$$

where $\mathbf{C} \in \mathbb{R}^{g\times n}$ is the measurement matrix and $\boldsymbol{\alpha}[k] \in \mathbb{R}^n$ is the state vector. D $\in \mathbb{R}^{g\times n}$ is a selection matrix and $\boldsymbol{\xi}[k]$ is an $n \times 1$ vector.

According to [30], the previously presented State-Space system can be written as an autoregressive model $AR(\infty)$ for long-memory models. Choosing a long enough truncation lag $m$ allows the evaluation of an approximation of the likelihood function. Thus, according to [28] the $AR(m)$ model can be represented in State-Space form as described in Appendix A.

Calculation of the maximum likelihood estimation of stationary generalized LRD models using parametric approaches can require high computational resources. Authors in [31] develop a Bayesian sampling algorithm for a bi-variate process with a stationary long-memory component. A sampling schema for stationary generalized long-memory models with one or more latent ARFIMA components was proposed to compute the maximum likelihood estimator in [32]. It was shown in [31] that the numerical computation of the estimation increases when more than two latent components are used. In the next sections, a State-Space model using an AR description is considered to represent the ARFIMA model according to [30].

### *C. Estimation of the considered model*

Analysis of the GOOSE network traffic using relevant invariants of the communication network presented in [33] shows presence of Long-Range Dependency (LRD) characteristics [7] that can modeled using a state-space representation of an ARFIMA model as presented in Section II-B. In the present section, estimation of the parameter vector of the considered model is tackled. Subspace methods can be considered for parameter estimation of state-space models for long-range fractionally integrated models as they can be practically implemented as described in [34]. Subspace procedures are based on a model reduction applied to an initial high-order vector autoregression estimate. Some of their main advantages are on one hand the possibility of handling problems of missing values or demeaning and de-trending [34] and on the other hand the efficiency of their numerical implementation.

Considering the specific representation of the State-Space model introduced in Eq. (9), an alternative estimation method needs to be adopted since there are constraints associated to the estimation problem. Parameter estimation of models as in Eq. (9) can be retrieved using the expectation maximization (EM) algorithm which is shown to be a robust solution as described in [35]. In the following, we present the definitions of conditional mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ of the state vector, respectively:

$$\boldsymbol{\alpha}[k|k-1] = E\left(\boldsymbol{\alpha}[k]|\mathbf{x}[k-1]\right) = \boldsymbol{\mu}, \tag{3a}$$

$$\mathbf{P}[k|k-1] = \mathrm{Var}\left(\boldsymbol{\alpha}[k]|\mathbf{x}[k-1]\right) = \boldsymbol{\Sigma} \tag{3b}$$

According to [35], an iterative maximum likelihood estimator of the parameters of the State-Space model is derived from the following log-likelihood:

$$\begin{aligned}\log L = &-\frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{\alpha}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\alpha}-\boldsymbol{\mu})\\ &-\frac{N}{2}\log|\mathbf{Q}|\\ &-\frac{1}{2}\sum_{k=1}^{N}(\boldsymbol{\alpha}[k]-\mathbf{A}\alpha[k-1])^{\mathrm{T}}\mathbf{Q}^{-1}(\boldsymbol{\alpha}[k]-\mathbf{A}\boldsymbol{\alpha}[k-1])\\ &-\frac{N}{2}\log|\mathbf{R}|\\ &-\frac{1}{2}\sum_{k=1}^{N}(\mathbf{x}[k]-\mathbf{C}\boldsymbol{\alpha}[k])^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{x}[k]-\mathbf{C}\boldsymbol{\alpha}[k])\end{aligned} \tag{4}$$

The observations $\{\mathbf{x}[0], \ldots, \mathbf{x}[N-1]\}$ are accessible but some hidden states are unknown. Thus, only the estimated log likelihood is available and calculated as in Appendix B.

In order to analyze the convergence properties of the Expectation Maximization (EM) algorithm, an introductory example of an ARFIMA$(1, d, 1)$ model reported in [28] is considered. Fig. 2 shows the data of an experiment using this exemplary process. The simulation data representing an experiment is depicted in the top part of Fig. 2 which is used for parameter estimation. The parameters of the State-Space (SS-AR) approximation of the ARFIMA model, are estimated using the EM algorithm. The bottom part of Fig. 2 shows the validation dataset (black) and the model predictions $\hat{x}[k]$.

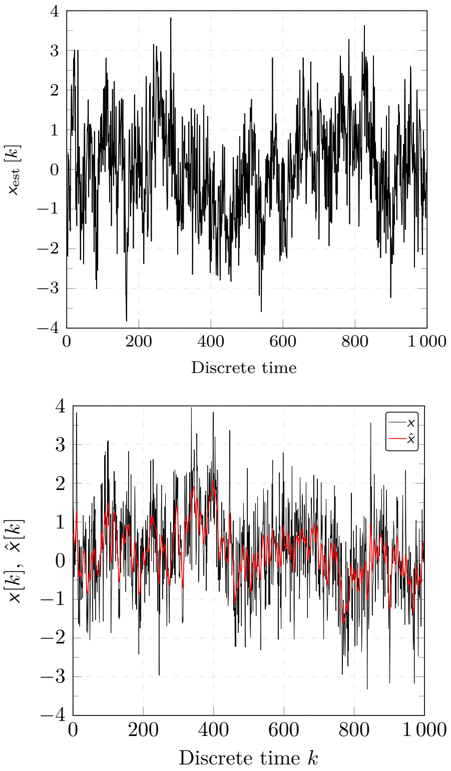

Fig. 2. Data from an experiment using and ARFIMA(1,d,1)

A total of 25 experiments with different WGN realizations is performed on the introductory example and the results of the maximization of the log-likelihood function in the estimation algorithm are shown in Fig. 3 for each experiment. For the considered use case, the log-likelihood function converges after approximately 20 iterations as depicted in Fig. 3.

The average value of the model parameters computed on the total number of experiments are the following:

$$\mathbf{A} = \begin{pmatrix} 0.33 & 0.21 & 0.24 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{5a}$$

$$\mathbf{C} = \begin{pmatrix} 1, 0, \ldots, 0 \end{pmatrix}, \tag{5b}$$

$$\mathbf{D} = 0, \tag{5c}$$

$$\mathbf{Q} = \begin{pmatrix} 0.47 & 0 & 0 \\ 0 & 1.73 & \\ 0 & 0 & 0.16 \end{pmatrix} \tag{5d}$$

$$\mathbf{R} = 1.02 \tag{5e}$$

Estimation and convergence properties of he log-likelihood function of the SS-AR model representing the GOOSE network traffic, are discussed in the following. The different values of the estimated parameters $a_i$ of the matrix $\mathbf{A}$ of the obtained SS-AR are shown respectively in Fig. 4. For each experiment, the parameters converge after some iterations which is consistent with the results shown in Fig. 3. However, the parameter $a_2$ ranges over a larger span of values in the first iterations in comparison with the other parameters. The cost

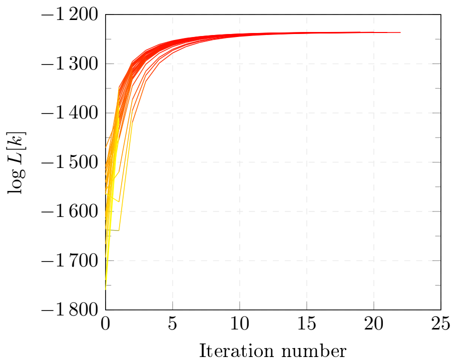

Fig. 3. Maximization of the log-likelihood function

function in the expectation-maximization (EM) algorithm is minimized for different combination of the parameters shown in Fig. 4. The convergence properties are similar to Fig. 3. The expectation-maximization algorithm can be adjusted by fixing the value of some parameters, i.e initial values, and computing the remaining ones in an iterative procedure. Thus, knowledge about the signal under analysis is required since the initial values of the parameters should be assigned based on the experience of the user. The estimation method is implemented numerically and the convergence can be improved by adjusting the choice of initial values.

## III. The Anomaly Detection Method EDA4GNeT

### A. General features of EDA4GNeT

The network traffic in industrial control and in energy systems is commonly assumed to be at steady-state [36]. However, as result of changes in the operating conditions of the industrial system or the grid, the network traffic might exhibit time-varying characteristics according to [20], [36]. Thus, one of the main features of the developed method is to account for the aforementioned fluctuations which allows its adaptability to the dynamics of the system. This guarantees that the test statistics are not affected by fluctuations in normal operation and remain within the expected range which reduces considerably the rate of false alarms and enhance the overall performance of our ED4GNeT detection method.

A second property that is taken in account for the design of our novel detection method is a recursive implementation. A similar property is required for the adaption of the network traffic model in real-time application such as the transmission of some particular GOOSE messages within IEC 61850 substations.

The adoption of a recursive implementation avoids unnecessary computational complexity and memory problems as the computation at $k$ are based at values on $k-1$. Thus, appending a new each collected sample of the network traffic to the training set at each iteration is avoided. The predictions of the state-space AR model developed in the present work are calculated recursively i.e. the prediction at the current sample is estimated using only previous samples circumventing the aforementioned limitations.

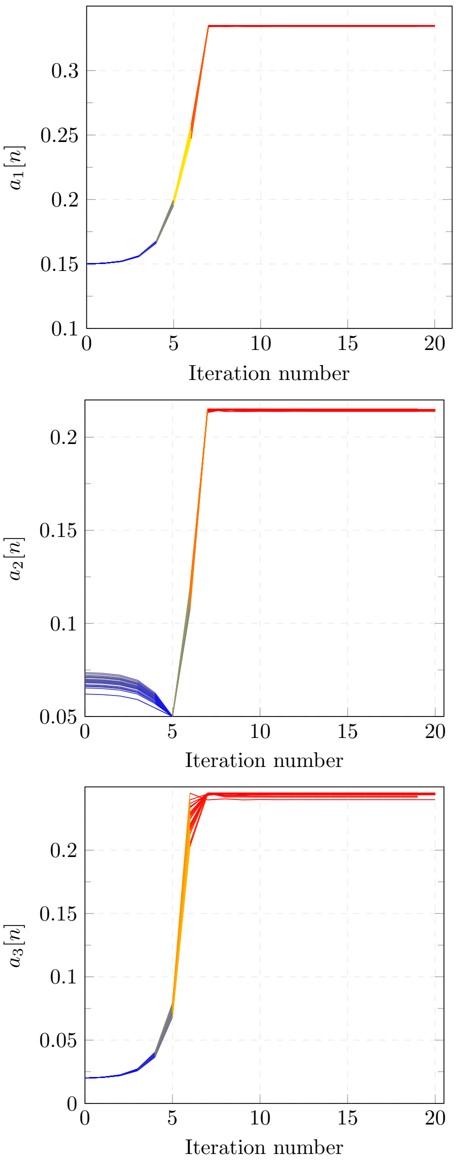

Fig. 4. Convergence of the first parameter $a_1$ of the $\mathbf{A}$ matrix in the SS-AR model

In addition to the previously described properties, the detection time is considered as a design requirement in EDA4GNeT method contrarily to available IDS where the detection delay is only considered as a performance metric. In fact, the detection delay corresponds to the time difference between the occurrence of an alarm and the actual time of an attack.

Early detection of GOOSE attacks in IEC 61850 substation network is possible with EDA4GNeT since it is based on a robust forecasting algorithm. This additional feature, that has not been considered in any of the previously available for IEC 61850 networks, is central in helping prevent power failure and revenues due to possible consequences of cyber-attacks on the grid such as fatigue damage or resonance attacks [3].

### B. Description of EDA4GNeT

A graphical description in the form of a block diagram is depicted in Fig. 5 to represent the EDA4GNeT method.

The analyzed network traffic is represented using an adequate mathematical model as shown in Fig. 5. The parameter of the chosen model are estimated in a further step. The modeling procedure is completed whenever a satisfactory model is obtained through the validation step represented by the fourth block in Fig. 5. The early detection is based on the multi-step ahead prediction vector stemming from the fifth block. The different steps as well as the parameters used for the modeling, prediction and detection, depicted in Fig. 5, are thoroughly explained in the following.

*1) Analysis and modeling of the network traffic:* Analysis of the characteristics of the network traffic in IEC 61850 substations show presence of Long-Range Dependency (LRD) properties that can be described by an ARFIMA model [7] to represent the network $\mathbf{x}_{\text{est}}[k]$. An ARFIMA model is an extension of the autoregressive fractionally integrated moving average (ARIMA) model that allows use of non-integer values of the differencing parameter $d$. More details can be found in the pioneer works of [37] and [38].

A state-space approximation of the ARFIMA model using an SS AR model introduced in [30] and explained in Section II-B, is selected to describe the communication in the modeling step.

*2) Estimation of the parameter vector:* The network traffic $\mathbf{x}_{\text{est}}[k]$ is used for the estimation of the parameter vector $\hat{\mathbf{\Theta}}$ that includes the parameters of the state-space model ($\mathbf{A}$, $\mathbf{Q}$, $\mathbf{R}$) defined in Eq. (9).

Due to the specific form of the matrices $\mathbf{A}$ and $\mathbf{C}$, an adapted version of the Expectation Maximization (EM) algorithm is developed which is one of the contributions of the present work. The values resulting from the estimation algorithm are used for further computations of the prediction values.

In order to guarantee the adaptability to the dynamics of the network traffic, use of KF with suitable initial values allows to efficiently fit the dynamics of the traffic and guarantee a good performance of EDA4GNeT. In fact, the KF guarantee optimal estimates in case of linear models with White Gaussian Noise (WGN) and its use is adapted for the estimation of the system state by minimizing the mean squared error.

The Expectation Maximization (EM) algorithm offers a good performance as a fast convergence is achieved after few iterations [35]. A discussion about the convergence of the EM algorithm is introduced in Section II-C.

The predictions of the GOOSE network traffic denoted by $\hat{\mathbf{x}}[k]$ are further computed using the parameter vector $\hat{\mathbf{\Theta}}$.

*3) Validation of the model:* Within the validation step, the model and residuals analysis are the considered criteria. In the validation step, the Normalized Root Mean Square Error (NRMSE) as well as the distribution of the residuals $\varepsilon_{\text{val}}[k]$ resulting from the difference between the real signal $\mathbf{x}_{\text{val}}[k]$ and the predicted one $\hat{\mathbf{x}}[k]$ are used to assess the quality of the model.

The accuracy of the obtained model is subject to repeating the computation of the EM algorithm within the estimation step until acceptable results are obtained. Whenever a satisfactory model is achieved after the validation step, the parameter vector $\hat{\mathbf{\Theta}}$ can be further used for the recursive computations of the predictions.

*4) Multi-step ahead prediction:* The “multi-step prediction” stage takes as input $\hat{\mathbf{\Theta}}$ and includes the computation of the KF equations to calculate the $j$ step-ahead prediction $\hat{\mathbf{x}}[k+j|k]$. For the value $j = 1$, $\hat{\mathbf{x}}$ represents the commonly known one-step ahead predictor. Further details are presented in Section III-C

The EDA4GNeT method can be used with one-step ahead predictor and a Cumulative Sum (CUSUM) test for the attack detection based on the residuals $\varepsilon_{\text{val}}[k]$. In fact, an anomaly observed in the measurement or in the transition equations can also be detected in the residuals [39].

For the early detection feature, the detector for EDA4GNeT is based on a novel score function based on the $j$-step ahead prediction, with $j$ chosen as $j > 1$, instead of the residuals. More details about the score function used for the detection test is described in Section III-C.

### C. Novel score function

One of the main advantages of an early detection of DoS attacks is to reduce operational costs by avoiding loss of availability of the network in substations.

The novel EDA4GNeT method offers an early detection feature with the help of a novel score function based on the j-step ahead prediction $\hat{\mathbf{x}}[k+j|k]$.

The model predictions are computed based on the estimation of the state described in Appendix C.

In the present part, we provide a detailed description of the detection step within the EDA4GNeT method including the novel score function.

The considered detection problem can be formulated as an early change point detection where the anomalies occur at unknown times resulting in changes in the statistical properties of the network traffic. The main challenge of the early detection is to ensure a satisfactory performance together with a reduced False Alarm (FA) rate.

The early detection in EDA4GNeT is based on a new score function, inspired by [40]. Indeed, the concept developed in [40] is extended for an early detection and a parametric approach with the test statistic $g[k+j]$ computed as follows:

$$\begin{aligned} k_{alarm} &= \min\{k : g[k+j] \geq \gamma\} \\ with\ g[k+j] &= \max(0, g[k+j-1] + S[k+j]) \end{aligned} \tag{6}$$

The novel score function $S[k]$ is based on the predictions of $\hat{\mathbf{x}}[k+j|k]$ and defined by

$$S(\hat{\mathbf{x}}[k+j|k]) = a_1\hat{\mathbf{x}}[k+j|k] + a_2\hat{\mathbf{x}}^2[k+j|k] - a_0 \tag{7}$$

The representation of the score function in a linear-quadratic form enables the positive design parameters. The design parameters $a_0, a_1$ and $a_2$ used in in the linear quadratic form of the score function, help account for changes in the mean and in the variance representing an anomaly.

The previously described steps are included in the EDA4GNeT algorithm as shown in Fig. 6. The different equations described in Section II to Section III-C are summarized as a pseudo-code in Fig. 6.

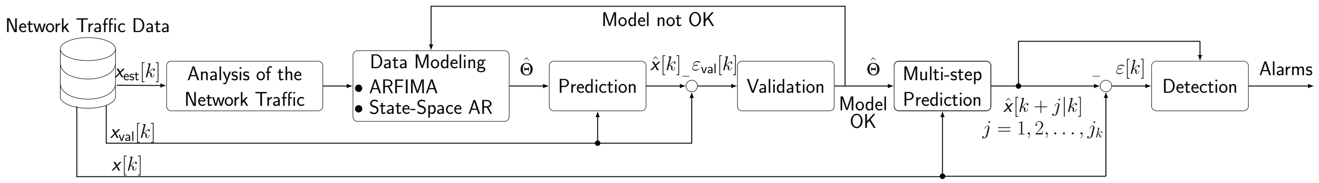

Fig. 5. Block diagram of the EDA4GNeT method

```
Algorithm 1: EDA4GNeT for early detection of GOOSE poisoning attacks
Input: Network traffic data $\{\mathbf{x}_{\text{est}}[k]\}$, order of the SS-AR model $a_m$,
       initialization values for EM algorithm $A_0$
Output: Early detection time for possible DoS attacks, $k_{alarm}$
Data: Network traffic, $\mathbf{x}^N = \{x[0], x[1], \ldots, x[N-1]\}$
Initialization: $g[i] = s[i] = 0,\ i = 0, 1, \ldots, k-1$
Estimation: Apply the EM-Algorithm as described by equations (6) to (10)
while True do
    Compute state vector at $k+j$ by
    $\boldsymbol{\alpha}[k+j] = \mathbf{A}^{j-1}\boldsymbol{\alpha}[k+1] + \boldsymbol{w}[k]$    (12a)
    $\boldsymbol{w}[k] = \mathbf{A}^{j-1}\mathbf{H}\boldsymbol{\eta}[k] + \mathbf{A}^{j-2}\mathbf{H}\boldsymbol{\eta}[k+1] + \cdots + \mathbf{H}\boldsymbol{\eta}[k+j-1]$    (12b)
    Compute the $j$-step ahead prediction using (12) as follows
    $\hat{\mathbf{x}}[k+j|k] = \mathbf{C}\mathbf{A}^j\hat{\boldsymbol{\alpha}}[k|k]$    (15)
    The novel score function $S[k]$ is obtained by
    $S(\hat{\mathbf{x}}[k+j|k]) = a_1\hat{\mathbf{x}}[k+j|k] + a_2\hat{\mathbf{x}}^2[k+j|k] - a_0$    (18)
    The test statistic is computed as
    $g[k+j] = \max(0, g[k-1+j] + S(\hat{\mathbf{x}}[k+j|k]))$    (17)
    if $g[k+j] > \gamma$ then
        $k_{alarm} \leftarrow k+j$ /* possible anomaly early detected */
    else
        $k = k+1$ /* continue searching for anomalies */
```

Fig. 6. The algorithm of the EDA4GNeT method

The considered threat model consists in a DoS attack resulting from a GOOSE poisoning attack. This attack is possible through masquerading a legitimate IED to send malicious packets as decribed in Section II-A.

## IV. Use Case

### A. Description of the use case

A synthesized dataset generated by the Advanced Digital Sciences Center (ADSC) [24] is adopted in the present work.

The simulated testbed, shown in Fig. 7 describes the operation of a 66/11 kV electrical substation model established according to recommendations described in IEC 61850. In fact, protocols from IEC 61850 standard are used including a Generic Object Oriented Substation Event (GOOSE) and Sampled Values (SV) communication via Ethernet VLAN between current transformers, voltage transformers and Intelligent Electronic Devices (IEDs). The MMS protocol at the station level is used for the connection between Human-Machine Interfaces (HMIs) and IEDs. A total of 18 IEDs, shown in Fig. 7, are included within the same multicast group.

Some of the datasets proposed in [24] consist of normal operation scenarios in substations that include disturbances such as a breaker failure or a busbar protection which are particularly relevant for our study as they include presence of specific GOOSE messages which are sent to address disturbance event changes.

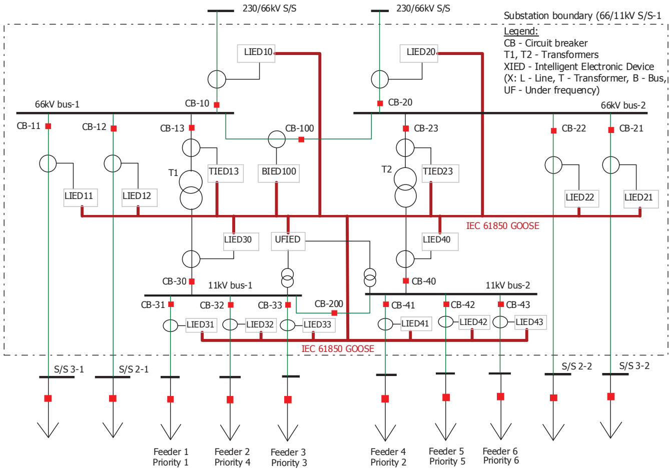

Fig. 7. The single-line diagram of a 66/11 kV substation [24]

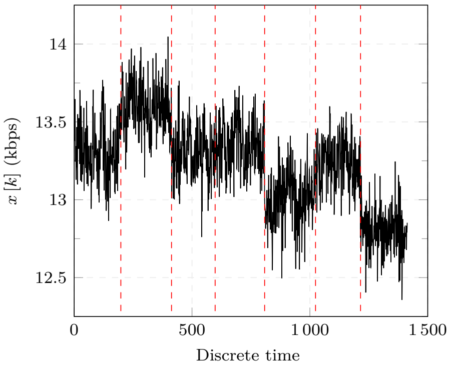

Fig. 8. Simulated GOOSE network traffic with several changes

The normal operation in an electrical substation might also include disturbances such as a breaker failure or a busbar protection where specific GOOSE messages are sent to address such event changes.

In order to thoroughly test the detection performance of the EDA4GNeT method, several experiments with different noise realizations are performed. The considered case study includes several DoS attacks with different characteristics including the duration and the amplitude of the changes. The aforementioned use case is depicted in Fig. 8.

A normal GOOSE network traffic is simulated based on the modeling procedure described in Section II. Each attack

TABLE II
COMPARISON OF DETECTION RESULTS USING EDA4GNET

| Change | Earliness of Detection* | | | Basic | | | | | | Composite | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | FPR [%] | | | FNR [%] | | | $C_{exp}$ | | | $C_{ID}$ | | |
| | [19] | [17] | EDA4GNeT | [19] | [17] | EDA4GNeT | [19] | [17] | EDA4GNeT | [19] | [17] | EDA4GNeT | [19] | [17] | EDA4GNeT |
| 1 | – | – | −29.51 | 0.7810 | 2.4796 | 0.4430 | 0.2983 | 0.9471 | 0.1692 | 0.1023 | **0.1013** | 0.0970 | 0.2985 | 0.1802 | 0.3034 |
| 2 | – | – | −30.24 | 1.3017 | 1.5143 | 1.0970 | 0.4972 | 0.5784 | 0.4190 | 0.1025 | 0.1032 | 0.0970 | **0.2995** | 0.1805 | **0.3212** |
| 3 | – | – | −30.88 | 1.5130 | 1.9984 | 0.9752 | 0.5779 | 0.7633 | 0.3725 | 0.1023 | 0.1029 | 0.0980 | 0.2890 | 0.1825 | 0.3102 |
| 4 | – | – | −30.52 | 0.7917 | 0.7029 | 0.7072 | 0.3024 | 0.2685 | 0.2701 | 0.1022 | 0.1030 | 0.0980 | 0.2890 | **0.1900** | 0.3211 |
| 5 | – | – | −30.43 | 0.9635 | 1.0282 | 0.6253 | 0.3680 | 0.3927 | 0.2389 | 0.1023 | 0.1050 | **0.0970** | 0.2901 | 0.1769 | 0.3103 |
| 6 | – | – | −29.97 | 0.9886 | 0.7047 | 0.4576 | 0.3776 | 0.2692 | 0.1748 | **0.1021** | 0.1052 | 0.0973 | 0.2973 | 0.1675 | 0.3051 |

consists of two changing levels and consequently six changes are represented in Fig. 8. While the first attack starts at $k = 198$ and lasts 215 samples, the second one starts at $k = 413$ and can hardly be distinguished from the normal traffic. The third and last attack starts at $k = 1024$ and finishes at $k = 1215$. It is worth noting that although the changes representing some of the attacks can be perceived in Fig. 8, it is difficult to guess with the naked eye the starting time. It can be also remarked that a change in the dynamics of the system was introduced from the sample $k = 809$. In electrical substations, such changes can occur in case of modification in the physical system that require adaption of the operating conditions.

### B. Performance metrics

In the present section, common as well as advanced performance criteria are introduced.

If $\mathcal{H}_1$, the hypothesis indicating the presence of an anomaly, is selected by the anomaly detection method when $\mathcal{H}_0$, indicating the absence of an anomaly, is true, a False Alarm (FA) or a false positive (FP) are raised. A false negative (FN) (i.e. miss) indicates that $\mathcal{H}_0$ is selected by the detection mechanism when the hypothesis $\mathcal{H}_1$ is true.

True positives (TPs) or a hit refer to the fact that the detector decides correctly for $\mathcal{H}_1$ whereas a true negative (TN) (i.e. correct rejection) indicates that $\mathcal{H}_1$ is correctly discarded.

The following basic performance assessment metrics are defined based on the aforementioned concepts:

When considering a case where the cost of FNs is high, the TPR also called DR or recall, helps give additional information about the detection. The TNR, also called specificity represents the proportion of correctly classified normal samples by the total number of samples of the entire dataset. The FPR represents the rate of FAs which is also referred to as type I errors in statistics. FNR or miss rate represents the proportion of type II error, i.e the hypothesis $\mathcal{H}_0$ is chosen when $\mathcal{H}_1$ is true.

The previously introduced metrics can be combined with other criteria to deduce advanced performance measures.

Combining basic metrics with additional criteria allow obtaining composite detection metrics [41] that can give an advanced evaluation of the performance of IDSs. It is, however, worth noting that the composite metrics do not replace the basic ones [42].

The first advanced metric $C_{exp}$ is defined as follows:

$$\begin{aligned} C_{exp} = min(C \cdot FNR \cdot \boldsymbol{B}, TNR \cdot (1 - \boldsymbol{B})) \\ + min(C \cdot TPR \cdot \boldsymbol{B}, TPR \cdot (1 - \boldsymbol{B})) \end{aligned} \tag{8}$$

Where $C$ is a user-defined parameter representing the ratio of the cost of a misdetected intrusion by the cost of an IDS generating an alert when an intrusion has not occurred.

The base rate $\boldsymbol{B}$ represents the probability that there is an intrusion in the considered dataset. For our experiments, a base rate of $\boldsymbol{B} = 0.1$ is assumed whereas the value of $C$ is set to be equal to 10 following [41]. Indeed, a misdetection of an anomaly within the communication of an electrical substation might result in an increased damage to the overall grid. Thus, the cost of a False Negative Rate (FNR) is considered to be much higher than the cost of False Positive Rate (FPR) as represented by $C$.

According to [41], a low value $C_{exp}$ indicates a better performance of the IDS which provides a practical way to relate the different basic detection metrics. It is, nevertheless, worth mentioning that this performance metric depends on the value of $C$ which is a subjective measure that can be challenging to set as it might depend on different factors including for instance the size and the location of the substation.

The second composite metric considered in the present work, namely the intrusion detection capability $C_{ID}$, was originally introduced in [42]. The $C_{ID}$ metric can provide a more objective evaluation of IDSs as it presents the ratio of the mutual information between the IDS input and output to the entropy of the input. Its main purpose is to have less uncertainty with respect to the input given the Intrusion Detection System (IDS) output. The $C_{ID}$ represents the fraction of correct guesses of an IDS and it is computed as the ratio between the correct guesses of alerts generated by the IDS by the total number of required binary guesses. In order to compare the performance of IDSs, the maximum value of intrusion capabilities $C_{ID}$ obtained for each method shall be compared between them.

### C. Results and discussion

The hardware setup for the performed experiments includes a computer equipped with an Intel processor i7-2.00 GHz and 32GB RAM. A total number of 25 Monte-Carlo simulations are performed for each threshold and under different realizations of WGN for each experiment.

In order to test the performance of the novel EDA4GNeT method, it is compared to the closest works on anomaly detection for IEC 61850 substations available in the literature based on Table I. In fact, methods in [7], [17], [19] are anomaly detection methods that meet at most the predefined requisites namely the inclusion of the specific features of the network traffic, the robustness against network variations and the detection of DOS attacks in the GOOSE as shown in Table I. As the novel method is based on a previous work developed in [7], the two closest counterparts to which we will compare our method are [19] and [17]. Additionally, the ARFIMA model proposed in [7] corresponds to an accurate representation of the network in IEC 61850, however it cannot provide an early detection of advanced DoS attacks since $j$-step ahead predictions cannot be computed from the proposed model. Consequently, to the best of our knowledge and as shown in Table I, none of the currently available anomaly detection methods based on a mathematical model accounting for the dynamics of the network traffic, are able to offer an early detection of DoS attacks resulting from GOOSE poisoning attacks.

In [19] and [17], statistical anomaly detection methods against attacks in IEC 61850 substations, are developed. The authors in [19], present a statistical detection based on a comparison of the residuals with a variance-based threshold while assuming that the network traffic in electrical substations can be represented as a DC level in white Gaussian noise model. In contrast to [19], where the residuals are obtained from a DC level embedded in WGN, a modified version with the residuals computed from the appropriate mathematical model developed in Section II-C are considered instead.

The approach presented in [17] is based on an anomaly score resulting from the comparison of the network traffic with a minimum and a maximum value extracted from real measurements. No details were provided in [17] for the choice of the user-defined parameters necessary for the anomaly detection score. Thus, empirically adapted values are chosen that would allow a high detection performance for the considered use case. Both approaches are implemented for comparison with EDA4GNeT method.

To validate the performance of the EDA4GNeT method, an average of the results of the different Monte-Carlo experiments is presented in Table II. The EDA4GNeT method offers an early detection of attacks of an average of 30 samples ahead according to Table II.

As shown in Table II, on average the False Positive Rate (FPR) of EDA4GNeT is less than 1 % for most of the changes. The False Positive Rate (FPR) of the detection algorithm presented in [17] and [19] are, in general, higher than our method.

The values of the FNR range between $0.17\,\%$ and $0.42\,\%$ for EDA4GNET. The developed method outperforms both counterparts in almost all the changes with the exception of the forth case which might be due to numerical precision. Due to the limitations of the model proposed in [19] for the considered case study, high FPR and FNR are obtained.

In Table II, the lowest values of $C_{exp}$ are depicted in bold and according to [41], the detection method with the lowest cost metric has the best performance. In fact, the results of the composite metrics $C_{exp}$ and $C_{ID}$, are consistent with the basic metrics as, for instance, they also reflect a slightly smaller cost $C_{exp}$ for the EDA4GNeT method.

As shown in bold in Table II for the method developed in [17], the lowest value of $C_{exp}$ is equal to $0.1013$, whereas for EDA4GNeT it corresponds to $0.097$. The values of the intrusion capability $C_{ID}$ of all the changes are higher for EDA4GNeT than for both counterparts.

The reason of the better detection performance of EDA4GNeT can be explained by the adequacy of the selected model for the description of the network traffic in IEC 61850 substations as well as the accuracy of our detector introduced in Section III-C.

In Table II, the earliness of detection represents the number of discrete time samples after which the anomaly is detected. Regardless of the amplitude and duration of the change i.e. the start or the end of the attack, EDA4GNeT is able to detect them in average $30$ samples in advance with an approximate detection rate of $99\,\%$. Compared with its counterparts [17], [19], EDA4GNeT offers a good compromise between the early detection feature and other detection performance statistics. It is worth mentioning that the performance of the EDA4GNeT method is evaluated considering the GOOSE protocol's timing requirements as the main focus is to develop a detection method able to deliver time-ahead alarms with satisfactory detection performance. However, considerations about network latency and packet losses as presented in [14], that might impact the performance of the EDA4GNeT method, will be considered for future work to yield an even more robust detection method.

## V. Conclusion

The security of critical infrastructures including electrical substations is a major concern that has been gaining increasing interest within the research community. In the present work, we tackle the challenge of enhancing the availability of the data in networks of IEC 61850 substations. Thus, the Early Detection of Attacks for GOOSE Network Traffic (EDA4GNeT) method is developed in order to detect DoS resulting from GOOSE poisoning attacks.

The novel anomaly detection system EDA4GNeT addresses limitations of available methods and achieves remarkable results in terms of a balance between the detection performance and earliness of detection. On one hand, its recursive implementation helps account for dynamic changes in the traffic and on the other hand, a robust statistical method based on a novel detection test introduced in Section III-C allows accurate detection of attacks.

To validate the performance of EDA4GNeT, we analyze a use case of the network in a $66/11\,$kV substations including the simulation of different attacks and we use basic and composite performance metrics are used for the evaluation of the method. Comparing the early detection method EDA4GNeT to the related works shows a superior detection performance with a detection rate of more than $99\,\%$ and a false positive rate of no more than around $1.1\,\%$ together with an average early detection of 30 samples ahead.

## Acknowledgement

A special thanks to the anonymous reviewers for their insightful comments and pertinent feedback. This work was supported by the Energy System Design (ESD) project and

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2023.3272749

by funding from the topic Engineering Secure Systems of the Helmholtz Association (HGF) and by KASTEL Security Research Labs (structure 46.23.02).

## APPENDIX A
## AR STATE-SPACE APPROXIMATION

The $AR(m)$ approximation can be presented as follows:

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & \cdots & a_m \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{9a}$$

$$\mathbf{C} = (1, 0, \dots, 0), \tag{9b}$$

$$\mathbf{D} = 0, \tag{9c}$$

$$\mathbf{H} = (1 \; 0 \cdots 0)^T \tag{9d}$$

where the parameters $a_i$ for $j = 1, ..., m$ are computed according to [28]. Thus, expressing an ARFIMA $(p, d, q)$ model using a truncated infinite AR expansion yields

$$\mathbf{x}[k] = \sum_{j=1}^{\infty} a_j \mathbf{x}[k-j] + \mathbf{e}[k] \tag{10}$$

## APPENDIX B
## ESTIMATION OF THE LOG-LIKELIHOOD

The estimated log-likelihood is expressed by:

$$\begin{aligned} G(\Theta) =& E\left(\log L | \mathbf{x}[1], \dots, \mathbf{x}[N]\right) \\ =& -\frac{1}{2}\log|\boldsymbol{\Sigma}| \\ & -\frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\Sigma}^{-1}(\mathbf{P}[0|N] + (\boldsymbol{\alpha}[0|N] - \boldsymbol{\mu})(\boldsymbol{\alpha}[0|N] - \boldsymbol{\mu})^{\mathrm{T}})\right] \\ & -\frac{N}{2}\log|\mathbf{Q}| - \frac{1}{2}\operatorname{Tr}\left[\mathbf{Q}^{-1}(\mathbf{F} - \mathbf{E}\mathbf{A}^{\mathrm{T}} - \mathbf{A}\mathbf{E}^{\mathrm{T}} + \mathbf{A}\mathbf{D}\mathbf{A}^{\mathrm{T}})\right] \\ & -\frac{N}{2}\log|\mathbf{R}| - \frac{1}{2}\operatorname{Tr}\Bigg[\mathbf{R}^{-1}\sum_{k=1}^{N}((\mathbf{x}[k] - \mathbf{C}\boldsymbol{\alpha}[k|N]) \\ & \quad (\mathbf{x}[k] - \mathbf{C}\boldsymbol{\alpha}[k|N])^{\mathrm{T}} + \mathbf{C}\mathbf{P}[k|N]\mathbf{C}^{\mathrm{T}}\Big)\Bigg] \end{aligned} \tag{11}$$

The maximization of $G(\Theta)$ is obtained with the following computations:

$$\mathbf{A}^{r+1} = \mathbf{E}\mathbf{D}^{-1} \tag{12a}$$

$$\mathbf{Q}^{r+1} = \frac{1}{N}\left(\mathbf{F} - \mathbf{E}\mathbf{D}^{-1}\mathbf{E}^{\mathrm{T}}\right) \tag{12b}$$

$$\begin{aligned} \mathbf{R}^{r+1} =& \frac{1}{N}\sum_{k=1}^{N}\Big((\mathbf{x}[k] - \mathbf{C}\boldsymbol{\alpha}[k|N])(\mathbf{x}[k] - \mathbf{C}\boldsymbol{\alpha}[k|N])^{\mathrm{T}} \\ & + \mathbf{C}\mathbf{P}[k|N]\mathbf{C}^{\mathrm{T}}\Big) \end{aligned} \tag{12c}$$

where,

$$\mathbf{D} = \sum_{k=1}^{N}\left(\mathbf{P}[k-1|N] + \boldsymbol{\alpha}[k-1|N]\boldsymbol{\alpha}[k-1|N]^{\mathrm{T}}\right) \tag{13a}$$

$$\mathbf{E} = \sum_{k=1}^{N}\left(\mathbf{P}[k|N] + \boldsymbol{\alpha}[k|N]\boldsymbol{\alpha}[k-1|N]^{\mathrm{T}}\right) \tag{13b}$$

$$\mathbf{F} = \sum_{k=1}^{N}\left(\mathbf{P}[k|N] + \boldsymbol{\alpha}[k|N]\boldsymbol{\alpha}[k|N]^{\mathrm{T}}\right) \tag{13c}$$

The quantities required in Eq. (13) are computed using the Kalman smoother. The Kalman smoother allows a recursive state estimation to compute the posterior distribution over the latent states of a linear state space model given some observed data. The Kalman smoother is proposed for state estimation based on the values of the signal $x[k]$. The equations for the Kalman smoother are expressed as following for $k = n, n-1, \dots, 1$:

$$\mathbf{J}[k-1] = \mathbf{P}[k-1|k-1]\mathbf{A}^{\mathrm{T}}(\mathbf{P}[k|k-1])^{-1} \tag{14a}$$

$$\begin{aligned} \boldsymbol{\alpha}[k-1|n] =& \boldsymbol{\alpha}[k-1|k-1] \\ & + \mathbf{J}[k-1](\boldsymbol{\alpha}[k|n] - \mathbf{A}\boldsymbol{\alpha}[k-1|k-1]) \end{aligned} \tag{14b}$$

$$\begin{aligned} \mathbf{P}[k-1|n] =& \mathbf{P}[k-1|k-1] + \mathbf{J}[k-1] \\ & (\mathbf{P}[k|n] - \mathbf{P}[k|k-1])\mathbf{J}[k-1]^{\mathrm{T}} \end{aligned} \tag{14c}$$

The initial values for the smoother are the final estimates of the filter. At each iteration, the rules in Eq. (12) are computed using Eq. (13).

## APPENDIX C
## COMPUTATION OF THE J-STEP AHEAD PREDICTION

The state vector at $k + j$ can be written as follows [43]:

$$\boldsymbol{\alpha}[k+j] = \mathbf{A}^{j-1}\boldsymbol{\alpha}[k+1] + \boldsymbol{w}[k] \tag{15a}$$

$$\begin{aligned} \boldsymbol{w}[k] =& \mathbf{A}^{j-1}\mathbf{H}\boldsymbol{\eta}[k] + \mathbf{A}^{j-2}\mathbf{H}\boldsymbol{\eta}[k+1] + \cdots \\ & + \mathbf{H}\boldsymbol{\eta}[k+j-1] \end{aligned} \tag{15b}$$

The general expression of the j-step ahead forecast with $j > 1$ of the state vector $\hat{\boldsymbol{\alpha}}[k+j|k]$ is expressed from the conditional expectation of Eq. (15)

$$\hat{\boldsymbol{\alpha}}[k+j|k] = \mathbf{A}^{j-1}\hat{\boldsymbol{\alpha}}[k+1|k] \tag{16}$$

The error of the forecast of the state vector can be calculated as follows:

$$\boldsymbol{\alpha}[k+j] - \hat{\boldsymbol{\alpha}}[k+j|k] = \boldsymbol{\alpha}[k+j] - \mathbf{A}^{j}\hat{\boldsymbol{\alpha}}[k|k] \tag{17}$$

The previously introduced equation Eq. (16) can be used to describe the $j$-step ahead forecasts of the observation vector $x[k+j]$.

$$\hat{\mathbf{x}}[k+j|k] = \mathbf{C}\mathbf{A}^{j}\hat{\boldsymbol{\alpha}}[k|k] \tag{18}$$

The error of the forecast calculated in Eq. (18) is:

$$\mathbf{x}[k+j] - \hat{\mathbf{x}}[k+j|k] = \mathbf{x}[k+j] - \mathbf{C}\mathbf{A}^{j}\hat{\boldsymbol{\alpha}}[k|k] \tag{19}$$

## References

[1] G. Elbez, H. B. Keller, and V. Hagenmeyer, "Authentication of GOOSE messages under timing constraints in IEC 61850 substations," in *6th International Symposium for ICS & SCADA Cyber Security Research 2019 6*, 2019, pp. 137–143.

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

[3] G. Elbez, H. B. Keller, and V. Hagenmeyer, "A new classification of attacks against the cyber-physical security of smart grids," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*, 2018, pp. 1–6.

[4] H. Yoo and T. Shon, "Challenges and research directions for heterogeneous cyber–physical system based on IEC 61850: Vulnerabilities, security requirements, and security architecture," *Future generation computer systems*, vol. 61, pp. 128–136, 2016.

[5] H. B. Keller, O. Schneider, J. Matthes, and V. Hagenmeyer, "Reliable, safe and secure software of connected future control systems-challenges and solutions," *at - Automatisierungstechnik*, vol. 64, no. 12, pp. 930–947, 2016, ISSN: 0178-2312, 2196-677X. DOI: 10.1515/auto-2016-0060.

[6] M. S. Mahmoud, H. M. Khalid, and M. M. Hamdan, *Cyberphysical Infrastructures in Power Systems: Architectures and Vulnerabilities*. Academic Press, 2021.

[7] G. Elbez, H. B. Keller, A. Bohara, K. Nahrstedt, and V. Hagenmeyer, "Detection of DoS attacks using ARFIMA modeling of GOOSE communication in IEC 61850 substations," *Energies*, vol. 13, no. 19, p. 5176, 2020.

[8] S. Cheung, B. Dutertre, M. Fong, U. Lindqvist, K. Skinner, and A. Valdes, "Using model-based intrusion detection for SCADA networks," in *Proceedings of the SCADA security scientific symposium*, Citeseer, vol. 46, 2007, pp. 1–12.

[9] U. K. Premaratne, J. Samarabandu, T. S. Sidhu, R. Beresh, and J.-C. Tan, "An intrusion detection system for IEC61850 automated substations," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2376–2383, 2010.

[10] T. Morris, R. Vaughn, and Y. Dandass, "A retrofit network intrusion detection system for modbus rtu and ascii industrial control systems," in *2012 45th Hawaii International Conference on System Sciences*, IEEE, 2012, pp. 2338–2345.

[11] H. Lin, A. Slagell, C. Di Martino, Z. Kalbarczyk, and R. K. Iyer, "Adapting bro into SCADA: Building a specification-based intrusion detection system for the dnp3 protocol," in *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop*, 2013, pp. 1–4.

[12] 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, 2016.

[13] M. Kabir-Querrec, S. Mocanu, J.-M. Thiriet, and E. Savary, "Power utility automation cybersecurity: IEC 61850 specification of an intrusion detection function," in *25th European Safety and Reliability Conference (ESREL 2015)*, CRC Press, 2015.

[14] A. Bohara, J. Ros-Giralt, G. Elbez, A. Valdes, K. Nahrstedt, and W. H. Sanders, "Ed4gap: Efficient detection for GOOSE-based poisoning attacks on IEC 61850 substations," in *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, IEEE, 2020, pp. 1–7.

[15] U. Inayat, M. F. Zia, S. Mahmood, H. M. Khalid, and M. Benbouzid, "Learning-based methods for cyber attacks detection in iot systems: A survey on methods, analysis, and future prospects," *Electronics*, vol. 11, no. 9, p. 1502, 2022.

[16] C. Kwon, W. Liu, and I. Hwang, "Security analysis for cyber-physical systems against stealthy deception attacks," in *2013 American control conference*, IEEE, 2013, pp. 3344–3349.

[17] Q. Yang, W. Hao, L. Ge, W. Ruan, and F. Chi, "Farima model-based communication traffic anomaly detection in intelligent electric power substations," *IET Cyber-Physical Systems: Theory & Applications*, vol. 4, no. 1, pp. 22–29, 2019.

[18] W. Shang, L. Li, M. Wan, and P. Zeng, "Industrial communication intrusion detection algorithm based on improved one-class SVM," in *2015 World Congress on Industrial Control Systems Security (WCICSS)*, IEEE, 2015, pp. 21–25.

[19] Y. Kwon, H. K. Kim, Y. H. Lim, and J. I. Lim, "A behavior-based intrusion detection technique for smart grid infrastructure," in *2015 IEEE Eindhoven PowerTech*, IEEE, 2015, pp. 1–6.

[20] M. F. Elrawy, L. Hadjidemetriou, C. Laoudias, and M. K. Michael, "Light-weight and robust network intrusion detection for cyber-attacks in digital substations," in *2021 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia)*, IEEE, 2021, pp. 1–5.

[21] T. S. Ustun, S. Hussain, A. Ulutas, A. Onen, M. M. Roomi, and D. Mashima, "Machine learning-based intrusion detection for achieving cybersecurity in smart grids using iec 61850 goose messages," *Symmetry*, vol. 13, no. 5, p. 826, 2021.

[22] S. Ashraf, M. H. Shawon, H. M. Khalid, and S. Muyeen, "Denial-of-service attack on iec 61850-based substation automation system: A crucial cyber threat towards smart substation pathways," *Sensors*, vol. 21, no. 19, p. 6415, 2021.

[23] A. Turner, *Tcpreplay*, 2003.

[24] P. P. Biswas, H. C. Tan, Q. Zhu, Y. Li, D. Mashima, and B. Chen, "A synthesized dataset for cybersecurity study of IEC 61850 based substation," in *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, IEEE, 2019, pp. 1–7.

[25] M. Strobel, N. Wiedermann, and C. Eckert, "Novel weaknesses in IEC 62351 protected smart grid control systems," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, IEEE, 2016, pp. 266–270.

[26] J. G. Wright and S. D. Wolthusen, "Stealthy injection attacks against IEC61850's GOOSE messaging service," in *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, IEEE, 2018, pp. 1–6.

[27] J. Hosking, "Fractional differencing modeling in hydrology," *JAWRA Journal of the American Water Resources Association*, vol. 21, no. 4, pp. 677–682, 1985.

[28] S. Grassi and P. S. De Magistris, "When long memory meets the Kalman filter: A comparative study," *Computational Statistics & Data Analysis*, vol. 76, pp. 301–319, 2014.

[29] W. Palma, *Long-memory time series: theory and methods*. John Wiley & Sons, 2007, vol. 662.

[30] N. H. Chan and W. Palma, "State space modeling of long-memory processes," *Annals of Statistics*, pp. 719–740, 1998.

[31] N.-J. Hsu, B. K. Ray, and F. BREIDT, "Bayesian estimation of common long-range dependent models," in *Proc. of Seventh Vilnius Conference on Probability Theory and Mathematical Statistics*, 1999, pp. 311–324.

[32] G. Mesters, S. J. Koopman, and M. Ooms, "Monte carlo maximum likelihood estimation for generalized long-memory time series models," 2011.

[33] S. Floyd and V. Paxson, "Difficulties in simulating the internet," *IEEE/ACm Transactions on Networking*, vol. 9, no. 4, pp. 392–403, 2001.

[34] D. Bauer, "Using subspace methods to model long-memory processes," in *International Conference on Time Series and Forecasting*, Springer, 2018, pp. 171–185.

[35] V. Digalakis, J. R. Rohlicek, and M. Ostendorf, "Ml estimation of a stochastic linear system with the em algorithm and its application to speech recognition," *IEEE Transactions on speech and audio processing*, vol. 1, no. 4, pp. 431–442, 1993.

[36] R. R. R. Barbosa, "Anomaly detection in SCADA systems: A network based approach," 2014.

[37] C. W. Granger and R. Joyeux, "An introduction to long-memory time series models and fractional differencing," *Journal of time series analysis*, vol. 1, no. 1, pp. 15–29, 1980.

[38] J. R. Hosking, "Modeling persistence in hydrological time series using fractional differencing," *Water resources research*, vol. 20, no. 12, pp. 1898–1908, 1984.

[39] M. Basseville, I. V. Nikiforov, *et al.*, *Detection of abrupt changes: theory and application*. prentice Hall Englewood Cliffs, 1993, vol. 104.

[40] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, 2014.

[41] A. Milenkoski, M. Vieira, S. Kounev, A. Avritzer, and B. D. Payne, "Evaluating computer intrusion detection systems: A survey of common practices," *ACM Computing Surveys (CSUR)*, vol. 48, no. 1, pp. 1–41, 2015.

[42] G. Gu, P. Fogla, D. Dagon, W. Lee, and B. Skoric, "Measuring intrusion detection capability: An information-theoretic approach," in *Proceedings of the 2006 ACM Symposium on Information, computer and communications security*, 2006, pp. 90–101.

[43] J. G. de Gooijer and A. Klein, "On the cumulated multi-step-ahead predictions of vector autoregressive moving average processes," *International Journal of Forecasting*, vol. 7, no. 4, pp. 501–513, 1992.

**Ghada Elbez** is head of the Secure Energy Systems (SES) research group at Karlsruhe Institute of Technology (KIT), Germany and coordinator of the KASTEL Security Lab Energy at the Institute for Automation and Applied Informatics (IAI). She received her M.Sc. degree in Electrical Engineering and Computer Science in 2016 from the Ecole Polytechnique de Lille, France and her PhD in 2022 from the Informatics Faculty at KIT. Her research interests include cyber-security of energy systems, defense mechanisms for early detection of attacks and security standards. She serves as reviewer and chair in different IEEE and ACM conferences. She actively contributes to some German standardization groups within VDI/VDE on the topics of cyber-security in energy systems.

**Klara Nahrstedt** is the Grainger Distinguished Chair in Engineering Professor in Computer Science Department, and Director of Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. Her research interests are directed towards multi-modal distributed systems, Quality of Service management, and trusted cyber-physical smart grid systems. She is the recipient of the IEEE Computer Society Technical Achievement Award, ACM SIGMM Technical Achievement Award, and she is a Fellow of IEEE, ACM, AAAS. She is a member of the German Academy of Sciences (Leopoldina Society), and USA National Academy of Engineering. Klara Nahrstedt received her Diploma degree in mathematics and numerical analysis from Humboldt University, Berlin, in 1985. In 1995 she received her PhD from the University of Pennsylvania in the Department of Computer and Information Science.

**Veit Hagenmeyer** is currently the Professor of Energy Informatics with the Faculty of Informatics, and the Director of the Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Karlsruhe, Germany. His research interests include modeling, optimization and control of sector integrated energy systems, machine learning based forecasting of uncertain demand and production in energy systems mainly driven by renewables, and integrated cyber-security of such systems.