Cyber Physical Security Analytics for Anomalies in Transmission Protection Systems

Arman Ahmed, *Student Member, IEEE*, Vignesh V. G. Krishnan, *Member, IEEE*, Seyedeh Armina Foroutan, *Student Member, IEEE*, Md. Touhiduzzaman, *Student Member, IEEE*, Caroline Rublein, *Student Member, IEEE*, Anurag Srivastava, *Senior Member, IEEE*, Yinghui Wu, *Senior Member, IEEE*, Adam Hahn, *Member, IEEE*, and Sindhu Suresh, *Senior Member, IEEE*

***Abstract*—Protection systems are one of the most critical components in the transmission system and are becoming more digital with ongoing automation. These digital systems are prone to vulnerabilities/attacks, and exploitation of these vulnerabilities may cause major impacts on the electric grid performance. Multiple alarms reported in the control center could be a result of the faults (expected operations) or failures in the protection system (anomalies/ unexpected operation). Situational awareness gained through sensors such as a phasor measurement unit (PMU) and data acquired through the cyber system provide an opportunity to develop continuous cyber-physical monitoring of the system. Note that relay data are not reported in the control center continuously. This paper presents a cyber-physical data analytics based technique to monitor transmission protection system and detect malicious activity. Initially, continuous monitoring of PMU data is utilized for data anomaly detection, which includes bad or missing data using long short-term memory (LSTM). Then, PMU data of interest are utilized for failure diagnosis, using a semisupervised deep autoencoder model. In this research, cyber anomalies are modeled by manipulating the setting/logic design of protective devices, and a ridge regression based classifier with a feature engineering pipeline is used to detect cyber anomalies. The results from the deep autoencoder model and ridge regression based classifier are then utilized for detailed investigation to find the root causes of the observed events assisted by the cyber log data from the protection devices. The algorithm is validated using a real-time simulation of the IEEE test system with industrial hardware relays and PMUs in the loop. Data analytics algorithm running on server utilizes these real-time data continuously for anomaly detection and classification for the developed use cases.**

***Index Terms*—Cyber anomalies, cyber-physical systems, cyber security, data analytics, digital protection, transmission protection systems.**

I. Introduction

PROTECTION system is a critical component in the power grid, which isolates the faulty components from the healthy system as quickly as possible. Malfunctions or failures in the protective devices can lead to isolation of some components of the healthy system along with the faulty components. Protection system maloperation has been ranked as the number one concern by the North American Electric Reliability Corporation to cause power blackouts [1]. Diagnosis of the root cause leading to failures in protection system is very important to restore the system to its normal operating condition. A part in protection system is said to maloperate, if it operates unintentionally or outside of its expected zone of protection. Such maloperations could be due to physical failures or cyber-induced reasons. With the increasing automation and digitization in the power grid, there are higher risks of cyber anomaly induced protection system maloperations and failures. Cyber-induced power system blackouts have been recently reported in the literature [2], including the attack on Ukrainian power grid [3]. On December 23, 2015, a group of attackers successfully intruded the Ukrainian substation and created an impact on a large regional distribution area followed by another attack in 2016. These attackers were able to manipulate the grid without the need of sophisticated cyber-physical malware payloads, such as those used in Stuxnet [2].

The impact of the protection system and hidden failures on bulk power system reliability was investigated in the work presented in [4]. A breaker-oriented bulk power system network model has been developed, which includes substation configuration as well as corresponding protection schemes. Yu and Singh [5], [6] have presented the vulnerability analysis of protection system failures. In [7], neural networks have been used to model the uncertainties involved in the relay and breaker operation messages to estimate the faulted section. A decision support system using the circuit breaker information for online fault section estimation in power systems has been presented in [8]. The existing methods for power system fault diagnosis do not systematically address the possible malfunctions or failures of

Manuscript received February 3, 2019; revised April 21, 2019; accepted May 7, 2019. Date of publication July 14, 2019; date of current version October 18, 2019. Paper 2019-IACC-0136.R1, presented at the 2018 IAS Annual Meeting, Portland, OR, USA, Sep. 23–27, and approved for publication in the IEEE Transaction on Industry Applications by the Industrial Automation and Control Committee of the IEEE Industry Applications Society. This work was supported in part by the Siemens and in part by the National Science Foundation award 1840192. *(Corresponding author: Anurag Srivastava.)*

A. Ahmed, S. A. Foroutan, M. Touhiduzzaman, A. Srivastava, Y. Wu, and A. Hahn are with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99163 USA (e-mail: arman.ahmed@wsu.edu; s.foroutan@wsu.edu; md.touhiduzzaman@wsu.edu; asrivast@eecs.wsu.edu; yinghui.wu@wsu.edu; a.hahn@wsu.edu).

V. V. G. Krishnan is with Washington State University, Pullman, WA 99163 USA. He is now with the Indian Institute of Technology Tirupati, Tirupati 517506, India (e-mail: v.venkatagopalakris@wsu.edu).

C. Rublein is with Washington State University, Pullman, WA 99163 USA (e-mail: clr2027@lockhaven.edu).

S. Suresh is with Siemens Corporate Research, Princeton, NJ 08540 USA (e-mail: sindhu-suresh@siemens.com).

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

Digital Object Identifier 10.1109/TIA.2019.2928500

protective devices [7], [9]. When a fault occurs in the system along with protection system failures, conflicting information and alarms makes the problem identification difficult for system engineers/operators. Therefore, an automated algorithm is required for protection system failure diagnosis, in order to precisely identify the malfunctions and failures of protective devices. Phasor measurement units (PMU) send streaming data and provide an opportunity to develop a failure diagnosis method in a dynamic environment. In [10], a real-time tool was presented to detect, classify, and locate transmission line faults, and indicate whether the line was tripped due to a maloperation of protective relays. Synchrophasor measurements were used to prevent zone 3 maloperations in [11]. Temporal causal diagram has been introduced in [12] for the root cause analysis in the power system.

With the increased deployments of the digital measurement devices, such as PMUs on the transmission systems, increasingly growing data have initiated a serious concern [13]. Especially, in a scenario of simultaneous multiple failures happening in a network, because of the increasing number of PMUs, relays, and breakers, it will be difficult to pinpoint the actual event, so a data analytic tool seems to be necessary. Manual failure diagnosis using such data can be more difficult, owing to the large number of data sources and protection devices. Recently, there has been a surge in research dedicated to measurement driven grid operation analysis using big data technologies for smart grid applications [13], [14]. Data analytic based approaches reduce human tuning efforts. As physical sensors, PMUs often contain imperfections that can result in bad or missing data. These incorrect data can be considered noise and can interfere with the analysis of actual event data. Hence, identifying anomalies in power grid data is extremely important for performing accurate analysis. Unfortunately, outlier detection in PMU data is a multifaceted problem. Since bad and missing data can take on any value, finding all anomalies is difficult since some of them could be nearly identical to the expected value. Even when the majority of anomalies fall far from the expected value, individual methods such as Chebyshev and linear regression have been only moderately accurate on their own, as in [15]. Therefore, it is important to develop anomaly detection algorithms to detect and filter any bad data in the measurements, before they are deployed for further applications, as proposed recently by many researchers.

Supervised methods, which include Bayesian methods [16], [17], support vector machines [17]–[20], nearest neighbor methods [17], [18], and AdaBoost [17] are implemented for anomaly detection. However, one major drawback of supervised methods is that they are only useful if the attack is previously observed. Detecting new attacks require the ability to establish the baseline of normal data and detect deviations of the new data from normal data.

In this paper, a data analytic approach using PMUs, breaker status, and streaming data has been developed for the failure diagnosis in transmission protection systems. This paper extends our algorithms in [21] by introducing long short-term memory (LSTM) networks to identify the anomalies in the PMU data stream. LSTM networks use a combination of all previous data to predict future values, with older inputs fading into a smaller proportion to impact the prediction over time. Then, a semisupervised deep learning algorithm "Deep Autoencoder" for detecting anomalies in PMU data is proposed A ridge regression based classifier with a feature engineering pipeline is proposed for detecting unauthorized intrusion on relay using the cyber data, which acts as ground truth for anomalies found in the PMU data. Finally, the outcomes from the analysis of physical data and cyber data are aggregated for further investigation and decision. The algorithms are designed to run in the control center after a fault has occurred in the presence of the malfunctioning protective system and specifying the reasons for the failure (physical induced or cyber attack induced). The final root cause is determined using the results from the data analytic algorithm, which is further validated using relay log files. The main contributions of this paper are as follows.

1) To propose LSTM networks to identify bad data in the PMU data stream.
2) To implement cyber attack models to simulate/emulate protection system maloperation due to cyber attacks.
3) To propose a semisupervised deep learning algorithm for detecting an anomaly in the PMU data.
4) To propose a ridge regression based classifier with a feature engineering pipeline for detecting an intrusion in digital relay.
5) To validate the algorithm on a real-time testbed incorporating with real-time digital simulator (RTDS), hardware relays, and proposed cyber attack models.

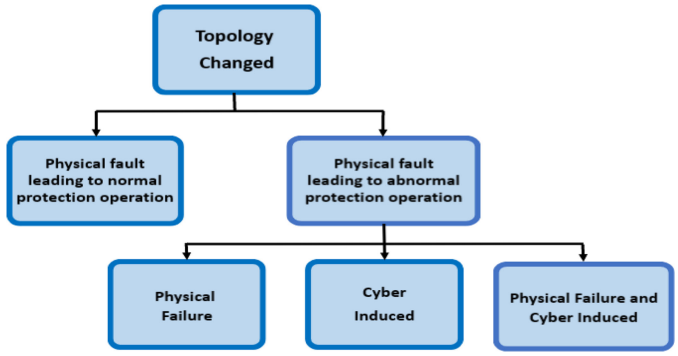

Fig. 1. Various possible failure cases in protection systems.

II. Failure in Protection Systems

The possible cases in protection system operations are shown in Fig. 1. These anomalies can change the topology of the transmission grid. Following sections briefly describe the various modes of protection system operations.

A. Case A: Physical Faults Leading to Normal Protection System Operation

In this case, the protection systems normally function in response to a fault in the power system. This could be a relay operation in zone 1, 2, or 3. The fault is cleared by opening of the circuit breaker.

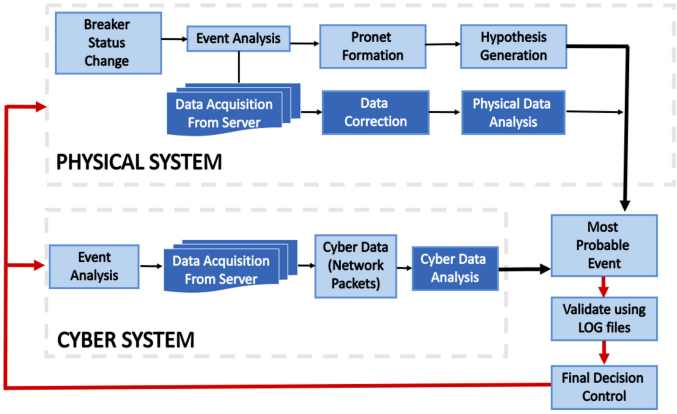

Fig. 2. Schematic of the proposed solution strategy.

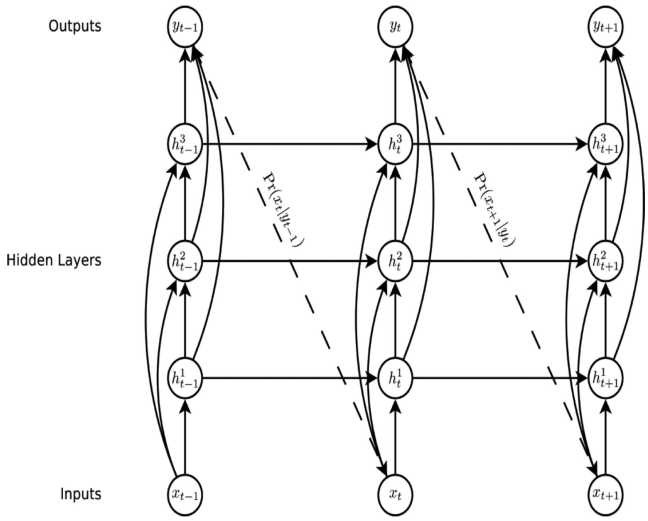

Fig. 3. LSTM layers.

B. Case B: Physical Faults Leading to Abnormal Protection System Operation

In this case, the protective systems do not respond normally to the fault. This could be a failure in the protection system due to following reasons.

1) Physical reasons such as bush failure and breaker mechanism failure [22]. Such malfunctions can usually be easily diagnosed.
2) Cyber attacks leading to the maloperation of the protection system. In this case, the attacker gains access into the substation control and protection network, and issues a malicious control command to operate a circuit breaker. Maloperations due to a cyber attack can be hard to be diagnosed. This may require both physical and cyber data.
3) Protection systems may fail due to both 1) and 2). Out of the maloperation, the attacker can maliciously control few protection systems and the rest can maloperate due to physical reasons. This scenario may need extensive analysis to arrive at a correct reason for the failure.

III. Proposed Solution Approach

Given a large number of protective devices in the power grid, it is usually hard to perform failure diagnosis manually. This (manual procedure) may involve an extensive postfailure analysis using relay log files and other data. Given the streaming physical data (from PMUs and other sensors) and with the advent of state-of-the-art data science techniques, the failure diagnosis can be made efficient. Fig. 2 shows the architecture of the proposed solution and the following sections describe the various steps in the proposed strategy for protection system failure diagnosis.

A. LSTM Network for PMU Anomaly Detection

An LSTM is a type of recurrent neural network composed of LSTM cells. They are well suited for analyzing sequential data, which helps in learning important events that are few and far between, as proved in [23]. In an LSTM cell, sequential nature of inputs is propagated to the adjacent future time steps as shown in Fig. 3. A layer of LSTM cells can be stacked onto another layer of LSTM cells, which further aids the LSTM network to learn the sequential latent structure of the data.

Inside each hidden layer, the LSTM cell contains a forget gate, candidate layer, input gate, and an output gate. The forget, input, and output gates concatenate the dot product of the input vector X_t and a weight vector U with the dot product of the previously hidden state H_{t-1} and another weight vector W; then, the sigmoid activation function is applied to the computed values. Equation (1) of the forget gate and the other gates use similar operations but with different weight vector values. The equation that determines the candidate layer is also identical to (1), but uses the hyperbolic tangent (tanh) as the activation function. The computations in the gates are performed as element-wise operations

$$f_t = \sigma(x) = \sigma(X_t * U_f + H_{t-1} * W_f) \tag{1}$$

where

$\boldsymbol{X}_t$ *input vector*;
$\boldsymbol{U}_f$ *weight vector of input at time step t in the forget gate* f_t;
$\boldsymbol{H}_{t-1}$ *previous hidden state vector;*
$\boldsymbol{W}_f$ *weight vector of previous hidden state in the forget gate* f_t;
$\sigma(\boldsymbol{f}(\boldsymbol{x}))$ *activation function applied element wise at each neuron of the encoder.*

Each of the gates calculates the current memory state C_t and current hidden state H_t according to the following equations:

$$C_t = f_t * C_{t-1} + I_t * \overline{C}_t \tag{2}$$

where

$\boldsymbol{f_t}$ *forget gate (sigmoid function);*
$\boldsymbol{C_{t-1}}$ *previous current memory state vector;*
$\boldsymbol{I_t}$ *input gate (sigmoid function);*
$\overline{\boldsymbol{C_t}}$ *candidate layer* (tanh *function)*

$$H_t = O_t * \tanh(C_t) \tag{3}$$

where

O_t *output gate (sigmoid function);*

C_t *current memory state vector.*

B. Triggering the Data Analytics for Failure Diagnosis

A trigger-based approach is used to initiate the proposed failure diagnosis approach. Continuous monitoring of circuit breaker status is done, which is typically available every few seconds. When the breaker status changes, physical and cyber data are acquired from the database server in the cloud (usually a few seconds after the event of breaker status change). The deep autoencoder algorithm then analyzes the physical data, and the ridge regression based classifier algorithm analyzes the cyber network data.

C. Hypothesis Generation

When a fault occurs in a network, under normal condition, it is expected that both relays at each end of the fault line trip their corresponding breakers. However, if at least one of these relays/breakers mal-operate because of a physical/cyber anomaly, then it is usually not easy to determine the reason of the failure, and several failure reasons can be interpreted based on the status of relays and breakers.

In the proposed failure diagnosis approach, the first step is to create these scenarios and develop Protection Network (ProNet) consisting of all lines and nodes (buses) adjacent to an open breaker. Each line with at least one open breaker can be a candidate for fault location. Based on other breakers' status, whether the current scenario can explain the current topology needs to be analyzed. Doing this by manual investigation can be cumbersome, as there may be many protective devices in the network.

D. Data Analytic Based Decision Making

The possible root causes of the failure can be further simplified with the help of data science techniques. This section describes the anomaly detection using streaming cyber and physical data.

1) Anomaly Detection Over PMU Using Deep Autoencoder: Deep Autoencoder is a feed-forward neural network consisting of three types of layers, namely input layer, hidden layer, and output layer connected in a sequential order, respectively, as shown in Fig. 4. A deep autoencoder comprises two symmetrical deep-belief networks that have multiple hidden layers [24].

The proposed deep autoencoder model, used in this paper, is used to label the PMU readings as normal (0) or as an anomaly (1), and it forms the primary basis for anomaly detection over PMU data. Encoder and decoder are two significant components of deep autoencoder, as shown in Fig. 4. The encoder compresses the input feature vector (input data) x to a compressed code represented as "ϕ," as shown in Fig. 4, whereas the decoder reconstructs the compressed code "ϕ" and outputs a reconstructed feature vector $\hat{x}$, which is of the same shape as the original input feature vector (input data) x. The proposed deep autoencoder model consists of seven stacked fully-connected neural network layers representing a combination of encoder and decoder. The proposed autoencoder architecture consists of seven layers consisting of 12, 10, 5, 2, 5, 10, and 12 number of neurons, respectively, as shown in Fig. 4.

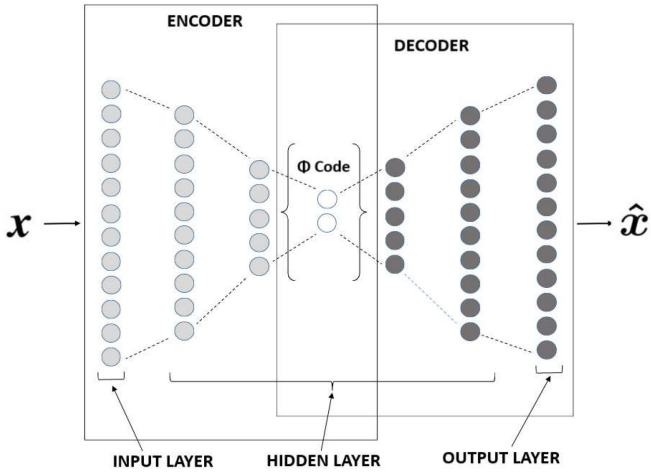

Fig. 4. Proposed architecture of the autoencoder model.

The output "ϕ" of the encoder can be represented as

$$\phi = \sigma(f(\boldsymbol{x})) = \sigma(\boldsymbol{W} * \boldsymbol{x} + \boldsymbol{b}) \tag{4}$$

where

$\boldsymbol{b}$ *bias vector for encoder;*

$\boldsymbol{W}$ *weight matrix for encoder;*

$\boldsymbol{x}$ *input feature vector for encoder;*

$\boldsymbol{\sigma}(f(x))$ *activation function applied element wise at each neuron of the encoder.*

The output $\hat{x}$ of the decoder can be represented as

$$\hat{\boldsymbol{x}} = \hat{\sigma}(g(\phi)) = \hat{\sigma}(\hat{\boldsymbol{W}} * \phi + \mathbf{c}) \tag{5}$$

where

$\boldsymbol{c}$ *bias vector for decoder;*

$\boldsymbol{W}$ *weight matrix for decoder;*

$\hat{\boldsymbol{\sigma}}(\boldsymbol{g}(\boldsymbol{\phi}))$ *activation function applied element wise at each neuron of the decoder;*

ϕ *compressed output of the encoder, which act as an input feature vector for the decoder.*

All neurons in the proposed architecture of deep autoencoder apply rectified linear unit [25] activation function to their respective inputs, except those neurons in the output layer as they apply "Sigmoid" [26] activation function to their respective inputs. Key points for deep autoencoder can be defined as follows.

1) Input feature selection for training Deep Autoencoder: Selection of input features plays an imperative role during training a machine learning algorithm. However, in deep neural network, the hidden layers automatically learn the underlying latent structure of the data. The proposed deep autoencoder model is provided with 12 input features of PMU data, which include 3-phase voltage and current phasor readings of PMU except the true label, which indicates that a particular PMU reading in the dataset is an anomaly (1) or not (0) as it is a semisupervised approach proposed in this research.

2) Normalizing selected input features: PMU data consist of voltage, current values that are of varying magnitude (scale). Algorithms such as neural network, which impose weights to input values, do not perform well with unnormalized (varying scale) input values. Hence, in order to optimize the performance of the proposed deep autoencoder model, the values in the input feature vector are normalized to the mean of zero and standard deviation of one, which helps the algorithm to learn faster during back propagation.
3) Training of Deep Autoencoder: During the process of training, in the proposed deep autoencoder model, a purity-based approach is used, i.e., only normal PMU measurements are used to train the model and the parameters ($\boldsymbol{W}, \boldsymbol{b}, \hat{\boldsymbol{W}}, \boldsymbol{c}$) are evaluated, optimized, and updated during back-propagation process using a stochastic optimization algorithm known as adaptive moment estimation (ADAM) [27]. ADAM optimizer [27] minimizes the reconstruction error loss between x and $\hat{x}$, which is evaluated using the *mean squared error* (MSE) given as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x} - \hat{\boldsymbol{x}}_i)^2. \tag{6}$$

The proposed deep autoencoder model is built using state-of-the-art Python deep learning libraries "Keras" and "Tensorflow." While building the proposed architecture early stopping is also implemented to prevent overfitting. The additional hyper parameters of the model such as learning rate and number of epochs for training the deep autoencoder model are set beforehand. Various values of hyper parameters are used to train the model to obtain the best performance, and the hyper parameters that give the best performance are chosen for training the deep autoencoder.

2) Intrusion Detection on Relay Using Ridge Classifier With a Feature Engineering Pipeline: Analyzing network traffic in order to identify regular and abnormal packet interactions plays an important role while detecting a cyber attack over a network. In the proposed cyber attack model, the attacker uses a shell code injection to take control over the control network. Finally, the attacker uses the manufacturing message specification (MMS) protocol to modify the relay configuration and opens the circuit breaker maliciously. Each network packet consists of protocols (GOOSE and MMS) that defines properties associated with the respective network packet. The attacker modifies MMS protocol, which includes various properties associated with network packet like IP address, port number, sequence, and acknowledge number, etc. To detect an intrusion in the protection system, this research proposes a ridge regression based classifier integrated with a feature engineering pipeline.

Network packets from the relay first pass into a feature engineering pipeline before being analyzed by the ridge classifier model. Feature engineering pipeline transforms the features of an individual network packet into a feature vector. The network packet consists of various properties associated with it, but five features that can be leveraged from the properties are "frame length," "source IP," "destination IP," "source port number," and "destination port number," to detect an intrusion. The pipeline holds all the authorized IP addresses and port numbers. For every network packet, the pipeline checks if the "source IP," "destination IP," "source port number," and "destination port number" belong to the authorized pool of IP addresses and port numbers, if yes then it assigns the id associated with the respective IP address and port number, else it assigns 0 indicating an unauthorized IP address and port number. This research considers the ordinal relationship of the "source IP," "destination IP," "source port number," and "destination port number," since communication verification happens in an ordinal manner and hence these features are not converted into one-hot encoded vectors. Fig. 5 shows network packet properties before feature engineering pipeline and Fig. 6 shows the properties after feature engineering pipeline.

frame_len	ip_src	ip_dst	tcp_srcport	tcp_dstport
188	192.168.0.23	192.168.0.16	6163	102
185	192.168.0.14	192.168.0.16	6196	102
109	192.168.0.23	192.168.0.16	6163	102
108	192.168.0.14	192.168.0.16	6196	102
243	192.168.0.16	192.168.0.14	102	6196

Fig. 5. Network packet properties before feature engineering.

frame_len	ip_src	ip_dst	tcp_srcport	tcp_dstport
188	2	1	1	0
185	0	1	2	0
109	2	1	1	0
108	0	1	2	0
243	1	0	0	2

Fig. 6. Network packet properties after feature engineering.

The output of the feature engineering pipeline goes to ridge classifier for intrusion detection, which either outputs "0" indicating no intrusion or "1" indicating an intrusion. Ridge classifier is trained using five features, which include "frame length," "source IP," "destination IP," "source port number," and "destination port number," which are obtained from feature engineering pipeline. Ridge classifier [28] is basically a generalized representation of regularized linear regression. In ridge classifier, the classification probabilities are estimated by minimizing the least square cost function with L2 norm defined as

$$\arg\min_{\theta} \sum_{i=1}^{n} ||\theta x^{(i)} - y^{(i)}|| \; + \; \lambda||\theta||_2^2 \tag{7}$$

where λ greater than 0 is a regularization term, which prevents overfitting. In ridge classifier [28], estimation of the parameters and conversion of probabilities to crisp values for classification is done using a solver called "stochastic averaging gradient descent."

TABLE I
LIST OF HYPOTHESIS

Physical Anomaly	Cyber Anomaly	Conclusion
1	1	Cyber Induced mal-operation
1	0	Physical Fault Induced mal-operation
0	1	Cyber Induced mal-operation With No Physical Event
0	0	Normal Operation

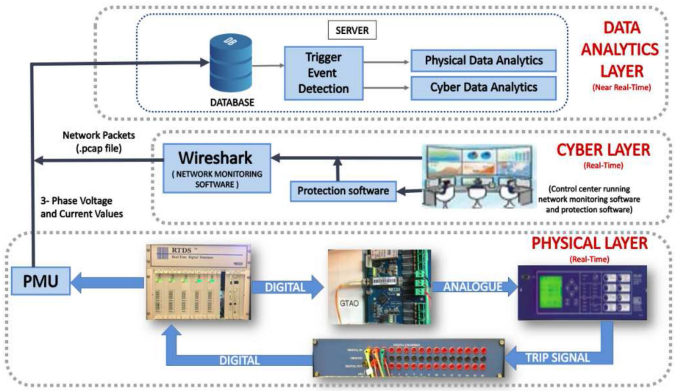

Fig. 7. Schematic of real-time simulation testbed.

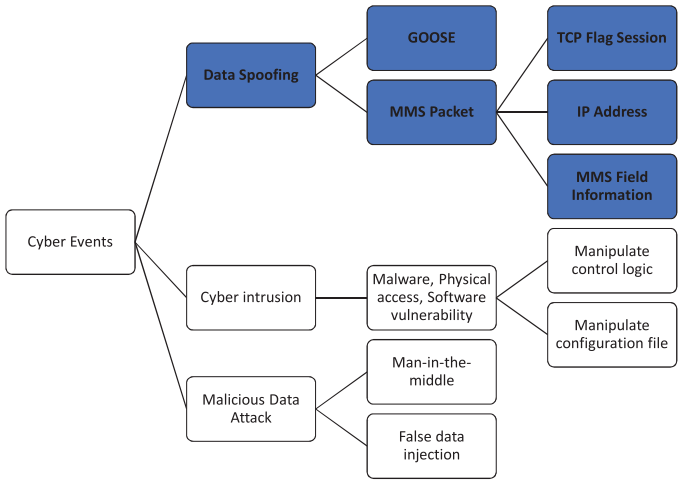

Fig. 8. Cyber event anomalies.

E. Final Decision and Control

Based on the results from physical and cyber data analytics, various conclusions can be drawn as shown in Table I. These results can further be validated using log files from relays, etc. It is to be noted that data analytics helps in making the anomaly detection faster by reducing the possible scenarios. However, the final conclusion and appropriate control decisions can be taken based on further manual validation using the log files of the relays. The advantage is that the operator does not have to look manually into log data of all the protective devices, but only of those devices, which the algorithm reports to be malicious.

IV. Real-Time Testbed

The real-time simulation testbed used to validate the proposed algorithm consists of three layers, namely physical layer, cyber layer, and data analytics layer as illustrated in the Fig. 7.

The physical layer is simulated in real time. It consists of RTDS, which implements custom hardware and software, specifically designed to perform real-time electromagnetic transient simulations and is used to simulate the power system network interfaced with a high-speed distance and directional protection relay. Current and voltage signals from RTDS are sent to relay using analog output channels in RTDS. Breaker trip and reclose signals from the physical relay are sent back to the RTDS. One of the circuit breakers in the RTDS simulation is controlled by a software relay model in RTDS, and the other is controlled by the physical relay connected to the simulator in hardware in the loop. The protective relay output contacts are connected to the RTDS simulator's digital interface panel to communicate breaker commands.

The cyber layer is simulated in real time. In cyber layer, the control center is running a protection software and Wireshark software. Wireshark software continuously extracts the network packets and streams them to the database in the data analytics layer.

The data analytics layer is near real time as the analysis on the data from physical and cyber layer is only triggered after a trigger event detection (change in breaker status). The data analytics layer is implemented on high performance computing server. The streaming data from physical layer (PMU: 3-phase voltage and current values) and data from cyber layer (Wireshark: network packets) are streamed and stored in the database and continuously monitored for an event *breaker status change*. As soon as the *breaker status change* event is detected, the data stored in the database are queried and analyzed for the final decision on event analysis.

V. Cyber Attack Modeling

The substation communication network has a remote access point such as virtual private network (VPN), dial-up, or wireless for control center operator or site engineer. An intruder always tries to access those points and may perform malicious cyber attack on substation protection relay and eventually trip the circuit breaker, which could result in a substantial blackout. In this paper, a prototype cyber attack similar to the Ukrainian event [3] is implemented in RTDS.

A. Cyber Anomalies

The substation automation system (SAS) is a target for cyber attack due to more substantial information and communication technology dependencies [29]. The SAS is equipped with different types of devices, such as network devices, user interface, server, global positioning system, firewall, intelligent electronic devices (IEDs), and remote access points. All these devices are vulnerable to various types of cyber attack such as data spoofing, man-in-the-middle attack, and data sniffing attack. Fig. 8 shows different types of cyber event anomalies that may exist in the SAS.

The SAS facilities use the IEC 61850 based protocol (GOOSE, SMV, and MMS) [30]. It is possible to perform a

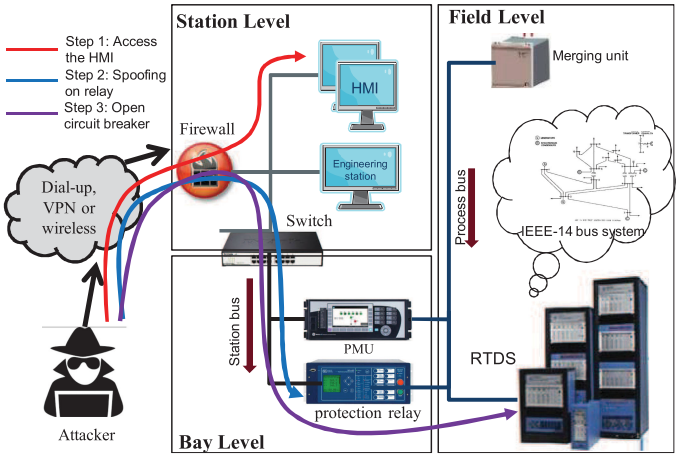

Fig. 9. Data spoofing attack on substation architecture.

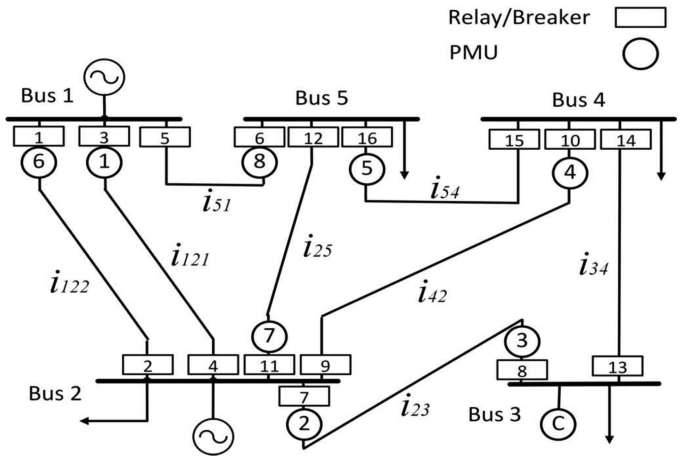

Fig. 10. Location of the PMUs and relays in IEEE 14-Bus System.

cyber attack on substation and control center through IEC 61850 protocol as it uses Ethernet-based communication and also lacks the encryption in field devices in the substation. GOOSE messages are used to send trip signals to the circuit breakers. By spoofing the MMS packet, an attacker can monitor and control the protection relay. MMS packet contains information such as TCP flag session (e.g., port number, sequence number, Ack number), IP address, and MMS field information (e.g., itemID, read/write status, etc.). All these information are vulnerable to cyber attack. In this paper, we mainly focus on cyber event anomalies related to the data spoofing attack on IEC 61850 protocol.

B. Cyber Attack Modeling

The modeling of cyber attack defined in this section is divided into three parts: first, a shell code injection, in which an email with malware is sent to the control center computer to gain unauthorized access over them; second, after getting access of two-level password of the email recipient's control center PC, the attacker analyzes the IEC 61850 protocol (GOOSE and MMS) packets to access the relay configuration session information. Finally, the attacker manipulates the relay configuration and opens the circuit breaker, causing disruptions in the power grid. Fig. 9 shows the overall view of data spoofing attack on substation architecture.

VI. Simulation Results

To validate the effectiveness of the proposed approach, IEEE 14-bus system is modeled as shown in the Fig. 10. Note that protection anomalies is expected in small subset of the system and not in large number of substations. Each PMU is receiving 3-phase voltage and current as listed in Table II.

A. LSTM-Based PMU Anomaly Detection

Data from Table II was acquired at a rate of 30 samples per second. Data anomalies were generated through the algorithm given in [15]. Modified values were inserted into normal PMU

TABLE II
PMU Information

PMU	Voltage Source	Current Source
PMU 1	V_1	I_{121}
PMU 2	V_2	I_{23}
PMU 3	V_3	I_{34}
PMU 4	V_4	I_{42}
PMU 5	V_5	I_{54}
PMU 6	V_1	I_{122}
PMU 7	V_2	I_{25}
PMU 8	V_5	I_{51}

TABLE III
LSTM Accuracy

Statistic	Voltage Mag.	Current Mag.	Timestamp	Voltage Ang.	Current Ang.
Recall	0.9998	1	0.1957	0.07933	0.04931
Precision	1	0.9993	1	1	1
False Positive	0.002	0	0.8043	0.9207	0.9507

data at various percentages, with some variation in the exact number of outliers generated. Additionally, these outliers could be made close to or far away from the normal values. In these experiments, anomalies were generated that differed from the expected value from 10% to 30 %. The LSTM network was first trained using a system of 30 hidden layers. The network was run on several different sets of anomalous data. Each set is focused on anomalies in a particular measurement, and the precision and recall were calculated accordingly. The network successfully identified anomalies in the magnitude of the voltage and current, as demonstrated in Table III. These results indicate that this network is excellent for identifying errors in relatively constant data. However, it marked many time stamps as anomalous. Since the network bases its predictions largely on its most recent input, training it for time stamps proposes some issues. The phase angles of the voltage and current readings suffered a similar problem. Similar to the issues with the time stamps, the LSTM recognizes the small increases between short portions of data. The identified anomalous data were replaced by most recent correct data. Advanced methods to replace the missing or bad

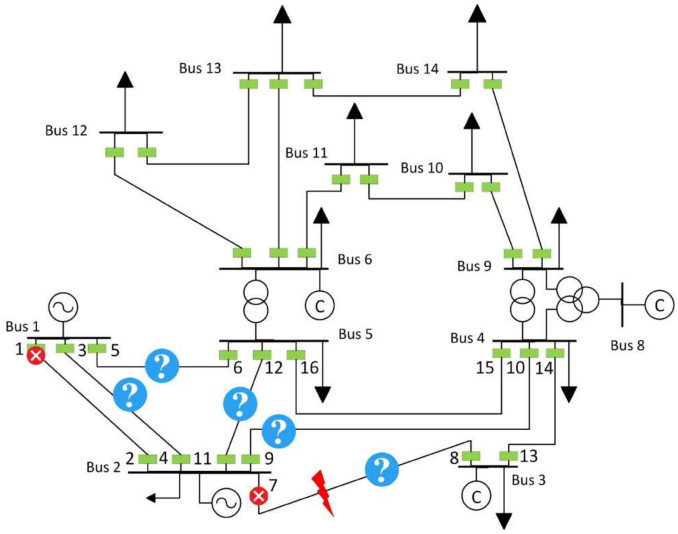

Fig. 11. IEEE 14-bus power system in abnormal condition.

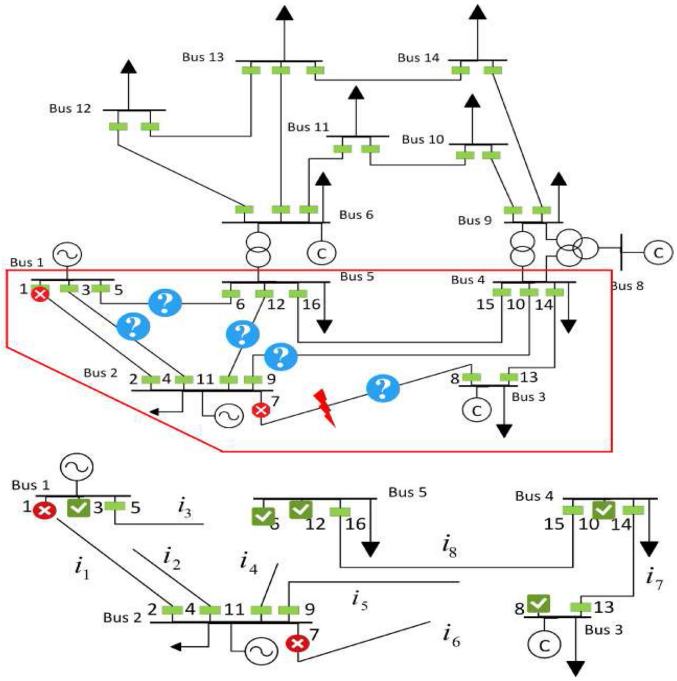

Fig. 12. ProNet and the status of the breakers.

data are currently being developed. These corrected data are, then, fed to the data analytics module.

B. Actual Scenario

For simulation purposes, it has been assumed that a fault occurred on line 2–3, but only breaker 8 has tripped to clear the fault. It is also assumed that the breaker 7 maloperated due to a cyber attack on the relay 7, which affected its functionality. In order to isolate the fault, relays 3, 10, and 12 responded to the fault in their zone 2 and tripped their corresponding breakers. Breaker 1 also malfunctioned because of the physical anomaly of the breaker. Hence, breaker 6 has tripped to isolate the fault. This is considered to be the actual scenario that had occurred as shown in the Fig. 11.

C. ProNet Formation and Hypothesis Formation

As a first step to correctly arrive at this conclusion (It is to be noted that the actual scenario is known for study purposes. In the actual case, the operator is unaware of the actual cause of the failure.), a ProNet has been created as shown in the Fig. 12. In this case, several possible scenarios can be interpreted based on the status of breakers (see Fig. 12). All these lines with at least one open breaker can be a candidate for the fault location. All possible scenarios explaining the current topology have been listed in Table IV. The actual event (Scn 0), that had occurred has been highlighted in Table IV.

D. Cyber Attack Implementation

For cyber attack, it has been assumed that no mitigation scheme have been employed. Initially, a brute force attack using a Telnet session to get access the two-level password of human–machine interface (HMI) was performed. After accessing the HMI, attack on MMS protocol is implemented. In this step, the MMS packet is scanned to determine the logical nodes (itemID) required to be manipulated to open the circuit breaker. $Wireshark$ has been used to extract the MMS packets. Fig. 13 shows the MMS packet information for a particular time stamp. In this particular time stamp, the logical node $BRK1CSWI1\$Co\$Pos\$Oper\$$ refers the breaker switch controller that is required to open the circuit breaker. Finally, the data spoofing attack on GOOSE protocol is performed to open the circuit breaker.

TABLE IV
POSSIBLE SCENARIOS

Scenario	Location of Fault	Initial incident	Consequential incident
	Line 2-3	Breaker 8 tripped Relay 7 malfunctioned	Breakers 3, 10, 12 tripped Relay 1 malfunctioned Relay 6 tripped
Scn 1	Line 2-4	Breaker 10 tripped Relay 9 malfunctioned	Breakers 3, 8, 12 tripped Relay 1 malfunctioned Relay 6 tripped
Scn 2	Line 2-1-2	Breaker 3 tripped Relay 4 malfunctioned	Breakers 8, 10, 12 tripped Relay 1 malfunctioned Relay 6 tripped
Scn 3	Line 1-5	Breaker 6 tripped Relay 5 malfunctioned	Relays 2,3,4 malfunctioned Breakers 8, 10, 12 tripped
Scn 4	Line 2-5	Breaker 12 tripped Relay 11 malfunctioned	Breakers 3, 8, 10 tripped Relay 3 malfunctioned Relay 6 tripped

E. Data Analytics Approach for Failure Diagnosis

In the next step of the solution approach, data analytics analyzes the physical data from PMU and cyber data to arrive at most possible explanation for the current network topology, as observed from the breaker status.

1) Physical Data Analytics: The physical data provided from PMU consist of 12 features, which includes 3-phase voltage and current phasor values. The physical dataset consist of 37500 PMU readings, which is further divided into three subdatasets

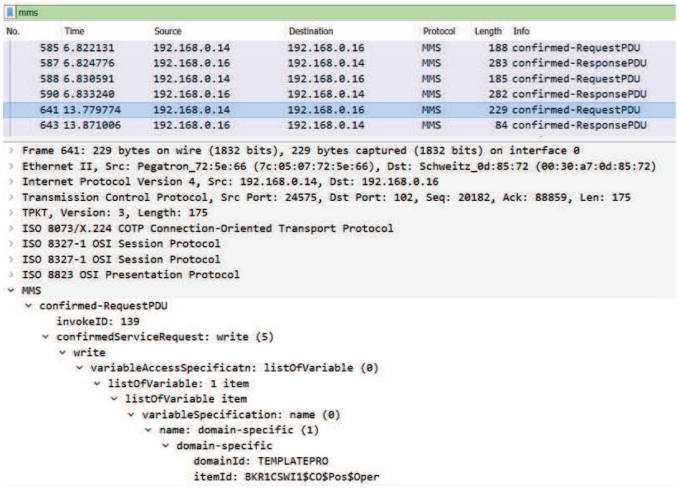

Fig. 13. MMS packer information of a particular time stamp.

TABLE V
DATASET DESCRIPTION

Dataset	Number of PMU Readings
Training Dataset	22250
Testing Dataset	11250
Validation Dataset	4000

TABLE VI
TYPES OF VALIDATION DATASET

Dataset	PMU readings (Normal)	PMU readings (Anomaly)
Validation Dataset	3979	21
SMOTE Validation Dataset	3979	3979

namely training, testing, and validation dataset as shown in Table V. The training and testing dataset consist of no anomalies, but the validation dataset consist of 4000 PMU readings out of which 21 readings represent anomalous PMU readings.

The validation dataset described in Table VI is unbalanced as it contains 3979 normal (majority) and 21 anomaly (minority) PMU readings. In order to balance out minority and majority PMU readings in the validation dataset, a state-of-the-art "synthetic minority oversampling technique" (SMOTE) presented in [31] is used, which helps in better evaluation of proposed deep autoencoder model in terms of accuracy, precision, recall, and *F*-measure.

In physical system, the deep autoencoder model shown in Fig. 2 is trained beforehand on training dataset from all eight respective PMUs for anomaly detection over the physical system. During an attack, "breaker status change" as shown in Fig. 2 acts as an initiating trigger event that initiates deep autoencoder model, which analyzes PMU data. The proposed deep autoencoder model is trained on training data and tested on test data as given in Table V and validated on "SMOTE validation dataset" as described in Table VI. The fundamental principle of training deep autoencoder on normal PMU measurements is to make the model learn and update weights of the neurons in the model during back propagation, such that the model is only able to reconstruct normal PMU values. Training the deep autoencoder only on normal measures of PMU will limit the model only to reconstruct normal measures of PMU, such that if an anomaly

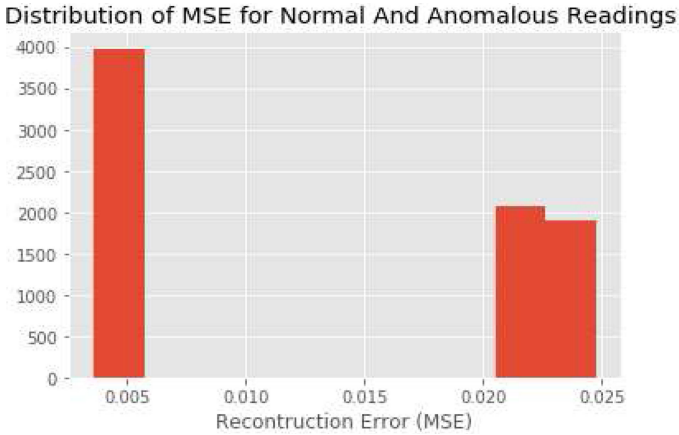

Fig. 14. Distribution of MSE for normal and anomaly values.

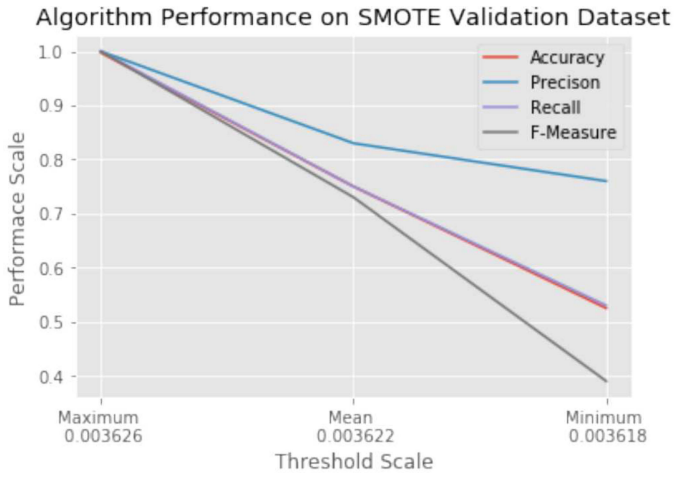

Fig. 15. Deep Autoencoder performance on varying threshold.

occurs over the system the deep autoencoder model will not be able to reconstruct the anomalous PMU readings. The metric MSE is also used to evaluate reconstruction error between input and output data. MSE is also used to decide a threshold value depending on the distribution and how well MSE for normal and anomalous readings are separated, as shown in Fig. 14.

The threshold for reconstruction error (MSE) is set empirically as discussed in [32]. The fundamental idea is to find a threshold point that best separates MSE of normal from that of anomalous readings, as shown in Fig. 14. Based on the selected threshold value of MSE, the PMU readings are labeled as normal-(0), which are less than threshold, and anomaly-(1), which are greater than threshold, during validation. Depending on the selected threshold value, the evaluation metrics like accuracy, precision, recall, and *F*-measure can vary. In Fig. 14, extreme left histogram represents MSE distribution for normal PMU readings and two histograms on extreme right show MSE distribution of anomalous PMU readings and the empty space between them represents the separation of normal-MSE and anomaly-MSE. To label a PMU reading in the validation dataset as normal-(0) or an anomaly-(1), three threshold values are evaluated as minimum, average, and maximum of test data MSE. Fig. 15 shows how accuracy, precision,

TABLE VII
DEEP AUTOENCODER PERFORMANCE

Threshold	Accuracy	Precision	Recall	F-measure
0.003618 (Minimum)	52.50 %	0.76	0.53	0.39
0.003622 (Mean)	74.99 %	0.83	0.75	0.73
0.003626 (Maximum)	99.74 %	1.0	1.0	1.0

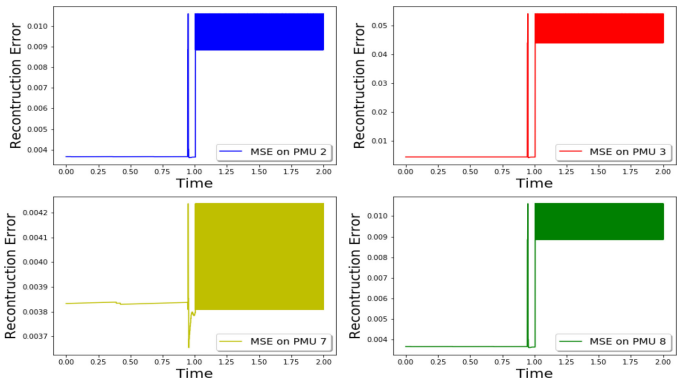

Fig. 16. Reconstruction error (MSE) for PMU 2, PMU 3, PMU 7, and PMU 8.

TABLE VIII
DEEP AUTOENCODER PERFORMANCE

PMU	Threshold (MSE-Maximum)	MSE (Anomaly)
PMU 1	0.003626	0.02337
PMU 2	0.003639	0.07648
PMU 3	0.004303	0.04887
PMU 4	0.003705	0.00437
PMU 5	0.003675	0.01113
PMU 6	0.003626	0.02337
PMU 7	0.003839	0.00401
PMU 8	0.003659	0.00968

recall, and *F*-measure can vary depending on selected threshold value.

Table VII further describes Fig. 15 about varying performance of the deep autoencoder model on the basis of the selected threshold value.

As observed from Table VII, maximum performance is achieved when maximum of MSE (test data) is selected as a threshold value. Hence, while detecting an anomaly over the physical system (PMU) maximum of test data, MSE is set as a threshold in order to obtain maximum accuracy.

In the physical system, eight different deep autoencoder models are trained on the respective PMUs and when an anomaly is triggered in the physical system, the models start analyzing respective PMU data and the results are shown in Fig. 16.

From Table VIII, data analytics concludes that PMU 2 and PMU 3 exhibit high MSE as compared to PMU 7 and PMU 8, which are far from the fault location. Further, when log files of the physical system are checked, it is verified that fault has occurred between PMU 2 and PMU 3. Fig. 16 shows the reconstruction error (MSE) scale for PMU 2, PMU 3, PMU 7, and PMU 8.

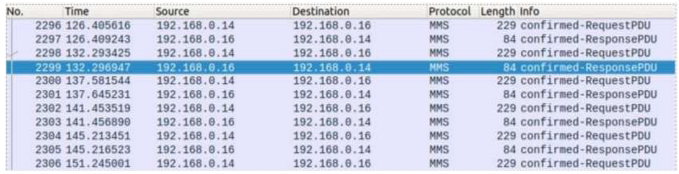

Fig. 17. Communication scenario during an attack.

2) Cyber Data Analytics: The IP addresses for system operator, relay, and attacker are 192.168.0.23, 192.168.0.16, and 192.168.0.14, respectively. Fig. 17 shows unauthorized IP address taking control over the relay and making a successful attack attempt over the cyber system. The cyber data (network packet) provided from relay unit are of ".pcap" format. The properties associated with the network packets that are extracted are shown in Fig. 5 and fed into the feature engineering pipeline. The engineering pipeline outputs formatted features for properties associated with network packets, as shown in Fig. 6. This is further fed into the ridge classifier to detect an intrusion. Ridge classifier is kept pretrained for intrusion detection. During an attack, it captures the network packets and analyzes them. The cyber data consist of 1238 network packets out of which 80% are used for training and 20% are used for testing using a stratified split. Stratified split is a variant of standard train-test split used in KFold cross validation, where the folds are made by preserving the percentage of samples for each class. Ridge classifier obtains 98.32 % accuracy on training data and 97.11 % on testing data. In the implemented cyber attack model, during the attack simulation, the cyber data analytic module gets triggered when the "breaker switch control" event is detected, followed by which, network packets form all relays are captured and analyzed for intrusion detection. During the simulation scenario, it was found that relay 7 was compromised by the attacker.

F. Detailed Investigation and Decision

For final conclusions on the failure diagnosis, consider Table VIII. It shows that PMU 2 and 3 show highest MSE among all PMUs, as shown in Fig. 16, although all the ProNet PMUs were affected by the fault. Based on the PMUs source information presented in Table II, it can be determined that the fault could have occurred in the line from bus 2 and 3. Also, based on the cyber anomaly detection, it can be concluded that there could have been a cyber attack on relay 7. The cyber data analytics for breaker 1 indicated that there is no intrusion and most probable cause for its malfunction could be a physical fault (insulation failure, etc.). Therefore, Scenario 0 in Table IV could be the most possible reason for the current topology as is given by the breaker status. Therefore, it can be concluded that breakers 1 and 7 have been intruded/maloperated, so the field engineers are needed to be sent to those exact substations and read the log files to arrive at the final conclusion.

VII. CONCLUSION

In this paper, a data analytics technique to monitor and detect malicious activity in the cyber-physical transmission protection

system are presented. The proposed method utilizes streaming PMU and cyber data, and breaker status data. Considering that the breaker status change as a trigger, multiple hypothesis theory is employed to generate hypothesis, which can explain the ongoing network condition in the system. First, bad data and missing data in the streaming PMU data streams are detected through LSTM networks. Next, the autoencoder model analyzes the PMU data for anomaly detection. Simultaneously, in cyber system, the ridge-based regression classifier analyzes network packets to detect an intrusion. Finally, the analysis of both modules are aggregated and fed to "most probable event" module to select the most probable reason for the event from the list of hypotheses. The final diagnosis and validations are made using the cyber log data of relays.

The main contributions of this paper are to implement cyber attack models to simulate protection system maloperation due to cyber attacks, to implement a semisupervised deep learning algorithm for anomaly detection over PMU, and to use ridge classifier for detecting an intrusion in relay. The proposed approach validated on a real-time testbed consisting of RTDS, hardware relay, and a cyber-attack model. The simulations on IEEE test system implemented on the test bed confirm that the proposed data analytics approach is able to diagnose abnormal operations in a transmission protection system caused by physical and cyber events. The proposed approach is scalable as protection system is local and focus on 3–4 substations for observed events. Possible future extension of the proposed data analytics technique includes an online learning technique for training deep autoencoder model and develop parallel implementation of the proposed technique.

ACKNOWLEDGMENT

The authors would like to acknowledge the technical support from Dr. B. Cui.

REFERENCES

[1] J. Jaeger and R. Krebs, "Protection security assessment—An important task for system blackout prevention," in *Proc. Int. Conf. Power Syst. Technol.*, Oct. 2010, pp. 1–6.

[2] N. Falliere, L. O. Murchu, and E. Chien, "W32.stuxnet dossier (v1.4)," Symantec, Mountain View, CA, USA, Tech. Rep., Feb. 2011.

[3] S. Soltan, M. Yannakakis, and G. Zussman, "REACT to cyber attacks on power grids," *IEEE Trans. Netw. Sci. Eng.*, 2018, *Preprints*.

[4] F. Yang, A. P. S. Meliopoulos, G. J. Cokkinides, and Q. B. Dam, "Effects of protection system hidden failures on bulk power system reliability," in *Proc. 38th North Amer. Power Symp.*, Sep. 2006, pp. 517–523.

[5] X. Yu and C. Singh, "Integrated power system vulnerability analysis considering protection failures," in *Proc. IEEE Power Eng. Soc. General Meeting*, Jul. 2003, vol. 2, pp. 706–711.

[6] X. Yu and C. Singh, "A practical approach for integrated power system vulnerability analysis with protection failures," *IEEE Trans. Power Syst.*, vol. 19, no. 4, pp. 1811–1820, Nov. 2004.

[7] G. Cardoso, J. G. Rolim, and H. H. Zurn, "Identifying the primary fault section after contingencies in bulk power systems," *IEEE Trans. Power Del.*, vol. 23, no. 3, pp. 1335–1342, Jul. 2008.

[8] Y.-C. Huang, "Fault section estimation in power systems using a novel decision support system," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 439–444, May 2002.

[9] H.-J. Lee, B.-S. Ahn, and Y.-M. Park, "A fault diagnosis expert system for distribution substations," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 92–97, Jan. 2000.

[10] A. Esmaeilian, T. Popovic, and M. Kezunovic, "Transmission line relay mis-operation detection based on time-synchronized field data," *Elect. Power Syst. Res.*, vol. 125, pp. 174–183, 2015.

[11] P. Kundu and A. K. Pradhan, "Synchrophasor-assisted zone 3 operation," *IEEE Trans. Power Del.*, vol. 29, no. 2, pp. 660–667, Apr. 2014.

[12] N. Mahadevan, A. Dubey, A. Chhokra, H. Guo, and G. Karsai, "Using temporal causal models to isolate failures in power system protection devices," *IEEE Instrum. Meas. Mag.*, vol. 18, no. 4, pp. 28–39, Aug. 2015.

[13] B. Yang, J. Yamazaki, N. Saito, Y. Kokai, and D. Xie, "Big data analytic empowered grid applications 2014—Is PMU a big data issue?" in *Proc. 12th Int. Conf. Euro. Energy Market*, May 2015, pp. 1–4.

[14] A. Kaci, I. Kamwa, L. A. Dessaint, and S. Guillon, "Synchrophasor data baselining and mining for online monitoring of dynamic security limits," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2681–2695, Nov. 2014.

[15] M. Zhou, Y. Wang, A. K. Srivastava, Y. Wu, and P. Banerjee, "Ensemble based algorithm for synchrophasor data anomaly detection," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2979–2988, May 2018.

[16] S. Pan, T. H. Morris, and U. Adhikari, "A specification-based intrusion detection framework for cyber-physical environment in electric power system," *IJ Netw. Sec.*, vol. 17, no. 2, pp. 174–188, 2015.

[17] R. C. B. Hink, J. M. Beaver, M. A. Buckner, T. Morris, U. Adhikari, and S. Pan, "Machine learning for power system disturbance and cyber-attack discrimination," in *Proc. 7th Int. Symp. Resilient Control Syst.*, 2014, pp. 1–8.

[18] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1773–1786, Aug. 2016.

[19] Y. Zhang, L. Wang, W. Sun, R. C. Green II, and M. Alam, "Distributed intrusion detection system in a multi-layer network architecture of smart grids," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 796–808, Dec. 2011.

[20] Z. Xiang, H. Guangyu, and W. Zhigong, "Masquerade detection using support vector machines in the smart grid," in *Proc. 7th Int. Joint Conf. Comput. Sci. Optim.*, 2014, pp. 30–34.

[21] A. Ahmed *et al.*, "Cyber physical security analytics for anomalies in transmission protection systems," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, Sep. 2018, pp. 1–8.

[22] N. T. Stringer and D. Waser, "An innovative method of providing total breaker failure protection," in *Proc. IEEE Ind. Appl. Conf. 30th IAS Annu. Meeting*, Oct. 1995, vol. 2, pp. 1165–1169.

[23] A. Nanduri and L. Sherry, "Anomaly detection in aircraft data using recurrent neural networks (RNN)," in *Proc. Integr. Commun. Navig. Surveillance*, Apr. 2016, pp. 5C2-1–5C2-8.

[24] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[25] A. F. Agarap, "Deep learning using rectified linear units (ReLU)," 2018, *arXiv:1803.08375*.

[26] S. Narayan, "The generalized sigmoid activation function: Competitive supervised learning," *Inf. Sci.*, vol. 99, no. 1/2, pp. 69–82, 1997.

[27] S. Ruder, "An overview of gradient descent optimization algorithms," 2016, *arXiv:1609.04747*.

[28] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[29] NERC, "High-impact, low-frequency event risk to the North American bulk power system," Jointly-Commissioned Summary Rep. North Amer. Electric Rel. Corporation U.S. Dept. Energy, Washington, DC, USA, Jun. 2010.

[30] R. Mackiewicz, "Technical overview and benefits of the IEC 61850 standard for substation automation," in *Proc. Power Syst. Conf. Expo*, Nov. 2006, pp. 623–630.

[31] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[32] N. Japkowicz *et al.*, "A novelty detection approach to classification," in *Proc. 14th Int. Joint Conf. Artif. Intell.*, 1995, vol. 1, pp. 518–523.