Detection of Stealthy Cyber-attack in Distributed DC Microgrids Based on LSTM Neural Network

1st Xingquan Fu
School of Cyber Science and Engineering
Southeast University
Nanjing, China
751176041@qq.com

2nd Mengfei Niu
School of Cyber Science and Engineering
Southeast University
Nanjing, China
mengfeiN@163.com

3rd Guanghui Wen
School of Mathematics
Southeast University
Nanjing, China
wenguanghui@gmail.com

2021 International Conference on Neuromorphic Computing (ICNC) | 978-1-6654-1287-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICNC52316.2021.9608150

***Abstract*—Distributed DC microgrids, which mainly consist of cyber layer and physical layer, can be classified as a classical cyber-physical system. Compared with the physical layer, the cyber layer has a significant risk of malicious cyber-attack due to its inherent vulnerability. This paper investigates a lightweight cyber-attack detection problem for distributed DC microgrids with bounded process and measurement noises. More precisely, the cyber-attack under consideration is assumed to be the stealthy cyber-attack. A long short-term memory (LSTM) neural network-based strategy is proposed to construct the attack detection scheme without using the information of the noises. Case studies indicate that the proposed attack detection strategy can work effectively in the following two cases: i) the statistical characteristics of system noises are unknown; ii) the magnitude of the cyber-attack is lightweight, which is detected extremely hard by traditional detectors such as χ^2 detector.**

***Index Terms*—Attack detection, stealthy cyber-attack, distributed DC microgrids, LSTM neural network**

I. Introduction

In recent years, distributed DC microgrids have been received widespread attention due to its good performance in resilience to faults and power quality [1]–[5]. However, cyber-attack may occur in microgrids, which can cause tremendous damage. Hence, the research on the security of distributed DC microgrids is also indispensable [6]. Note that the distributed DC microgrids are mainly consisted of sensors, actuators and Distributed Generation Units (DGUs), where data interact among all those components over wireless communication channels. Therefore, malicious attackers may take advantage of the vulnerability of wireless transmission to launch malicious cyber-attack and cause incalculable repercussions.

Attack detection, also known as attack monitoring, is committed to discover whether the system is suffered from cyber-attack [7]–[13]. Attack detection plays an important role in resisting cyber-attack, which are the prerequisite for attack defense such as isolating the attacked node [14] or shutting down the system [15]. The application of state estimation in attack detection methods such as χ^2 detector [16] is quite frequent because the internal state of the system must change under cyber-attack. In different cases, the methods of state estimation are various. For example, Kalman filter [17] is used in the case of white gaussian noise and Luenberger observer [18] can be chosen under the condition that noises can be ignored. In this paper, we employ an unknown input observer (UIO) [19] to deal with exogenous input in the distributed DC microgrid. When information about the noises is sufficiently known, such as the probability density function of the noises are known or the noises are bounded and its boundary is known, a threshold function about the state estimate is constructed and system is considered to be under attack if value of the threshold function is greater than the stated threshold [7], [8]. However, when the system is assaulted by lightweight cyber-attack, the traditional attack detection algorithms are inefficient. The rationale is that magnitude of the cyber-attack is so tiny that the function value is lower than the specified threshold.

The increase in computing capabilities and theory brings fast development of artificial intelligence, so as deep learning. LSTM neural network has become one of the most popular techniques of deep learning since it is firstly proposed by Hochreiter [20], which has been applied in many fields such as load prediction, text generation and so on [21]–[24]. As a classical Recurrent Neural Network (RNN), LSTM neural network can handle the problem of time-series classification. Meanwhile, attack detection might be considered to be a binary categorization with a normal or attacked system. Motivated by this, LSTM neural network might be applied to attack detection.

This paper contributes a new attack detection strategy to deal with stealthy cyber-attack, which mainly contains two steps. Firstly, a time-series about the state observed by the UIO is obtained. Secondly, the time-series is put into a trained LSTM neural network to determine whether the system has been attacked. In particular, the system noises are bounded but the boundaries are unknown, which is different from most existing studies. To the best of authors' knowledge, it is the first time to combine UIO and LSTM neural network in attack detection algorithm.

This paper continues as follows: Section II formulates the dynamic model of a DC Microgrid and describes the detection of stealthy cyber-attack. Section III presents UIO, LSTM neural network and proposed attack detection algorithm. The distributed DC microgrid parameters and experimental results of attack detection are given in Section IV. Finally, Section V describes the conclusion.

Notion: $\mathbb{R}$ stands for the set of real numbers. For a matrix

2021 International Conference on Neuromorphic Computing (ICNC)

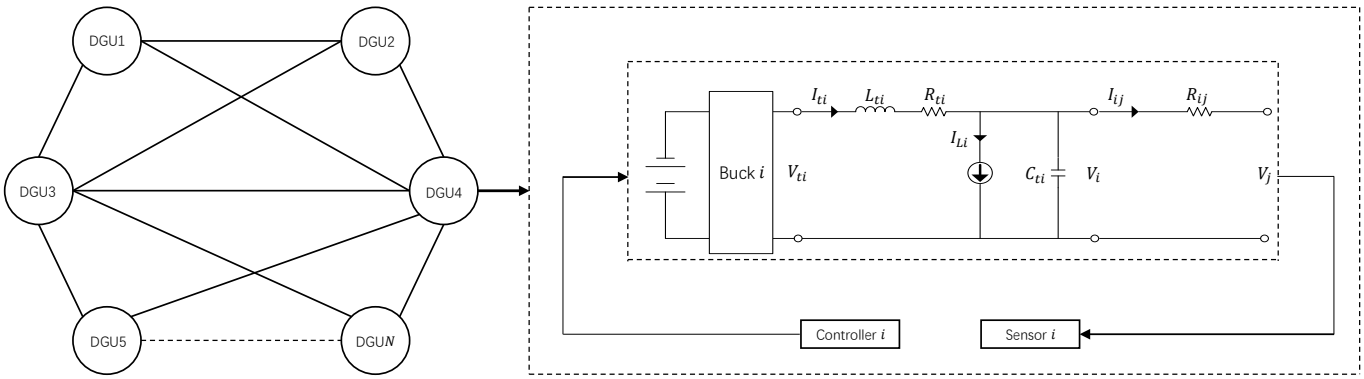

Fig. 1. Overall architecture of the distributed DC microgrid. On the left, the system consists of N DGUs; on the right, the graph representing the circuit diagram of DGUi. Meanwhile, sensor$i$ receives data from DGUi and controller$i$ sends data to DGUi.

$M \in \mathbb{R}^{m\times n}$, M^T describes transpose of M. I is an identity matrix with compatible dimensions. Symbol $|\cdot|$ represents the Euclidean norm of a vector. Symbol $\|\cdot\|$ denotes the spectral norm of a matrix.

II. Problem Formulation

A. Dynamic Model of a Distributed DC Microgrid

We focus on line-independent plug-and-play DGUs, which consist of a Buck converter and RLC filter [1]. As shown in Fig. 1, a distributed DC microgrid consists of N interconnected DGUs and the ith subsystems can be described as following:

$$S_i : \begin{cases} \dot{x}_i = A_{ii}x_i + B_i u_i + M_i d_i \\ \qquad + \sum_{j\in\mathcal{N}_i} A_{ij}x_j + w_i \\ y_i = H_i x_i + v_i \end{cases} \tag{1}$$

where S_i represents the dynamics of the ith DGU, $x_i = [V_i, I_{ti}]^T$, $u_i = V_{ti}$, $d_i = I_{Li}$, y_i represent the ith DGU state, control input, exogenous input and measurable output, respectively. $\mathcal{N}_i$ is the set of neighbors of the ith DGU and H_i is a measurement matrix with compatible dimensions. Process disturbance w_i and measurement disturbance v_i are bounded, but the bounds are unknown, i.e.,

$$|w_i| \le \bar{w}_i, \qquad |v_i| \le \bar{v}_i \tag{2}$$

where $\bar{w}_i$ and $\bar{v}_i$ represent the boundaries of w_i and v_i, respectively. Note that $\bar{w}_i$ and $\bar{v}_i$ are unknown positive scalars.

Remark 1: Motivated by relevant studies on distributed DC microgrids [1]–[3], the Quasi-Stationary Line (QSL) model is considered [1]. Where A_{ii}, B_i and M_i represents system matrix, coupling matrix of input u_i and coupling matrix of unknown input d_i, respectively. And all of aforementioned matrices are linear time invariant parameters of the ith subsystem. Specifically, the matrix A_{ij} represents the interconnection between the ith subsystem and the jth subsystem. The matrices A_{ii}, B_i, A_{ij} and M_i are defined respectively as:

$$A_{ii} = \begin{bmatrix} \sum_{j\in\mathcal{N}_i} -\frac{1}{R_{ij}} & \frac{1}{C_{ti}} \\ -\frac{1}{L_{ti}} & -\frac{R_{ti}}{L_{ti}} \end{bmatrix}, \quad B_i = \begin{bmatrix} 0 \\ \frac{1}{L_{ti}} \end{bmatrix},$$
$$A_{ij} = \begin{bmatrix} \frac{1}{R_{ij}C_{ti}} & 0 \\ 0 & 0 \end{bmatrix}, \qquad M_i = \begin{bmatrix} -\frac{1}{C_{ti}} \\ 0 \end{bmatrix}.$$

B. Detection of Stealthy Cyber-attack

Because of the vulnerability of wireless network, the cyber layer is more vulnerable to attack than the physical layer. Each DGU is connected by electrical wire, as opposed to the data transmission between the actuators (sensors) and the DGUs by wireless network. As a result, malicious attackers can tamper with data from actuators or sensors. The above behaviors are called actuator attack and sensor attack such that

$$\mathcal{A}_i : \begin{cases} \widehat{u}_i = u_i + \tilde{u}_i \\ \widehat{y}_i = y_i + \tilde{y}_i \end{cases} \tag{3}$$

where $\mathcal{A}_i$ represents the Cyber-attack on the ith subsystem. $\tilde{u}_i$ and $\tilde{y}_i$ are the false data injected into the ith actuator and sensor respectively. The variable $\widehat{u}_i$ and $\widehat{y}_i$ denote the falsified data.

Assumption 1: There are only actuator attack and sensor attack in the cyber-attack.

Assumption 2: The malicious attacker can attack any actuator or sensor.

Definition 1 (Attack detection): Attack detection is one of binary classification of system into the classes {Normal, Attacked}. The system is normal if $|\tilde{u}|$ and $|\tilde{y}|$ are both equal 0 at all time, where $\tilde{u} = [\tilde{u}_1, \tilde{u}_2, ..., \tilde{u}_N]^T$ and $\tilde{y} = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_N]^T$. Otherwise, the system is attacked.

Definition 2 (Stealthy cyber-attack): Cyber-attack is called stealthy if the magnitude of the injected false data is so small that most detection mechanisms fail to detect its presence, which means that $|\tilde{u}_i|$ and $|\tilde{y}_i|$ are small positive numbers for a period of time.

The purpose of attack monitoring is to design a set of detection mechanisms to judge whether the system has been attacked. This paper contributes a attack detection algorithm based on LSTM neural network to target stealthy cyber-attack effectively.

III. Proposed Algorithm

The proposed attack detection mainly relies on LSTM neural network, which is one of most popular methods of deep-learning. This attack detection algorithm mainly includes two parts. Firstly, we apply a classical unknown input observer (UIO) to obtain state estimate $\hat{x}$ for the distributed DC microgrid. Secondly, a residual term about the state estimation is used as the input to the LSTM neural network.

A. Unknown Input Observer

The model of normal distributed DC microgrids can be represented as:

$$\begin{aligned} \dot{x} &= Ax + Bu + Md + w \\ y &= Hx + v \end{aligned} \tag{4}$$

where $x = [x_1, x_2, ..., x_N]^T$, $u = [u_1, u_2, ..., u_N]^T$, $d = [d_1, d_2, ..., d_N]^T$, $w = [w_1, w_2, ..., w_N]^T$, $v = [v_1, v_2, ..., v_N]^T$ and $y = [y_1, y_2, ..., y_N]^T$. The overall system matrices A, B, C and M are shown below:

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NN} \end{bmatrix}, B = \begin{bmatrix} B_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & B_N \end{bmatrix},$$

$$H = \begin{bmatrix} H_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H_N \end{bmatrix}, \quad M = \begin{bmatrix} M_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_N \end{bmatrix}.$$

In the distributed DC microgrid, d refers to the current created by consumption of energy. Consequently, the exogenous input d is considered as an unknown input, and a classical UIO could be applied to deal to this unknown signal. The state estimate $\hat{x}$ can be acquired from dynamics of the mentioned UIO:

$$UIO : \begin{cases} \dot{z} = Fz + Ty + Gu \\ \hat{x} = z - Ly \end{cases} \tag{5}$$

where z is an intermediate variable to obtain the state estimate $\hat{x}$. In addition to, The parameters of UIO dynamics F, T, G and L are matrices with compatible dimensions, they are designed as following:

$$0 = (I + LH)M \tag{6a}$$

$$G = (I + LH)B \tag{6b}$$

$$F = (I + LH)A - KH \tag{6c}$$

$$K = T + FL \tag{6d}$$

Where matrix F is Hurwitz. Necessary and sufficient conditions for the existence of matrix K are required [19]:

1)The matrix M and the matrix (HM) have full column rank;

2)The pair $((I + HL)A, H)$ is detectable.

We denote the state estimation error as $e = \hat{x} - x$ and its dynamics can be written as $\dot{e} = Fe$ without considering the influence of noises. In this situation, the state estimation error converges exponentially to the origin because matrix F is Hurwitz.

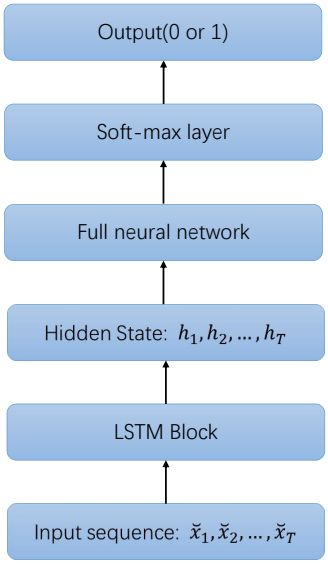

Fig. 2. The LSTM neural network framework

B. LSTM Neural Network

In a typical LSTM neural network architecture, we denote $\left\{\breve{x}_1, \breve{x}_2, ..., \breve{x}_T\right\}$ as an input sequence, where $\breve{x}_t$ represents a vector of real values $(\forall t = 1, 2, ..., T)$. The important elements of LSTM neural network consist of four basic computing units f_t, i_t, g_t, o_t, hidden state h_t and cell state s_t. The formulations are given as follow [20]:

$$f_t = \sigma(W_{xf}\breve{x}_t + W_{hf}h_{t-1} + b_f) \tag{7a}$$

$$i_t = \sigma(W_{xi}\breve{x}_t + W_{hi}h_{t-1} + b_i) \tag{7b}$$

$$g_t = \tanh(W_{xg}\breve{x}_t + W_{hg}h_{t-1} + b_g) \tag{7c}$$

$$o_t = \sigma(W_{xo}\breve{x}_t + W_{ho}h_{t-1} + b_o) \tag{7d}$$

$$h_t = \tanh(s_t) \circ o_t \tag{7e}$$

$$s_t = f_t \circ s_{t-1} + i_t \circ g_t \tag{7f}$$

where the parameters W_{xf}, W_{hf}, W_{xi}, W_{hi}, W_{xg}, W_{hg}, W_{xo} and W_{ho} are weight matrices for the neural network, and the parameters b_f, b_i, b_g and b_o are basics. All of parameters are updated by error back propagation. Symbol $\circ$ represents the element-wise multiplication. σ and $\tanh$ stand for sigmoid and tanh activation function. Sequence $\{\mathrm{h}_1,\mathrm{h}_2,...,\mathrm{h}_T\}$ can be used as the input to next layer of neural network.

As shown in Fig. 2, the proposed LSTM neural network is composed of an LSTM layer, a fully connected neural network and a soft-max layer. Generally speaking, LSTM neural network has an excellent effect on the classification of time series as a classical RNN. In addition, attack detection can be regard as a binary classification problem. Based on the above two cases, LSTM neural network could be used in the attack detection mechanism.

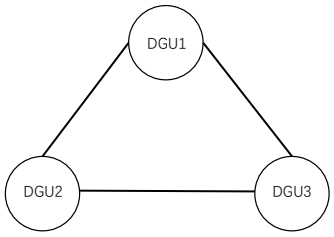

Fig. 3. Scheme of the distributed DC microgrid composed of 3 DGUs.

C. Attack Detection Algorithm

In this section, we propose a new algorithm to detect stealthy cyber-attack. To begin with, let us consider the distributed DC microgrid satisfies condition 1) and condition 2) of the UIO and the state estimation $\hat{x}$ can be obtained. We denote a variable $r := |y - H\hat{x}|$, r can be written as $|Hx - H\hat{x} + v| \leq \|H\|\,|e| + |v|$ under the normal system. The start time of the attack detection is t_0 and r is sampled at time-intervals of T_0. As a result, the sequence $\{r_{t_0}, r_{t_0+T_0}, r_{t_0+2T_0}, ...\}$ is acquired.

The sequence $\{r_{t_0+kT_0}, r_{t_0+(k+1)T_0}, ..., r_{t_0+(k+T)T_0}\}$ is the input of trained LSTM neural network, where k could be any positive integer. If the output of classification is 1, the system is under attacked. Otherwise, the system is normal. These steps are formally stated in Algorithm 1.

Algorithm 1 The proposed attack detection algorithm.

1: Design an UIO for (4);
2: Fix the sampling period T_0, detection start time t_0 and set $k = 0$;
3: **while** $k \geq 0$ **do**
4: Calculate $\{r_{t_0+kT_0}, r_{t_0+(k+1)T_0}, ..., r_{t_0+(k+T)T_0}\}$;
5: Input $\{r_{t_0+kT_0}, r_{t_0+(k+1)T_0}, ..., r_{t_0+(k+T)T_0}\}$ into the trained LSTM neural network;
6: **if** output $= 0$ **then**
7: The system is normal;
8: $k++$;
9: **else**
10: The system is under attack;
11: Break;
12: **end if**
13: **end while**

Remark 2: Any attack detection algorithm is ineffective, while the magnitude of the cyber-attack is negligible compared to the noises.

IV. Case Study

A. Model of the Distributed DC Microgrid

As shown in Fig. 3, The distributed DC microgrid model composed of three DC DGUs connected is considered. We set R_{12}=R_{13}=R_{23}= 2Ω, R_{t1}=R_{t2}=R_{t3}= 0.4Ω, L_{t1}=L_{t2}=L_{t3}= $2mH$ and C_{t1}=C_{t2}=C_{t3}= $2mF$. The measurement matrices are all set to the identity matrix, i.e., H_1=H_2=H_3=I. Each component of process and measurement disturbances obey the uniform distribution between [-0.1, 0.1].

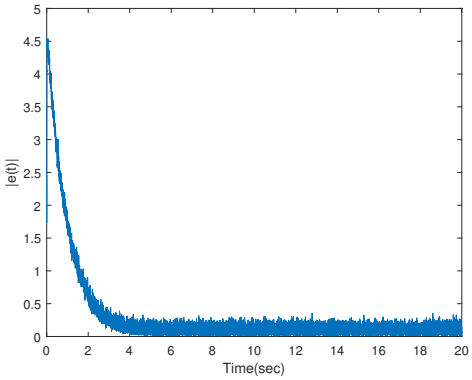

Fig. 4. Euclidean norm of the estimation error when the distributed DC microgrid is normal.

B. Simulation for the UIO

The given distributed DC microgrid satisfies condition 1) and condition 2) of the UIO, which means that there are solutions to the matrix K that makes equations (6a) to (6d) work. The simulation with $K = I$ for the UIO, it is depicted in Fig. 4. It can been seen that the UIO has a perfect effect on state estimation when the magnitude of noises is small. However, the state estimate $\widehat{x}$ can't asymptotically converge to x because of the noises.

C. Training of LSTM Neural Network

The framework of LSTM neural network is built on MATLAB R2020b and Deep Learning Toolbox 14.1 is used. The hardware platform is a desktop PC with 2.70 GHz Intel i7 processor and 8 GB memory.

For the proposed attack detection algorithm, length of the input sequence is $T = 5$ and the number of LSTM cells is selected as 5. LSTM neural network model adapts small batch training method, in which the batch size is 6. The learning rate is defined as $\alpha = 0.0005$. Meanwhile, ADAM optimizer is chosen to train the LSTM neural network model. The LSTM neural network is trained over 1.4×10^4 episodes under the magnitude of cyber-attack is in the range of [0.180,0.215], which are slight.

D. Test Result of the Proposed Attack Detection Algorithm

As shown in Fig. 5, there are eighteen different magnitude of cyber attack in the test. The eighteen sets of data are divided into three groups, the first to sixth tests are the first group, the seventh to twelfth tests are the second group and the thirteenth to eighteenth tests are the third group. Cyber-attack in the first group are almost negligible, cyber-attack in the second group are lightweight and cyber-attack in the third group are of a larger magnitude. It should be emphasized that grouping is based on comparison with the magnitude

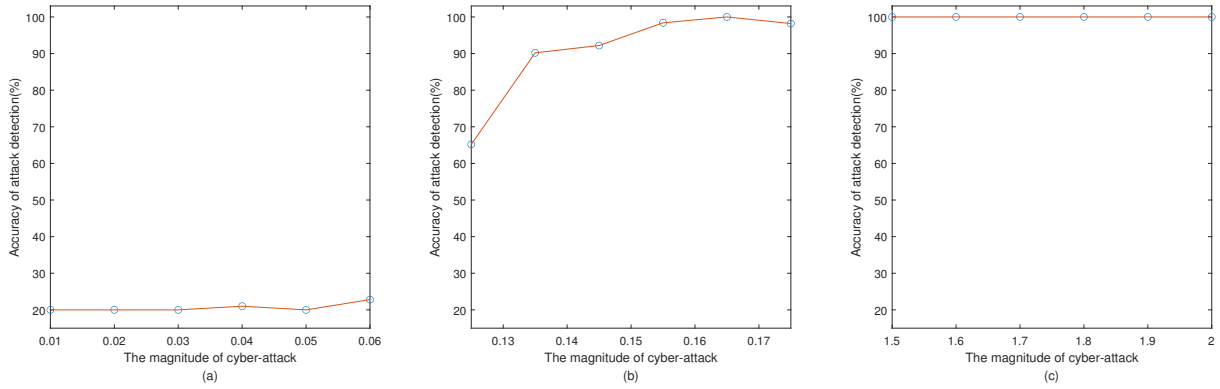

Fig. 5. Accuracy of the proposed attack detection algorithm. In each plot, the magnitude of the cyber-attack is different. (a)The magnitude of cyber-attack is small enough that compared to the impact of noises, it can be ignored. (b)The cyber-attack are stealthy. (c)The magnitude of cyber-attack is large enough to be clearly distinguishable from the noises.

of noises. Each set of tests consists of 500 samples, 100 samples from the normal system and 400 samples from the attacked system. As can be seen from Fig. 5(a), the accuracy of attack detection is about 20% when the cyber attacks are trivial. One-fifth of the samples are from normal systems, which meant that the proposed attack detection algorithm does not detect the presence of cyber-attacks at all. The accuracy of attack detection is greater than 90% when the magnitude of cyber-attack is greater than 0.135 from Fig. 5(b). In other words, the proposed attack detection algorithm is effective when the magnitude of cyber-attack is greater than 0.135. The simulation results in Fig. 5(c) illustrates that if the magnitude of the attack is large enough, which can be clearly distinguished from the noises, accuracy of the attack detection can reach 100%.

V. Conclusion

In this article, a new stealthy cyber-attack detection algorithm, for the distributed DC Microgrid, is proposed. The stealthy cyber-attack detection is based on state estimates, which have two problems. On the one hand, the magnitude of cyber-attacks is lightweight. On the other hand, the statistical characteristics of process and measurement noises are unknown. In order to deal with these problems, the proposed attack detection algorithm, in this paper, combines UIO and LSTM neural network. In the first step, UIO is used to obtain a time series about state estimation. In the second step, this time series is used as the input of LSTM neural network to judge whether the cyber-attack exists.

Acknowledgment

The authors would like to acknowledge Jian Qin for helpful discussions on LSTM neural network.

References

[1] M. Tucci, S. Riverso, J. C. Vasquez, J. M. Guerrero, and G. Ferrari-Trecate, "A decentralized scalable approach to voltage control of dc islanded microgrids," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 6, pp. 1965–1979, 2016.

[2] S. Riverso, F. Sarzo, and G. Ferrari-Trecate, "Plug-and-play voltage and frequency control of islanded microgrids with meshed topology," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1176–1184, 2014.

[3] M. Tucci, S. Riverso, J. C. Vasquez, J. M. Guerrero, and G. Ferrari-Trecate, "Plug-and-play decentralized model predictive control for linear systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2608–2614, 2013.

[4] M. Tucci, S. Riverso, and G. Ferrari-Trecate, "Line-independent plug-and-play controllers for voltage stabilization in dc microgrids," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 3, pp. 1115–1123, 2017.

[5] S. Bansal, M. N. Zeilinger, and C. J. Tomlin, "Plug-and-play model predictive control for electric vehicle charging and voltage control in smart grids," in *53rd IEEE Conference on Decision and Control*. Los Angeles, CA, USA: IEEE, 2014, pp. 5894–5900.

[6] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 20–23, 2015.

[7] A. Barboni, H. Rezaee, F. Boem, and T. Parisini, "Detection of covert cyber-attacks in interconnected systems: a distributed model-based approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3728–3741, 2020.

[8] F. Pasqualetti, F. Drfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[9] Y. Guan and X. Ge, "Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 48–59, 2017.

[10] A. Hoehn and P. Zhang, "Detection of replay attacks in cyber-physical systems," in *2016 American Control Conference*. Boston, MA, USA: IEEE, 2016, pp. 290–295.

[11] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "Coding sensor outputs for injection attacks detection," in *53rd IEEE Conference on Decision and Control*. Los Angeles, CA, USA: IEEE, 2014, pp. 5776–5781.

[12] W. Ao, Y. Song, and C. Wen, "Adaptive cyber-physical system attack detection and reconstruction with application to power systems," *IET Control Theory & Applications*, vol. 10, no. 12, pp. 1458–1468, 2016.

[13] S. Sahoo, S. Mishra, J. C. H. Peng, and Dragievi, "A stealth cyber-attack detection strategy for dc microgrids," *IEEE Transactions on Power Electronics*, vol. 34, no. 8, pp. 8162–8174, 2018.

[14] T. Yang, C. Murguia, M. Kuijper, and Neši, "A multi-observer approach for attack detection and isolation of discrete-time nonlinear systems," in *2018 Australian & New Zealand Control Conference*. Melbourne, VIC, Australia: IEEE, 2018, pp. 346–351.

[15] M. N. Kurt, O. Ogundijo, C. Li, and X. Wang, "Online cyber-attack detection in smart grid: A reinforcement learning approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5174–5185, 2018.

[16] R. K. Mehra and J. Peschon, "An innovations approach to fault detection and diagnosis in dynamic systems," *Automatica*, vol. 7, no. 5, pp. 637–640, 1971.

[17] G. Bishop and G. Welch, "An introduction to the kalman filter," *Proc of SIGGRAPH, Course*, vol. 41, no. 8, pp. 27 599–23 175, 2001.
[18] X. Hu, F. Sun, and Y. Zou, "Estimation of state of charge of a lithium-ion battery pack for electric vehicles using an adaptive luenberger observer," *Energies*, vol. 3, no. 9, pp. 1586–1603, 2010.
[19] M. Darouach, M. Zasadzinski, and S. J. Xu, "Full-order observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 39, no. 3, pp. 606–609, 1994.
[20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[21] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on lstm recurrent neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2017.
[22] D. Pawade, A. Sakhapara, M. Jain, N. Jain, and K. Gada, "Story scrambler-automatic text generation using word level rnn-lstm," *International Journal of Information Technology and Computer Science*, vol. 10, no. 6, pp. 44–53, 2018.
[23] H. Sak, A. W. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Proceedings of the Annual Conference of International Speech Communication Association*, 2014.
[24] M. Sundermeyer, R. Schlter, and H. Ney, "Lstm neural networks for language modeling," in *Thirteenth Annual Conference of The International Speech Communication Association*, 2012.