AnSMart: A SVM-based anomaly detection scheme via system profiling in Smart Grids

2021 2nd International Conference on Intelligent Engineering and Management (ICIEM) | 978-1-6654-1450-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICIEM51511.2021.9445353

Deepti Saraswat
Department of CSE
Institute of Technology, Nirma University
Ahmedabad, Gujarat, India
deepti.saraswat@nirmauni.ac.in

Pronaya Bhattacharya
Department of CSE
Institute of Technology, Nirma University
Ahmedabad, Gujarat, India
pronoya.bhattacharya@nirmauni.ac.in

Mohd Zuhair
Department of CSE
Institute of Technology, Nirma University
Ahmedabad, Gujarat, India
mohd.zuhair@nirmauni.ac.in

Ashwin Verma
Department of CSE
Institute of Technology, Nirma University
Ahmedabad, Gujarat, India
ashwin.verma@nirmauni.ac.in

Ashwani Kumar
Department of CSE
Sreyas Institute of Engineering and Technology
Nagole, Hyderabad, India
dr.ashwanikumar@sreyas.ac.in

***Abstract*—Anomaly detection techniques analyze consumer spend patterns and grid load profiles to predict possible deviations from normal behavior. However, as the measured data is time-varying, profiling the measured drifts becomes complex owing to the amount of raw generated data. Motivated fromthe aforementioned discussions, in this paper, we propose a scheme, *AnSMart*, to predict deviations in smart grid (SG) data through obtained profiling operations. The scheme operatesin two phases. In the first phase, valuable grid features are extracted from internal system call lists made by grid kernels after load profiling operations are completed over a day. Then,in the second phase, based on data logs, vector differences are computed for call vectors and malign scenarios are identified. The data is fed to the support vector machine (SVM) modelfor training, and compromised grid behavior is classified. SVM predicts metrics deviation from normal grids. The deviation is measured depending on parameters like- call vector signal, call- sizes, call-list deviations, and call-return values collected fromthe open-source *libiec61850* library that consists of resource-rich (RR) and resource-limited (RL) libraries for both compromised and uncompromised grids. Based on different cases, a total of** 50 **experiments were conducted. The obtained F-score is** 0.926 **and the accuracy of** 92.5**% is obtained based on system-call stacks and grid operating system (OS) behavior that outperforms the conventional anomaly-based approaches on SG.**

Index Terms*—Smart Grids, Anomaly Detection, System Profiling, Stack call traces.

I. INTRODUCTION

Conventionally, the power grids were designed to generate raw electricity power through a central generator that distributes energy units via transmission lines to different customers through step-down transformer operations. How-ever, with the rising surge of embedded sensors and cyber-physical systems, the energy requirements of grid operations have increased exponentially. According to a recent study, the estimated numbers of sensors in grid operations are expectedto rise to 800 million by 2025 [1]. To handle the growingdemands, centralized and manual distribution of power traditional grids is not a viable choice. With the rise of smartsensors in Internet of Things (IoT) ecosystems, the centralized traditional grids are upgraded to support decentralized smart meters installed at customer premises and electrical substations. The smart meters support energy-efficient, self- sustainable and low-powered metering operations [2]. Thus, decentralized control allows efficient load-balancing and minimizes fluctuations in power grid operations.

Smart meters communicate through power distribution units, peer grids, and other electrical substation units to facilitate on-demand energy transfer units based on scheduled load requirements measured through the day. The communicationis carried out through open wireless channels that support low-powered communication over high-performance switcheddevices that cover a large spatial range [3]. The open communication channels can be intercepted by malicious entities to perform network attacks to manipulate sensor readings. The malicious adversary occupying the smart grid could be a direct controller (such as an IED), indirect controller (spoofs measurement data) or a surveillance device (gathers sensitive/confidential data and measurements). These entities concern the security of consumer as well as grid centricoperations. In a recent study by North East Group over 125 countries, network attacks amounts to a loss of 96 billion which is levied on energy stakeholders [1]. Artificial Intelligence (AI) based techniques have gained prominencein decentralized grid ecosystems to measure deviations in consumer load patterns and overall load to measure anomaliesin grid behavior [4]. Smart meters supports bi-directional transactional profiling that handles the electricity flow from grid stations to meters, and vice-versa. However, the operations inside the smart meters are unknown to normal users, and hence secure profiling of meters is required.

In SG, transfer of electrical units is done through embedded sensors. The sensors are required to interact with grid controller units through open wireless channels. However, dueto weak encryption and authorization schemes, the grid

978-1-6654-1450-0/21/$31.00 ©2021 IEEE

data might get compromised owing to anomalous grid behavior. The classification of such anomalous grid behavior is a difficult task as it is difficult to identify grid components that supply false updates to the controller units. Such malign behavior includes control of current levels, voltage units and power distribution control; in extreme cases leads to unit overloading and grid cutoff. Thus, to address the limitation, proper authorization and encryption schemes are required to ensure uncompromised grid behavior. Moreover, SG employs automatic restoration and healing mechanisms that can intelligentlycontrol the grid operations, and allow fair pricing of energy units [16].

Table I: Comparative analysis of *AnSMart* with state-of-the-art schemes

Authors	**Year**	**Objective**	**Advantages**	**Limitations**
Agrawal *et al.*[5]	2015	Analysis of various data mining techniques to detect anomaly in the system to provide better understanding	Different behavior can be easily detected by data mining techniques.	Hybrid approach is not sufficient to overcome the limitations.
Kanovsky *et al.* [6]	2015	The detailed description of detection of counterfeit component is analyzed.	Using evaluation fake parts are easily detected	Not able to detect the fake parts using production test
Ahmed *et al.*[7]	2016	In-depth analysis of category wise anomaly detection	Various techniques for the detection of anomaly is used to categorize and analyze with intrusion detection data-set	Interaction and communication are not proper between instances of anomaly
Kosek[8]	2016	Contextual anomaly-based detection of malicious voltage control in grid	Voltage control and voltage distribution is easily detected	Less intrusion detection system analysis. More modules are required to add for power system control
Babun *et al.*[9]	2017	A light-weight system-level approach to detect counterfeit smart devices	Statistical analysis, functional call, tracing of system are used to detect the fake parts	Not deal with traditional security metrics like, recall, precision and accuracy.
Chamie *et al.*[10]	2018	Proposed micro phased measurement units as phasor measurement to detect anomaly in power grid system	Based on the transient property of the SG anomalies are detected which makes system scalable and more capable	The detection algorithm needs to place closer to the end-user to get better result
Karimipour *et al.*[11]	2019	A feature extraction scheme based on unsupervised anomaly detection based on statistical data	Suitable for large scale smart grid system that can identify the real faults from the disturbance with an accuracy of 99%.	User needs to spend time to classify the data and label the classes manually
Roy *et al.*[12]	2019	A detailed taxonomy of various machine learning (ML) models to analyze imbalance attacks.	The classifiers used in the study exhibited a better classification model with low false negative, which is a satisfactory performance	The effectiveness of the existing classifier can be improved by including the network packets and attacks that are unique to SG
Panthi *et al.*[13]	2020	Multiple ML algorithms are employed to detect and differentiate cyber-attacks and natural disturbance in power system	Assessed attacks which uses a deceptive technique to hide behind the normal technique and the analysis is done on data generatedby *IEEE 3-bus* system.	Classification of system disturbance is sometimes difficult using J-Ripper and One-R classification method.
Elmrabit *et al.*[14]	2020	A detailed analysis of ML algorithm is evaluated based on their ability to detect abnormal behavior over the network	The evaluation is carried out on three public data-sets and experimental work is performed through high-performance computing facility and the result is analyzed and presented	It's very difficult to identify the best ML algorithm to identify anomalies, as it depends on the type of data generated
Mokhtari *et al.*[15]	2021	A novel solution is proposed for anomaly detection based on measurement intrusion detection	A supervised ML model is used to classify normal and abnormal activities, which shows excellent performance in detecting data-sets anomaly	Not deal with traditional security metrics like, recall, precision, and accuracy.
Proposed	2021	Anomaly detection scheme in SG based on observed drifts in system call record vectors	Exhaustive set of parameters based on library call statistics are selected for RR, and RL devices	Attack scenarios are not considered

Thus, to detect the compromised smart devices, a scheme *AnSMart* is proposed in the paper. The scheme takes into account the kernel-level call values gathered from smart meter OS. The gathered information consists of profiling statistics like call-type, the length and duration of system calls, library functions, and call-sequences to the smart meters from SG. Based on the gathered information, the different metrics are fedinto SVM-based classifier. The data is analyzed for compromised & normal grids and benign devices are identified [17]. Apart from real-data, data is also collected from *IEC61850* test-bed protocol suite which is resistant against denial-of-service attacks [18].

A. Research Gap

In SG communications, researchers globally have proposed solutions to detect anomalies based on data mining schemes [5], contextual anomalies [8], phasor measurements [10], and many more. The proposed solutions have taken into account the anomaly behavior based on transient properties of grid ecosystems. As the amount of generated raw data is huge, the learning models tend to become inefficient over time, and bias measurements are required to be adjusted manually. This adds an overhead to the design of critical SG ecosystems for attack classifications, and subsequently, the detection of anomalous behavior classes. Thus, to address the research gap, the paper proposes a scheme that measures drifts based on call vector data, and computes the measured deviations from normal behavior, based on gathered call records from OSkernel logs. The vector length differences are computed, and the measured drifts are evaluated based on hamming distance, with consideration of different case scenarios. Once the malignbehavior is identified, an alert notification is raised to the grid user.

B. Research Contributions

The following are the research contributions of the article.

1) We present a scheme *AnSMart* that classifies anomalous behavior in compromised SG through a collection of log records from OS, libraries, and kernel logs.

2) Based on the collected log records, we present a training model based on an SVM-classifier that is trained on system call vector metrics that differs in the behavior of normal and compromised grids based on content, length, ordering, and type. For the same, we have computed the vector length difference between two call logs and presented scenario-based classification.
3) Once a grid is classified as anomalous, or non-anomalous,we present a notification scheme that raises an alarm trigger through sensors. The performance evaluation of the scheme is carried through confusion matrix, accuracy,and F-score based on data collected from resource-rich (RR) and resource-limited (RL) libraries. The obtained results indicate the viability of proposed scheme.

C. *Article Structure*

The structure of the article is as follows. Section II presents a comparative analysis of the proposed scheme with existing state-of-the-art schemes. Section III illustrates the proposed scheme. Section IV presents the performance analysis of the scheme based on the entire library call list details. Section V concludes the article.

II. State-Of-The-Art

In this section, we discuss the recent state-of-the-art schemes. A comparative analysis of existing schemes with theproposed scheme is presented in Table I. Issues such as inadequacies in grid infrastructure, cyber-security, storage concern, data management, communication issue, stability concerns, sensitivity to timing accuracy, and interaction among various components in the system possess challenges to efficient and reliable SG [19]. In addition to this SG operates under many uncertainties which could be arises from various factors such as quality of transmitting data, synchronization of devices, and capacity of computing resources [20], [21], [22]. By considering all these factors, growing concerns about the SG are security and durability, and this has been justified by the threats on the existing SG. In 2014, approximately 35% of the energy company in Europe experienced the tangible attacks recorded by intrusion detection systems [23].

To address the further issues, much research has been conducted in the same field; several works continued targeting theanalysis of the various threats and techniques to handle them. Agrawal *et al.* [5] presented various data mining techniques which helps analyzing the best technique for particular systembased on different parameters, i.e a detailed taxonomy on various kinds attack. Kanovsky *et al.* [6] described the key component of identifying counterfeit components; the main aim is to inform academician and industry professionals, how to identify the counterfeit components using some standard methods. Ahmed *et al.* [7] presented a detailed survey onnetwork anomaly detection techniques, and also discussed the main four detection technique based on classification and clustering. Kosek [8] presented a contextual anomaly detection method which observes the behavior of control actions and power system impacts, and thoroughly tested on ongoing voltage attack.

III. *AnSmart*: The Proposed Scheme

In this section, we propose *AnSMart*, a statistical scheme for detection of anomalous device constituting a SG environment.The need to develop such architecture comes from the accurate detection of the nowadays malicious activities posing specific danger and security threats as well as subsequent analysis to enable smoother operation of grids. The proposed scheme collects the system, library and kernel logs from the software layer controlling the grid and performs the statistical analysis using SVM classification and convolution network. Fig. 1 shows the division of proposed scheme in three layers explained in subsequent sections as below.

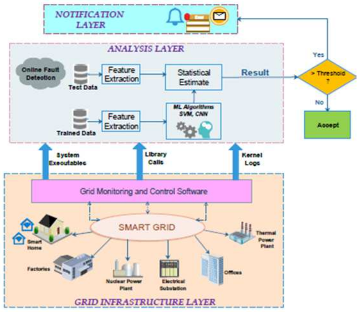


Fig. 1: *AnSMart*: The proposed scheme

A. *Grid Infrastructure Layer*

This layer consists of SG, supply, and consumer elements. The SG consists of a control and monitoring system, actuators,sensing elements controlled by computers through intelligent automation software. SG enables efficient, self-healing, reliable, secure, and less disturbed electricity transmission with reduced operations and cost. The supplier elements are typically electrical distribution systems (EDS), power plants (thermal, hydroelectric, nuclear), electrical substations, etc., and consumer elements are smart homes, offices, business places, factories, etc. The high-level software control and monitors the grid and gathers the data utilizing two-way digitalcommunication and provide output in form of logs. The latter is processed at the system level utilizing system executables, library calls, and kernel logs. These components are forwardedto the analysis layer for further processing.

B. *Analysis Layer*

In the proposed system, a classifier model is used to identify the compromised device (which performs the malicious activity). The classes used in our ML model takes into account the considerations as follows.

1) Number of calls for each type
2) Average number of calls for each type
3) A number derived from the call vector processing

Fig. 2 provides the detailed picture about the data flow of call vector signal generation up to the detection. The vector data is initially bifurcated in various frames through the process of *Framing*. Through a sliding operation (A_{frame}), these framesare mapped into various column vectors; the last vector being used as target or prediction vector. A predictor is used tocompare each individual data vector with the prediction vector.A successful prediction will produce a binary output of '0' and an unsuccessful prediction will produce a binary output of '1'.

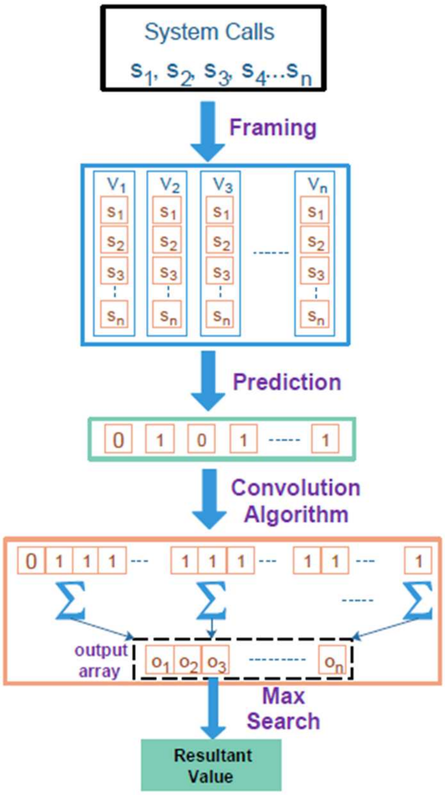
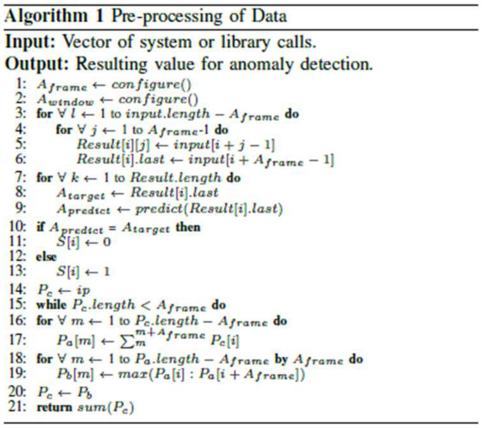


Fig. 2: Call vector signal flow diagram

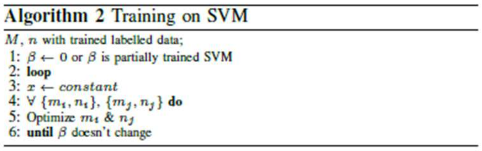
The produced binary outputs are vector processed and passed to a convolution plus maximum value search machinewhere the convolution kernel adds the binary values mapped by sliding processor A_{frame} and provides a constant value. It is to be noted that during the sliding process, the resulting vector length is reduced in each iteration since the maximumvalue search procedure utilizes non-overlapping frames. Ifthe resulting value is greater than the decision threshold, the process identifies an anomaly and vice-versa. The thresholdis decided based on the worst-case value (±3σ) of the loadprofile observed during the 24-hour window and a span ofseven days. Algorithm 1 describes the complete overview of the proposed scheme. Algorithm 2 explains the training of data on SVM. At the analysis layer, data is obtained from the grid management layer, combines the data logs derived from system and library call vectors, and applies ML algorithms to derive statistical deviation estimate of result to identify anomalous based on the electrical profile logged over a day. For the same, call vectors $C_V = \{ v_1, v_2, v_3, ... v_n\}$ are designed which are finite vectors of library or system calls at an instance$t$ when a computing unit completes its operation in-state Q_t. Here, we analyze the metrics of how two call vectors are different based on content, length, ordering, and type. We define a set of computation functions as can be explained below.



Algorithm 1 Pre-processing of Data

Input: Vector of system or library calls.
Output: Resulting value for anomaly detection.

```
$A_{frame} \leftarrow configure()$
$A_{window} \leftarrow configure()$
for $\forall\ l \leftarrow 1$ to $input.length - A_{frame}$ do
    for $\forall\ j \leftarrow 1$ to $A_{frame}$-1 do
        $Result[i][j] \leftarrow input[i + j - 1]$
        $Result[i].last \leftarrow input[i + A_{frame} - 1]$
for $\forall\ k \leftarrow 1$ to $Result.length$ do
    $A_{target} \leftarrow Result[i].last$
    $A_{predict} \leftarrow predict(Result[i].last)$
if $A_{predict} = A_{target}$ then
    $S[i] \leftarrow 0$
else
    $S[i] \leftarrow 1$
$P_c \leftarrow ip$
while $P_c.length < A_{frame}$ do
for $\forall\ m \leftarrow 1$ to $P_c.length - A_{frame}$ do
    $P_a[m] \leftarrow \sum_{m}^{m+A_{frame}} P_c[i]$
for $\forall\ m \leftarrow 1$ to $P_a.length - A_{frame}$ by $A_{frame}$ do
    $P_b[m] \leftarrow max(P_a[i] : P_a[i + A_{frame}])$
$P_c \leftarrow P_b$
return $sum(P_c)$
```



Algorithm 2 Training on SVM

```
$M$, $n$ with trained labelled data;
$\beta \leftarrow 0$ or $\beta$ is partially trained SVM
loop
$x \leftarrow constant$
$\forall\ \{m_i, n_i\}, \{m_j, n_j\}$ do
Optimize $m_i$ & $n_j$
until $\beta$ doesn't change
```

1) ***Vector Difference***: Vector difference V_D (v_1, v_2) computes how two vector calls are different according to theirtype of calls they inhibit. Let M be the set of calls in v_1, and$N$ be the set of calls in v_2. The function V_D (v_1, v_2) computes the unique calls present in v_1 but not in v_2. Formally stated,
$V_D(v_1, v_2) = |M - N|$.
2) ***Vector Length Difference:*** Vector length difference $V_{LD}(v_1, v_2)$ computes the difference of number of system calls contained by two vectors. $V_{LD} = 0$ means two call vectorsare of same length, $V_{LD} \neq 0$ means number of system callsin one vector is greater than other vectors.
3) ***Euclidean Distance:*** Euclidean distance E_L (v_1, v_2) combines both type and length difference between two vectors.Let M be the set of calls in v_1, and N be the set of calls in v_2. It defines the difference in number of calls made at time t. It can be stated as $E_L(v_1, v_2) = |v_{1(M)} - v_{2(N)}|$.

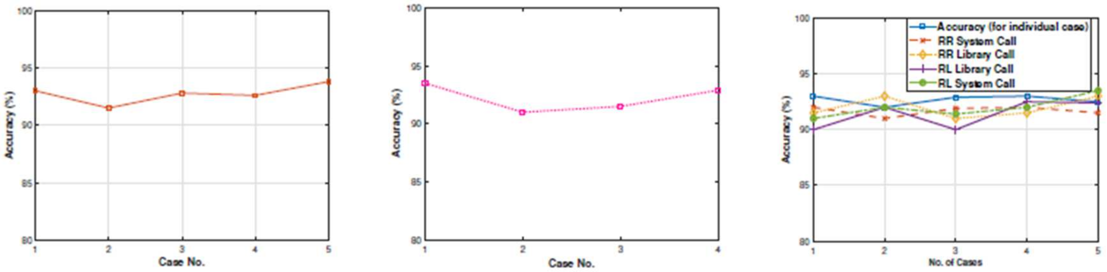


Fig. 3: Simulation Results of *AnSMart* scheme: (a) Accuracy of proposed scheme for individual cases, (b) Accuracy of data collected from resource rich (RR) and resource limited (RL) devices. Here case 1 to 4 represents RR-library, RR-system, RL-library and RL-system respectively, and (c) Accuracy of proposed scheme for RR and RL calls

4) *Hamming Distance:* Hamming Distance H_D (v_1, v_2) computes how many number of operations required to make two call vectors identical. $H_D(v_1, v_2) = 0$ means two vectors are identical. Following scenarios are considered as critical cases which also affect the performance evaluation of theproposed system. V_m represents the malign state and V_b represents benign state.

- $V_D(V_m, V_b) > 0$ means malign state makes a call that is not present in benign state.
- $V_D(V_m, V_b) < 0$ means benign state makes a call that is not present in malign state.
- $V_{LD}(V_m, V_b) \neq 0$: Both v_1 and v_2 have same type of calls, only differ in lengths. Later cases assumes $V_{LD}(V_m, V_b) = 0$.
- $E_L(V_m, V_b) = 0$: Both v_1 and v_2 are of same length and also same have type of calls but differs in individual internal distribution. Later case assumes $E_L(V_m, V_b) = 0$.
- $H_D(V_m, V_b) \neq 0$: Here all conditions of $E_L(V_m, V_b)$ are satisfied but only difference is that their orders are different.

C. *Notification Layer*

This layer provides the notification to the interactive control and monitoring display system based on the detailed computation of results by the analysis layer. If the value exceedsthe threshold, it activates the sensor system and alerts the user through an alarm, text message, or e-mail. The sensor is placedacross the grid, providing comprehensive real-time big-datato the central computer. Moreover, the layer is responsiblefor sensitizing the control centre regarding any failure, maintenance, interference, tampering, electrical leakages, missing meter reads as well as avoids field visits for any necessary disconnects/reconnects, configurations and firmware updates by allowing them to be performed remotely.

IV. *AnSmart*: PERFORMANCE EVALUATION

The section presents the schematics of the experimental methodology and setup and then presents the simulation results.

A. *Experimental Setup*

The ML-based classifier operation is treated as an unbiased *experiment* consisting of *events*. As an event, the SG device transits on the concerned state releases a call vector and go back to the idle state. The number of events x defined for sample space S for experiments $\{E_1, E_2, E_3, \ldots E_n\}$ is randomly chosen (defined by random variable X) according to Gaussian distribution function with mean μ and variance σ^2 as indicated in equation 1. A probability density function of the total number of calls is used to indicate the variance of the latter. The mean and variance are chosen on a trial and error basis to ensure that the distribution and experiments are not affected by the number of outcomes.

$$P(X = x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in S(E) \qquad (1)$$

The SG device after initializing the process control makes the transition to malicious state (V_m) or benign state (V_b) with probabilities p_m and p_b respectively. At V_m or V_b, the device provides a call vector and then moves back to the idle state V_I. A total of 50 experiments were carried out for each case as mentioned in the previous section, with the following criteria.

- 25 experiments with $p_b = 1$ and $p_m = 0$. It is assumed that real data is arriving from authentic devices.
- 25 experiments with $p_b = 0.99$ and $p_m = 0.01$.

B. *Data Set*

The proposed scheme is tested on real data obtained through open source IEC61850 library i.e. *libiec61850* that implementsan SG device. We consider two types of devices viz. resource- rich (RR) and resource-limited (RL). RL devices have simple hardware and software architecture with minimum memory& computational complexity (e.g. PLC, RTU). RR deviceshave high-end system configuration and significantly higher memory than RL devices (e.g. IED, PMU). The data-set from RR and RL device simulates attacks such as data manipulationand information leakage.

C. *Simulation Results*

As per the experimental setup, we now present the confusion matrix, as presented in Table II. The SVM learning classifier is used to identify the optimum resulting value of the differentiated call vector to support compromised devices.

However, the classifier is trained on authenticated call vectors. The decision model is trained with 2/3 of the experiments, and theremaining 1/3 is used for testing. Fig. 3 (a) shows the results as per variables defined above. It is clear that from cases 1 to 5, the proposed scheme can demarcate between anomalous andnon-compromised devices based on differential measurement of call vectors with much enhanced accuracy. The experiment was also run with benign and malicious components having equal call vector lengths. The F-score computed is 0.926 which tends to match the average accuracy number of *AnSMart* of 92.5% for cases 1 to 5 and better compared to the previous work [24]. Fig. 3(b) depicts that the algorithm efficientlyuses both system and library calls in resource-rich devices. Fig. 3(c) shows such comparison of individual cases withRR and RL system call vectors and library call vectors. The average accuracy obtained in this case also is around 92.3% that verifies the viability of the proposed scheme.

V. CONCLUSION & FUTURE WORK

In this paper, we propose a scheme, *AnSMart* that presents an SVM-classifier-based scheme to classify anomalous behavior of SG ecosystems. The scheme presents a convolution technique based on call vector data frames. Based on the category of SG (compromised or normal), drifts are observed in generated libraries and system calls, and vector length differences are computed. Then, based on malign and benign states, the hamming distance setup is formulated for different scenarios. The results are fed into an SVM-classifier model, that forms classification labels on whether the grids are compromised, or not. The scheme is tested on real data obtained through the *libiec61850* library, which consists of grid information of RR and RL calls. The performance results indicate the scheme viability. As part of future work, the authors would presenta secure grid ecosystem where grid transactions are recordedin blockchain to mitigate attack vectors from an adversary. As attack probability gets mitigated, we would present an anomaly-based intrusion detection model that collects datafrom the blockchain that improves the overall resiliency of the SG ecosystems.

REFERENCES

[1] G. Fenza, M. Gallo, and V. Loia, "Drift-aware methodology for anomalydetection in smart grid," *IEEE Access*, vol. 7, pp. 9645–9657, 2019.

[2] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid — the newand improved power grid: A survey," *IEEE Communications Surveys Tutorials*, vol. 14, no. 4, pp. 944–980, 2012.

[3] P. Bhattacharya, A. K. Tiwari, and R. Srivastava, "Dual buffers optical based packet switch incorporating arrayed waveguide gratings," *Journalof Engineering Research*, vol. 7, no. 1, 2019.

[4] A. Kumari, S. Tanwar, S. Tyagi, N. Kumar, M. S. Obaidat, and J. J. P. C. Rodrigues, "Fog computing for smart grid systems in the 5g en- vironment: Challenges and solutions," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 47–53, 2019.

[5] S. Agrawal and J. Agrawal, "Survey on anomaly detection using data mining techniques," *Procedia Computer Science*, vol. 60, pp. 708–713, 2015.

[6] A. Kanovsky, P. Spanik, and M. Frivaldsky, "Detection of electronic counterfeit components," in *2015 16th International Scientific Confer- ence on Electric Power Engineering (EPE)*, pp. 701–705, IEEE, 2015.

[7] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *Journal of Network and Computer Applications*, vol. 60, pp. 19–31, 2016.

[8] A. M. Kosek, "Contextual anomaly detection for cyber-physical securityin smart grids based on an artificial neural network model," in *2016 JointWorkshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG)*, pp. 1–6, IEEE, 2016.

[9] L. Babun, H. Aksu, and A. S. Uluagac, "Identifying counterfeit smart grid devices: A lightweight system level framework," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2017.

[10] M. El Chamie, K. G. Lore, D. M. Shila, and A. Surana, "Physics-based features for anomaly detection in power grids with micro-pmus," in*2018 IEEE International Conference on Communications (ICC)*, pp. 1– 7, 2018.

[11] H. Karimipour, S. Geris, A. Dehghantanha, and H. Leung, "Intelligent anomaly detection for large-scale smart grids," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, pp. 1–4,2019.

[12] D. D. Roy and D. Shin, "Network intrusion detection in smart grids for imbalanced attack types using machine learning models," in *2019 In- ternational Conference on Information and Communication TechnologyConvergence (ICTC)*, pp. 576–581, 2019.

[13] M. Panthi, "Anomaly detection in smart grids using machine learning techniques," in *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, pp. 220–222, 2020.

[14] N. Elmrabit, F. Zhou, F. Li, and H. Zhou, "Evaluation of machine learn-ing algorithms for anomaly detection," in *2020 International Conferenceon Cyber Security and Protection of Digital Services (Cyber Security)*, pp. 1–8, 2020.

[15] S. Mokhtari, A. Abbaspour, K. K. Yen, and A. Sargolzaei, "A machine learning approach for anomaly detection in industrial control systems based on measurement data," *Electronics*, vol. 10, no. 4, 2021.

[16] R. Singh, A. Singh, and P. Bhattacharya, "A machine learning approach for anomaly detection to secure smart grid systems," in *Advancementsin Security and Privacy Initiatives for Multimedia Images*, pp. 199–213, IGI Global, 2021.

[17] P. Bhattacharya, S. Tanwar, U. Bodke, S. Tyagi, and N. Kumar, "Bindaas:Blockchain-based deep-learning as-a-service in healthcare 4.0 applica- tions," *IEEE Transactions on Network Science and Engineering*, pp. 1–1,2019.

[18] R. E. Mackiewicz, "Overview of iec 61850 and benefits," in *2006 IEEE Power Engineering Society General Meeting*, pp. 8 pp.–, 2006.

[19] R. Kappagantu and S. A. Daniel, "Challenges and issues of smart grid implementation: A case of indian scenario," *Journal of Electrical Systems and Information Technology*, vol. 5, no. 3, pp. 453–467, 2018.

[20] A. Verma, P. Bhattacharya, U. Bodkhe, A. Ladha, and S. Tanwar, "Dams: Dynamic association for view materialization based on rule mining scheme," in *The International Conference on Recent Innovations in Computing*, pp. 529–544, Springer, 2020.

[21] K. C. Sou, H. Sandberg, and K. H. Johansson, "On the exact solution to a smart grid cyber-security analysis problem," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 856–865, 2013.

[22] A. Anwar and A. N. Mahmood, "Vulnerabilities of smart grid state esti- mation against false data injection attack," *Renewable energy integration*,pp. 411–428, 2014.

[23] J. Pagliery, "Hackers attacked the us energy grid 79 times this year," *CNN Money. Cable News Network. Retrieved*, vol. 16, 2015.

[24] C. Kaygusuz, L. Babun, H. Aksu, and S. Uluagac, "Detection of compromised smart grid devices with machine learning and convolutiontechniques," pp. 1–6, 05 2018.