Misbehaviour Detection for Smart Grids using a Privacy-centric and Computationally Efficient Federated Learning Approach

Muhammad Akbar Husnoo†, Adnan Anwar†, Nasser Hosseinzadeh‡, Robin Doss† , Biplab Sikdar*
†*School of Information Technology, Deakin University, Geelong, VIC 3216, Australia*
‡*School of Engineering, Deakin University, Geelong, VIC 3216, Australia*
**Department of Electrical and Computer Engineering, National University of Singapore, Singapore*
Email: {mahusnoo, adnan.anwar nasser.hosseinzadeh robin.doss}@deakin.edu.au, elebisik@nus.edu.sg

***Abstract*—Federated Learning (FL)-based Intrusion Detection Systems (IDSs) have recently surfaced as viable privacy-preserving solution to decentralized grid zones. However, conventional synchronous FL methods face technical challenges including the lack of consideration of communication delays and straggler nodes. To level the playing field, we propose a novel power system misbehaviour detection framework that leverages semi-asynchronous federated learning and dynamic aggregation. Specifically, our framework introduces an adaptive learning rate mechanism in the semi-asynchronous FL setting, allowing for efficient model updates and mitigating the impact of stragglers on the training process. Experiments conducted on publicly available Mississippi State University and Oak Ridge National Laboratory Power System Attack (MSU-ORNL PSA) Dataset demonstrate that our adaptive learning semi-asynchronous FL framework achieves superior attack detection rate while safeguarding data confidentiality and minimizing the negative effects of practical world communication latency and straggler nodes. Furthermore, our proposed method shows a significant 40% improvement in training time compared to conventional synchronous FL methods, showcasing the effectiveness and efficiency of our recommended approach.**

***Index Terms*—Federated Learning, Intrusion Detection, Internet of Things (IoT), Smart Grid**

I. Introduction

The rapid adoption of Internet of Things (IoT) devices along with revolutionary advancements in communication technology within modern power systems has sparked an unparalleled surge in data synthesis by edge devices. However, the progression of cyberattack fabrication methods alongside poor attack mitigation strategies, has resulted in the vulnerability of Smart Grids (SG) to cyber threats launched by malicious entities seeking illicit monetary and/or political benefits [1]. Consequently, it becomes imperative to design effective defenses to safeguard the valuable assets of SGs to counteract such forms of cyber threats.

In the recent past, several academic works [2], [3] have proposed innovative protection strategies against cyber threats in Smart Grids (SGs) using cutting-edge machine learning and deep learning techniques. Nevertheless, centralized intrusion defense solutions suffer from storage limitations, communication bottlenecks, and privacy concerns. To address these challenges, Federated Learning (FL)-based cyberattack countermeasures [4] have emerged as promising privacy-preserving solutions. These approaches leverage distributed learning by limiting data sharing and allowing cooperative on-device model training at the edge. For example, a federated false data injection attack detection mechanism using transformers was introduced in [5], showcasing its effectiveness. Despite these advancements, current FL-based cyberattack countermeasures predominantly rely on synchronous aggregation protocols such as FedAvg [6] and its extensions. These protocols involve the central orchestrator broadcasting the model, waiting for updates from all participating clients, and aggregating the local updates until convergence. However, resource-constrained Intelligent Electronic Devices (IEDs) in real-world scenarios present limitations. Studies [7] have highlighted issues with existing FL-based cyberattack detection methods. Firstly, delays in wireless SCADA communication protocols like IEC 61850 and unexpected dropout of client nodes lead to stragglers and communication inefficiency. Secondly, faster nodes experience global learning suspensions and timeouts while waiting for all updates. Thirdly, resource wastage occurs when competent client nodes remain idle due to node selection in large-scale setups. To address these challenges, asynchronous FL methods [8] have been proposed as alternatives. However, they assume physical homogeneity of data, which is unrealistic within a practical SG scenario due to continual data sensing by IEDs.

In this paper, we present semi-asynchronous federated intrusion detection framework with a dynamic aggregation to distinguish between adversarial cyberattacks and natural power system disturbances in decentralized power grids, while factoring in the presence of straggler client nodes. Thus, the major breakthroughs of this manuscript are as follows: 1) To counter the drawbacks of physical homogeneity and resource availability, we introduce a lightweight and privacy-preserving on-device collaborative semi-asynchronous intrusion detection framework in relation to power control systems. Our framework incorporates dynamic federated aggregation techniques to adap-

979-8-3503-1090-0/23/$31.00

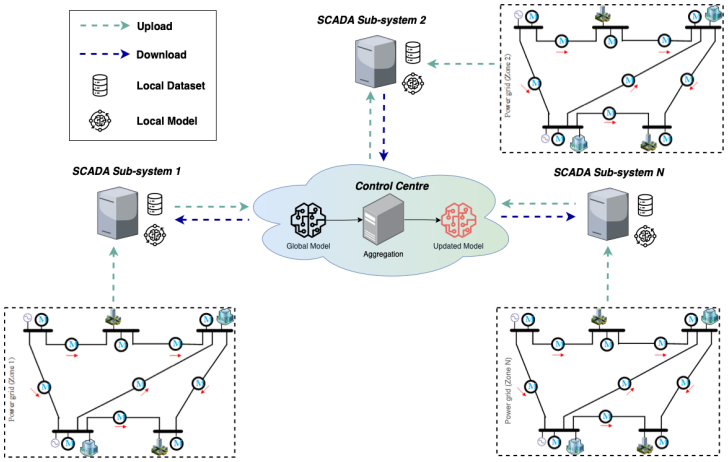

Fig. 1. Overview of our proposed misbehaviour detection power system model. The proposed framework with adaptive learning rate is employed for the training process. The framework incorporates an adaptive learning rate scheme, where the learning rate is initialized as η_0 and is decayed by a factor γ after each model aggregation. This adaptive learning rate allows the framework to dynamically adjust the influence of each device's update on the aggregated model. After reaching a predetermined cut-off time, the control center initiates the aggregation of model parameters.

tively adjust the aggregation process based on the availability and reliability of participating client nodes. By dynamically adapting the aggregation mechanism, this mitigates the impact of straggler nodes and communication delays, leading to improved convergence speed and efficiency. 2) To improve the anomaly detection rate in power systems, we leverage a representation learning-based Deep Auto-Encoder (DAE) model within the proposed framework. By utilizing the power of deep learning, the Deep Auto-Encoder model effectively learns and captures complex patterns and features from power system data, enabling accurate discrimination between cyber-attacks and natural disturbances. The dynamic aggregation mechanism ensures that the learned knowledge from individual client nodes is efficiently aggregated to enhance the overall detection performance. 3) Finally, we extensively evaluate the performance of our suggested approach using the Mississippi State University and Oak Ridge National Laboratory Power System Attack (MSU-ORNL PSA) Dataset [9]. The evaluation demonstrates the robustness of the solution in handling physical heterogeneity and resource constraints. Comparative analysis with state-of-the-art models reveals that it achieves a superior detection rate with an accuracy of 93.2% and reduces the training time by 40% in the presence of straggler nodes. These results validate the effectiveness and efficiency of our proposed dynamically aggregated semi-asynchronous FL framework for intrusion detection in power control systems.

The remaining structure of this article is organized as follows: Section II briefly introduces the power system model and highlights the key challenges faced by synchronous FL-based cyberattack detection methods. The proposed attack detection module and our solution is presented in Section III. Section IV presents the experimental scenarios used to demonstrate the reliability of our proposed approach through evaluation on publicly accessible datasets. Finally, Section V wraps up the paper with brief remarks.

II. Power System Model & Problem Formulation

Throughout this manuscript, we examine a generic decentralized smart grid system that is divided into a variable number of grid zones denoted as $N \in \mathbb{R}^+$. As depicted in Fig. 1, we briefly discuss the three primary actors in our system model as follows: 1) *Control Centre*: The control centre serves as the central orchestrator responsible for coordinating and monitoring grid operations. It is assumed to possess adequate computational infrastructure and acts as the central orchestrator of the federated intrusion detection system. 2) *SCADA Sub-systems*: The SCADA sub-systems are client nodes equipped with advanced data collection capabilities, playing a vital role in monitoring and collecting data from Intelligent Electronic Devices (IEDs). These SCADA sub-systems function as client nodes, jointly training local models on their respective datasets and updating model parameters to the control centre. 3) *Grid Zones*: The power grid zones are equipped with multiple IEDs such as relays and synchrophasors, as well as sensor networks continuously capturing power-related data, including current, phase angle, and voltage. Additionally, we assume that the SCADA sub-systems are honest and non-colluding, meaning they do not communicate with each other laterally to ensure data integrity and confidentiality, model quality, fairness etc.. Furthermore, the IEDs are connected to the SCADA sub-systems through stable high-speed communication networks to enable continuous data transmission.

The primary objective of our proposed solution is to optimize the objective function g_k for the global model parameter $P \in \mathbb{R}^M$. This objective function is defined as $g_k(P) = \frac{1}{N_k} \sum_{j=1}^{N_k} \ell\left(c_k^j, d_k^j; P\right)$, where $\ell(.) : \mathbb{R}^M \times \mathbb{R} \rightarrow \mathbb{R}$ represents the loss function that needs to be optimized for the data point $\left(c_k^i, d_k^i\right)$. Furthermore, the overall objective function G for all K distributed datasets can be expressed as $G(P) = \sum_{k=1}^{K} \frac{N_k}{N} g_k(P)$, which is the sum of local loss functions for each k. Here, N represents the total number of data samples

used for training. To achieve model convergence, popular federated learning aggregation algorithms such as FedAvg, FedSGD, and their extensions [10] are employed to aggregate local weights from all federated client nodes over multiple training rounds.

Current state-of-the-art literature on FL-based misbehaviour detection [5], [11], [12] lack consideration of communication delays between the central orchestrator and the client nodes. However, in the practical world, wireless communication protocols (such as IEC 61850 protocol) used for data transmission are subject to communication delays and unreliability which may be due to unforeseen circumstances including natural disasters, demographical locations, high magnetic fields, and so on. Definitively, in terms of misbehaviour detection within decentralized smart grid systems, communication latency is an urgent issue which is barely studied in the context of federated attack detection. Furthermore, active node dropouts due to power or connectivity issues is frequent. The scarcity of literature [13] that take into consideration the effect of stragglers signifies the urgent need to design fault-resilient intrusion detection frameworks for power system anomaly detection which is tackled throughout this paper.

III. Proposed Method

Throughout this section, we first describe the deep learning model considered for cyberattack detection followed by the proposed robust asynchronous FL-based intrusion detection framework.

A. Attack Detection Module

Our primary objective is to tackle the issue of intrusion detection in power systems by utilizing a Deep Auto-encoder (DAE) model, as illustrated in Figure 2. DAEs are highly effective unsupervised neural networks designed for representation learning, with the aim of achieving an identity mapping between input and output data. Our proposed DAE network architecture consists of two symmetrical Deep Belief Networks (DBNs): the first five shallow layers form the encoding portion of the network, while the subsequent five shallow layers constitute the decoding part. Restricted Boltzmann Machines (RBMs) form the fundamental building blocks of DBNs and are employed in each layer due to their generative modelling capabilities, amongst others. These RBMs enable bidirectional communication, acting as hidden layers for the preceding nodes and visible layers for the subsequent nodes. However, within a single layer, there is no lateral communication among nodes. The RBMs are trained using the Contrastive Divergence Algorithm. In our model, rectified linear unit (ReLU) activation functions are applied to all DAE layers to improve the effectiveness of RBMs. ReLU activation is easier to train and often yields better results. The reconstruction disparity, quantified by the Mean Squared Error (MSE), serves as an indication of the presence of cyber attacks. In our case, if the reconstruction loss r exceeds a predefined threshold τ, it suggests the occurrence of a cyber attack. During the training process, we sort the reconstruction errors for the training set in ascending order, selecting the value of t at the inflection point in the loss distribution. This value is employed as a threshold for classification purposes. To facilitate classification, softmax layers are added following the RBM stack. We set the learning rate and batch size to 10^{-3} and 100 respectively for training. Our DAE model is the foundation for cyberattack classification in our research article.

B. Semi-asynchronous FL Framework with Dynamic Aggregation

Algorithm 1: Proposed Framework

Input: Initial learning rate η_0, learning rate decay factor γ, Active SCADA sub-system K , local data sample D_k for $k \in [1, K]$, time for aggregation T_a, time cost for SCADA sub-systems $[s_0, s_1, ..., s_k]$

Initiate common unanimous model parameter P_0, parameter buffer B and learning rate $\eta = \eta_0$;

for *each training round* t *in* $T_{cl} \in (1, n)$ **do**

At Control Centre:
Get local gradient (g_k^t, t_k) from each node k;

$s_k \leftarrow$ adjust t_k; $B \xleftarrow{+} (g_k^t, t_k)$;

if *Time for aggregation* T_a *is attained* **then**

$B' \leftarrow$ group by model version t and compute aggregation;

foreach *(*g_k^t, t_k*) in* B' **do**

Add new global g and t;

end

Broadcast freshly aggregated global parameters (g, t) to respective SCADA sub-systems k;

Clear B;

Decay learning rate: $\eta \leftarrow \gamma \cdot \eta$;

end

At SCADA Sub-System:
for *each* $k \in [K]$ **do**

Calculate local gradient $g_k^t = \nabla f_k(Z^t)$ using D_k;

Calculate time cost of each client t_k;

Scale local gradient by learning rate: $g_k^t \leftarrow \eta \cdot g_k^t$;

Update (g_k^t, t_k) to control centre;

end

end

Output: Global Model, M

As depicted in Figure 1, our proposed solution focusses on a federated power system setup where K distributed SCADA

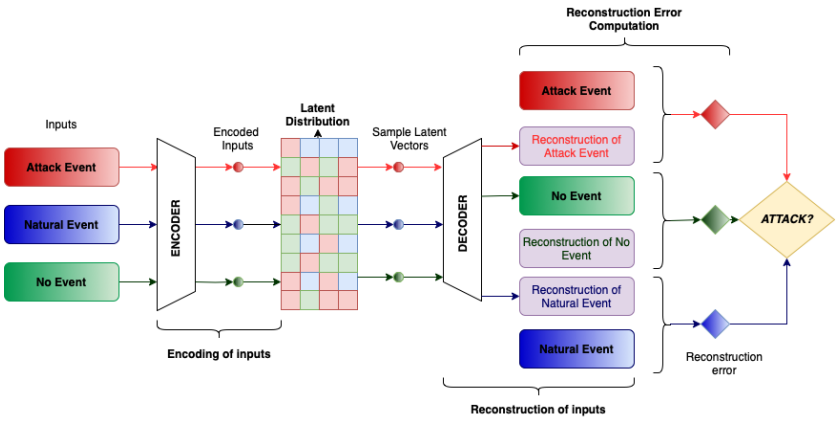

Fig. 2. Graphical overview of the proposed misbehaviour detection model architecture with data samples as they go through the the anomaly detection process. Our proposed misbehavior detection model comprises three key components: the Encoder, Latent Distribution, and Decoder. The Encoder compresses input data into a lower-dimensional latent representation. The Latent Distribution enforces a bottleneck, ensuring a compressed knowledge representation for further analysis. The Decoder reconstructs the original input based on the latent vectors generated by the Latent Distribution. By comparing the reconstructed output with the original input, the model facilitates anomaly detection. This schematic provides a concise overview of our misbehavior detection model architecture and the data flow during the anomaly detection process.

sub-systems are connected to the control centre through communication protocols. We acknowledge the presence of transmission delays and errors in the communication between the SCADA sub-systems and the control centre, which hampers stable and reliable communication due to the heterogeneous nature of the edge-based SCADA sub-systems. To address these challenges, we propose a semi-asynchronous federated learning (FL) framework that leverages adaptive learning rate and model update aggregation. The semi-asynchronous framework overcomes the limitations of synchronous FL, such as stragglers and the training time associated with waiting for all local model updates. However, it introduces the issue of staleness in local updates, which can hinder learning performance. To handle staleness, we incorporate an adaptive learning rate algorithm that dynamically adjusts the influence of each sub-system's update on the aggregated model. The learning rate, denoted as η, is initialized at η_0. After each model update aggregation, the learning rate is decayed by a factor γ: $\eta \leftarrow \gamma \cdot \eta$. This adaptive learning rate ensures that recent model updates have a stronger impact on the aggregated model, thereby mitigating staleness issues. By combining the semi-asynchronous framework, model update aggregation, and adaptive learning rate algorithm, we strike a balance between reducing staleness and optimizing the learning process in the federated power system setup. This approach overcomes the challenges posed by transmission delays and errors, while enhancing the overall convergence and performance of the FL-based intrusion detection system model.

During each communication round t of the proposed framework with adaptive learning rate (as detailed in Algorithm 1), the control centre and the SCADA sub-systems perform the following steps: 1) The control centre obtains the gradients (g_k^t, t_k) from each SCADA sub-system k. 2) If the aggregation time T_a is reached, the control centre clusters the model updates by the time version and computes the aggregation. The newly aggregated global parameters $(g,\ t)$ are then updated to the source SCADA sub-systems. 3) The control centre also updates the model parameter buffer B and the time cost s_k for each SCADA sub-system. 4) At the SCADA sub-systems, each sub-system k computes the local gradient $g_k^t = \nabla f_k(Z^t)$ using its local training set D_k. 5) The sub-systems scale their local gradient by the current learning rate: $g_k^t \leftarrow \eta \cdot g_k^t$. 6) The scaled gradients $(g_k^t,\ t_k)$ are then updated to the control centre. 7) Once the aggregation time T_a is reached, the control centre aggregates the model updates, updates the global parameters, and distributes the updated parameters to the sub-systems. The learning rate is also decayed by multiplying it with the decay factor γ: $\eta \leftarrow \gamma \cdot \eta$. By incorporating adaptive learning rate into the framework, the learning rate is adjusted during the training process. Decay factor γ can be used to control the rate at which the learning rate decreases over time. This adaptive learning rate allows the framework to adaptively change the influence of each device's update on the aggregated model, leading to improved convergence and performance in the federated learning process.

IV. Experimental Validations and Discussions

Throughout this section, we initially describe the data used for the experimental validation followed by the empirical aftermath that substantiate the attack detection effectiveness, robustness and computation efficiency of our propounded framework as discussed in Section III.

A. Dataset Description and Feature Engineering

We evaluate our proposed architecture using the publicly available industrial control system dataset, Mississippi State University and Oak Ridge National Laboratory Power System Attack (MSU-ORNL PSA) Dataset [9]. It is a three-class dataset which consists of 15 data files exhibiting matching feature composition whereby scenarios of *Natural Events* (sporadic Single Line to Ground (SLG) disturbances), *Attack Events*

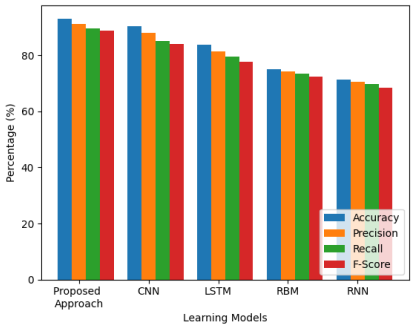

Fig. 3. Average intrusion detection rate of proposed framework vs. state-of-the-art federated models using four metrics: accuracy, precision, recall and F-Score metrics.

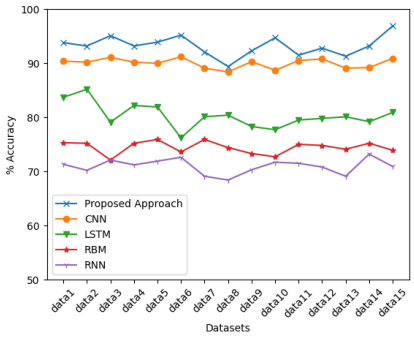

Fig. 4. Detection accuracy of proposed approach vs. state-of-the-art federated models over several data files.

(data injection attacks, remote tripping command injection attacks, etc.) and *No Events* have been simulated and recorded. The real-time synchrophasor measurement data collected at a sample rate of 120 per second from four Phasor Measurement Units (PMUs) constitutes of 128 features including frequency, current phase angle, voltage phase magnitude, etc. We impute missing values present in the dataset using K-Nearest Neighbour (KNN) machine learning algorithm in view of avoiding information loss. To train our proposed attack detection module, a basis of 100 features has been chosen via Principal Component Analysis (PCA) to cull out unnecessary features and decrease training complexity. Next, we normalize the input features into the range (0,1). Lastly, we divide the dataset into a train-test split of 70% and 30% respectively.

B. Intrusion Detection Performance

TABLE I
AVERAGE INTRUSION DETECTION RATE OF PROPOSED FRAMEWORK VS. STATE-OF-THE-ART MODELS

	Federated Models				
Metric (%)	***Prop. Approach***	***CNN***	***LSTM***	***RBM***	***RNN***
Accuracy	93.2	90.4	83.7	75.1	71.4
Precision	91.3	88.1	81.4	74.2	70.6
Recall	89.7	85.2	79.5	73.6	69.9
F1-Score	88.9	84.2	77.7	72.5	68.4

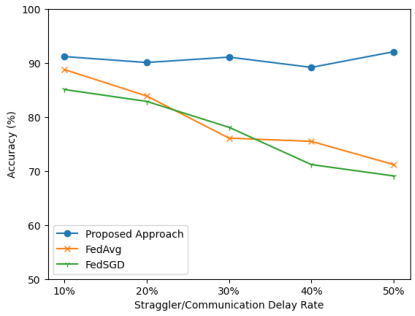

Fig. 5. Resilience of proposed approach vs. synchronous FL aggregation solutions based on accuracy with rising number of affected nodes.

First, we validate and contrast the intrusion detection performance of our novel against four state-of-the-art models namely Convolutional Neural Network (CNN), Restricted Boltzman Machine (RBM), Long Short Term Memory (LSTM) and Recurrent Neural Network (RNN), which are all trained in a similar federated set-up. The global models, obtained after training and fine-tuning, are then evaluated on the test set split. As illustrated in Fig. 3 and Table I, we compare the detection rate of our proposed framework using the four main classification metrics namely accuracy, precision, recall and F-1 Score. Anomalies, being distinct from normal patterns, can elevate prediction errors and contribute to higher averaged loss values as defined by our objective function. From these experimental results, we note that our proposed approach achieves superior detection performance as opposed to all competing models considered. Specifically, our framework achieved the highest overall accuracy of 93.2% as opposed to CNN (90.4%), LSTM (83.7%), RBM (75.1%) and lastly, RNN (71.4%). Likewise, in terms of other classification evaluation metrics, our work outperforms the other rivalling federated models. We also notice that RNN produces the lowest cyberattack detection performance amongst all methods. Furthermore, we assess the cyberattack detection performance of our proposed architecture on the data files forming part of the three-class MSU-ORNL PSA Dataset. Experimental validations, as depicted in Fig. 4 reveal that our proposed architecture outshines other popular cyberattack detection algorithms For instance, the accuracy of our solution on dataset 1 is 93.8% as opposed to CNN (90.4%) or RBM (75.3%). Identically, for dataset 15, our proposed solution achieves an accuracy rate of 93.2% as compared to LSTM (79.2%) or RNN (73.1%). Therefore, we can establish that our proposed semi-asynchronous federated malicious attack detection with dynamic aggregation attains ample intrusion detection performance in contrast to that of state-of-the-art deep learning algorithms whilst promising data privacy and minimizing the effect of communication delays and stragglers.

C. Resilience to stragglers

Considering the negative constraints of transmission delays and stragglers within the practical SG scenario, we propose

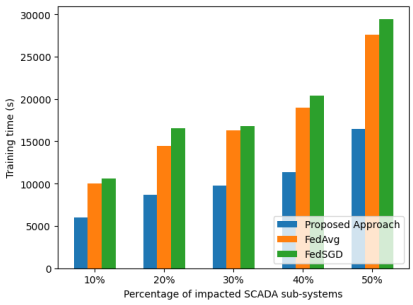

Fig. 6. Computation time of proposed approach vs. synchronous FL aggregation methods on varying number of affected SCADA sub-system nodes.

to solve this challenge using a semi-asynchronous FL-based approach with dynamic adaptive learning rate. To replicate the effects of delayed clients, we initially configure $K = 10$ and initiate a respite mechanism for a selected number of client nodes such that they are unresponsive for a certain amount of time. Next, we contrast the robustness of our solution against that of two classical synchronous FL aggregation solution FedAvg and FedSGD. As illustrated in Fig. 5, we perceive that the performance of our framework tends to stay relatively consistent with growing number of impacted nodes. On the flip side, synchronous FL-based aggregation algorithms experience a drastic intrusion detection performance downturn with rising number of straggler SCADA sub-systems. Therefore, we can conclude that our semi-asynchronous attack detection algorithm is robust against communication latency and straggler nodes.

D. Computational Efficiency

Consequently, we assess the computational efficiency through training time taken by the semi-asynchronous federated approach with dynamic aggregation as discussed in Section III. Explicitly, we compare the time for training taken for model convergence of our proposed solution against that of FedAVG and FedSGD in varying scenarios of impacted nodes. As presented in Fig. 6, we note that there is comparatively a decrease in training time to reach model convergence by our algorithm. In particular, it can be highlighted that the proposed solution achieves around 40% improvement in computational efficiency in comparison with rivalling aggregation methods. However, we also note an increase in the computation time of our proposed approach with increasing number of affected nodes which is related to more communication rounds needed for model convergence with increasing number of straggler nodes.

V. Conclusion

Throughout our work, we initially put forward a computation efficient and privacy-centric misbehaviour detection strategy to alleviate the real-world impacts of communication unreliability and straggler nodes within decentralized power grid systems. Moreover, we present a representation learning-based auto-encoder for accurate discernment between cyberattacks, power faults and normal operations. Lastly, we comprehensively validate the performance of our proposed solution by utilizing the publicly available MSU-ORNL PSA Dataset in terms of detection rate, robustness and computation efficiency. The experimental results reveal that our proposed approach achieves superior misbehaviour detection performance as compared to other state-of-the-art deep learning models whilst achieving a 40% improvement in computation time in contrast to classical synchronous aggregation methods. In future, we plan to explore robust techniques to enhance the robustness of our federated framework as one of our previous works [14] suggests their extreme vulnerability to Byzantine attacks.

References

[1] M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "False data injection threats in active distribution systems: A comprehensive survey," *Future Generation Computer Systems*, vol. 140, p. 344–364, Mar 2023.

[2] A. Anwar, A. Mahmood, B. Ray, M. A. Mahmud, and Z. Tari, "Machine learning to ensure data integrity in power system topological network database," *Electronics*, vol. 9, no. 4, p. 693, Apr 2020.

[3] A. Anwar, A. N. Mahmood, and Z. Shah, "A data-driven approach to distinguish cyber-attacks from physical faults in a smart grid," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, ser. CIKM '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 1811–1814.

[4] X. Xu, W. Liu, Y. Zhang, X. Zhang, W. Dou, L. Qi, and M. Z. A. Bhuiyan, "Psdf: Privacy-aware iov service deployment with federated learning in cloud-edge computing," *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 5, oct 2022.

[5] Y. Li, X. Wei, Y. Li, Z. Dong, and M. Shahidehpour, "Detection of false data injection attacks in smart grid: A secure federated deep learning approach," *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4862–4872, 2022.

[6] M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "Fedrep: Towards horizontal federated load forecasting for retail energy providers," in *2022 IEEE PES 14th Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 2022, pp. 1–6.

[7] W. Wu, L. He, W. Lin, R. Mao, C. Maple, and S. Jarvis, "Safa: A semi-asynchronous protocol for fast federated learning with low overhead," *IEEE Transactions on Computers*, vol. 70, no. 5, pp. 655–668, 2021.

[8] L. Cui, Y. Qu, G. Xie, D. Zeng, R. Li, S. Shen, and S. Yu, "Security and privacy-enhanced federated learning for anomaly detection in iot infrastructures," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3492–3500, 2022.

[9] U. Adhikari, S. Pan, T. Morris, R. Borges, and J. Beaver. [Online]. Available: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets

[10] A. Nilsson, S. Smith, G. Ulm, E. Gustavsson, and M. Jirstrand, "A performance evaluation of federated learning algorithms," in *Proceedings of the Second Workshop on Distributed Infrastructures for Deep Learning*, ser. DIDL '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–8.

[11] M. A. Husnoo, A. Anwar, H. T. Reda, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "Feddisc: A computation-efficient federated learning framework for power systems disturbance and cyber attack discrimination," *Energy and AI*, p. 100271, May 2023.

[12] M. Wen, R. Xie, K. Lu, L. Wang, and K. Zhang, "Feddetect: A novel privacy-preserving federated learning framework for energy theft detection in smart grid," *IEEE Internet of Things Journal*, vol. 9, no. 8, pp. 6069–6080, 2022.

[13] M. A. Husnoo, A. Anwar, H. T. Reda, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "Fedisa: A semi-asynchronous federated learning framework for power system fault and cyberattack discrimination," in *IEEE INFOCOM 2023 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2023, pp. 1–6.

[14] M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "A secure federated learning framework for residential short term load forecasting," *IEEE Transactions on Smart Grid*, pp. 1–1, 2023.