Deeply Hidden Moving-Target-Defense for Cybersecure Unbalanced Distribution Systems Considering Voltage Stability

Mingjian Cui, *Senior Member, IEEE* and Jianhui Wang, *Senior Member, IEEE*

***Abstract*—A recent proactive defense mechanism, named as moving-target-defense (MTD), has been proposed as a prevailing topic that is capable of actively changing transmission line reactance to preclude cyberattacks. However, the MTD strategy has seldom been studied for the unbalanced AC distribution system in the existing literature. Towards the end, this paper proposes a deeply-hidden MTD (DH-MTD) to elaborately hide both the self and mutual reactance of each phase at the transmission line installed with D-FACTS devices. Both the branch and injection power phasor measurement functions are integrated into DH-MTD in the cyberattack scenario and under the normal operating condition, while the system voltage stability is ensured. The proposed DH-MTD model is solved using a nonlinear least square (NLS) method based on the trust-region algorithm due to the non-Gaussian noise assumption. Also, we cope with the MTD allocation (MTDA) problem using a data-driven normalized PDF peak residual (NPPR) index. The effectiveness of the proposed DH-MTD method is demonstrated in the unbalanced IEEE 123-bus distribution system against both branch and node cyberattacks.**

***Index Terms*—Moving-target-defense, unbalanced distribution system, voltage stability, nonlinear least square.**

I. Introduction

POWER systems are undergoing more and more severe cyberattacks that significantly affect the operating reliability and security [1]. For instance, the Kudankulam nuclear power plant, the newest and largest such power station in India, was hacked using malware designed for data extraction in October 2019 [2]. Furthermore, malicious cyberattacks are threatening the secure and stable operation of distribution systems through the increasing development of micro-phasor measurement units (micro-PMUs) [3]. By injecting false data intelligently designed as a type of malicious cyberattacks to remain stealthy, erroneous measurements may be bypassed if the comprehensive knowledge of the distribution system topology is known to attackers [4].

Recently, as an aspect of defense-in-depth techniques that can increase the cost of cyberattacks for adversaries, the hidden moving target defense (HMTD) could utilize redundancy and diversity of the solution tools and system measurements [5]. The definition of "hidden" MTD means to actively hide parameters of transmission lines by using the distributed flexible AC transmission system (D-FACTS) devices, which can perturb the effective line reactance of transmission lines where they are installed. The concept of "hidden" MTD has already been adopted in several publications. For instance, Tian *et al.* [6] in 2018 defined the "hidden" MTD as changing the line susceptance of transmission lines but also keeping stealthiness even when the attackers are capable of checking the activation of D-FACTS. Liu *et al.* [7] in 2018 defined the "hidden" MTD as optimally changing the branch reactance in AC network to minimize the system loss as well as line power flow differences. The proactive defense mechanism against cyberattacks developed by HMTD is to actively perturb parameters of transmission lines by using D-FACTS devices. This is because D-FACTS can change the effective line reactance of transmission lines on which they are installed [8] that allow producing moving targets defending against cyberattacks.

In the literature, MTD strategies have been proven to be complementary and synergistic with the application in the distribution system state estimation (DSSE) [9]. Tian *et al.* [6] proposed an HMTD approach to maintain the power flow by analyzing the completeness and stealthiness of MTD. Liu *et al.* [7] designed a secure MTD that could maximize the likelihood of an attack detection and identification while minimizing the effect on the operational power loss on transmission lines. Zhang *et al.* [10] minimized the dimension of the stealthy attack space and maximized the number of covered buses by exploiting MTD on an appropriate set of branches. Li *et al.* [11] investigated the possibilities of proactively detecting the high-profile false data injection (FDI) attacks on power grid state estimation by using MTD. Liu *et al.* [12] constructed an HMTD strategy in combination with network reconfiguration by minimizing the system loss and line power flow differences before and after the HMTD. Yao and Li [13] developed a state estimation algorithm based on the selection of random measurements inspired by MTD to prevent and mitigate stealthy cyber-attacks. Lakshminarayana *et al.* [14] invalidated the knowledge that the attackers use to mask the effects of the physical attack by actively performing MTD. Though the general perception points that MTD would not interfere with the functionalities of DSSE algorithms, very few research has utilized MTD for this application in distribution systems, let alone the unbalanced distribution system.

To distinguish our main contributions and differences from the existing literature, Table I compares a variety of MTD strategies that are discussed in the state-of-the-art literature. As shown in this table, most of conventional MTD models focus on the balanced transmission-level DC system based on

M. Cui and J. Wang are with the Department of Electrical and Computer Engineering at Southern Methodist University, Dallas, TX 75275, USA (email: {mingjiancui, jianhui}@smu.edu).

Manuscript received, 2020.

TABLE I
LITERATURE COMPARISON FOR MTD-BASED MODELS

References	Trans./Distr.-level system	Balanced/ Unbalanced	DC/AC model	Gaussian noise assumption
[10], [11], [14], [15], [16]	Trans.-	Balanced	DC	Gaussian
[6], [7]	Trans.-	Balanced	DC+AC	Gaussian
[13]	Trans.-	Balanced	AC	Gaussian
[12]	Distr.-	Balanced	AC	–
This paper	Distr.-	Unbalanced	AC	non-Gaussian

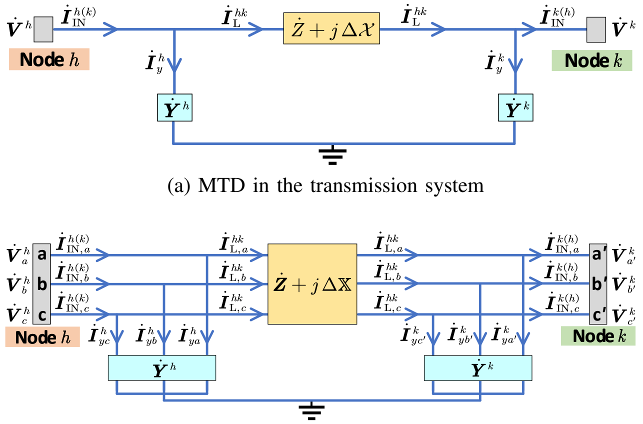
the Gaussian noise assumption. However, the DC system is too idealistic compared with the AC system and may not be applicable in practice. Also, the Gaussian noise assumption is too idealistic as the ambient noise does not always conform to it. Unlike the traditional MTD research, we propose a deeply-hidden MTD (DH-MTD) that is the first of its kind to consider the unbalanced AC distribution systems, while considering the non-Gaussian noise[1].

For the distribution-level system, to the best of our knowledge, only Liu *et al.* [12] explored the HMTD in distribution systems so far. However, they do not extend HMTD to the unbalanced distribution system yet, which inspires our work in this paper. Most of the conventional MTD strategies only consider perturbing the reactance of the transmission-level system with simplification to single-phase equivalents. Such a simplifying hypothesis of the equivalent single-phase model is not reliable for the three-phase unbalanced distribution system. In addition, for the simplified single-phase balanced distribution system, it is too arbitrary to assume that cyberattacks are performed on three phases of the transmission line simultaneously. This is because professional attackers may prefer to stealthily compromise a single phase. This case makes the research of an unbalanced three-phase system more practical than the balanced system [17]. For conventional Arbitrarily-Hidden MTD (AH-MTD) strategies, the simplified single-phase reactance perturbation (RP) is uniform but arbitrary from the perspective of three phases on the transmission line. Towards this end, we propose a DH-MTD strategy to elaborately hide both self and mutual reactance of the three-phase transmission line. The main contributions and novelties of this paper include:

(i) The proposed DH-MTD is the first of its kind that considers the unbalanced AC distribution systems. Based on this, the system voltage unbalancing status can be achieved in the cyberattack scenario, while the voltage stability is ensured.

(ii) An MTD model that does not require any Gaussian measurement noise assumption is proposed. Unlike the traditional weighted least square (WLS) algorithm that depends on the Gaussian noise, the proposed DH-MTD can effectively process the non-Gaussian noise (including Gaussian noise that is set as a special form of non-Gaussian noise in this paper).

(iii) A data-driven MTD allocation (MTDA) method is proposed to choose the exact transmission line to perturb its reactance. A prior-information-based metric is designed to solve the MTDA problem.



$\dot{V}^h$
$\dot{I}^{h(k)}_{\mathrm{IN}}$
$\dot{I}^{hk}_{\mathrm{L}}$
$\dot{Z}+j\Delta\mathcal{X}$
$\dot{I}^{hk}_{\mathrm{L}}$
$\dot{I}^{k(h)}_{\mathrm{IN}}$
$\dot{V}^k$
Node h
$\dot{I}^h_y$
$\dot{I}^k_y$
Node k
$\dot{Y}^h$
$\dot{Y}^k$
$\dot{V}^h_a$
$\dot{V}^h_b$
$\dot{V}^h_c$
a
b
c
$\dot{I}^{h(k)}_{\mathrm{IN},a}$
$\dot{I}^{h(k)}_{\mathrm{IN},b}$
$\dot{I}^{h(k)}_{\mathrm{IN},c}$
$\dot{I}^{hk}_{\mathrm{L},a}$
$\dot{I}^{hk}_{\mathrm{L},b}$
$\dot{I}^{hk}_{\mathrm{L},c}$
$\dot{Z}+j\Delta\mathbb{X}$
$\dot{I}^{hk}_{\mathrm{L},a}$
$\dot{I}^{hk}_{\mathrm{L},b}$
$\dot{I}^{hk}_{\mathrm{L},c}$
$\dot{I}^{k(h)}_{\mathrm{IN},a}$
$\dot{I}^{k(h)}_{\mathrm{IN},b}$
$\dot{I}^{k(h)}_{\mathrm{IN},c}$
a'
b'
c'
$\dot{V}^k_{a'}$
$\dot{V}^k_{b'}$
$\dot{V}^k_{c'}$
Node h
$\dot{I}^h_{yc}$
$\dot{I}^h_{yb}$
$\dot{I}^h_{ya}$
$\dot{I}^k_{yc'}$
$\dot{I}^k_{yb'}$
$\dot{I}^k_{ya'}$
Node k
$\dot{Y}^h$
$\dot{Y}^k$

(a) MTD in the transmission system

(b) MTD in the unbalanced distribution system

Fig. 1. PI models of single-line and three-phase branches with the complex scalar RP $\Delta\mathcal{X}$ and the complex matrix RP $\Delta\mathbb{X}$ in the transmission and distribution systems.

The organization of this paper is as follows. In Section II, the theoretical methodology of the proposed DH-MTD and its solution are introduced. The data-driven DH-MTD allocation method is described in Section III. Evaluation metrics are discussed in Section IV. Case studies and result analysis performed on the unbalanced IEEE 123-bus distribution system are discussed in Section V. Concluding remarks are summarized in Section VII.

II. Theoretical Methodology and Solution

For the current transmission system, only a single-line diagram instead of a three-line diagram is modeled for applying MTD. For the distribution system in this paper, we model the three-phase unbalanced branch by separately adding the reactance to each line parameter. To illustrate the main distinction of MTD in the transmission and distribution systems, Fig. 1 shows PI models of single-line and three-phase branches with the complex scalar RP $\Delta\mathcal{X}$ and the complex matrix RP $\Delta\mathbb{X}$. In the transmission system, existing MTD models [6], [11], [12] consider a complex scalar RP $\Delta\mathcal{X}$ in a single-line diagram. However, in the unbalanced three-phase distribution system, the single-line diagram cannot work and more detailed information of branch parameters of each phase is highly desired. Thus, the complex matrix RP $\Delta\mathbb{X}$ with both the self and mutual reactance must be elaborately designed.

A. Deeply-Hidden MTD Modeling

In this section, we describe the mathematical model of the proposed DH-MTD to defend against cyberattacks in unbalanced distribution systems. Based on the MTD definition [7], we introduce an RP matrix ($\Delta\mathbb{X}$) that can be actively perturbed by operators through D-FACTS installed on the branch. Fig. 1b shows PI models of single-line and three-phase branches with the complex scalar RP $\Delta\mathcal{X}$ and the complex matrix RP $\Delta\mathbb{X}$.

[1] Note that as the Gaussian noise is taken as a special form of the non-Gaussian noise in this paper, it also applies to the proposed MTD model.

Under previous normal operating condition, the branch and injection power phasor measurements $[\dot{\boldsymbol{S}}^{hk}_{\mathrm{L},a}, \dot{\boldsymbol{S}}^{hk}_{\mathrm{L},b}, \dot{\boldsymbol{S}}^{hk}_{\mathrm{L},c}]^{\mathrm{T}}$ at branch $h{\to}k$ and $[\dot{\boldsymbol{S}}^{h}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}^{h}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}^{h}_{\mathrm{IN},c}]^{\mathrm{T}}$ at node h are given by:

$$\begin{bmatrix}\dot{\boldsymbol{S}}^{hk}_{\mathrm{L},a}\\ \dot{\boldsymbol{S}}^{hk}_{\mathrm{L},b}\\ \dot{\boldsymbol{S}}^{hk}_{\mathrm{L},c}\end{bmatrix} = \begin{bmatrix}\dot{\boldsymbol{V}}^{h}_{a}\\ \dot{\boldsymbol{V}}^{h}_{b}\\ \dot{\boldsymbol{V}}^{h}_{c}\end{bmatrix} \begin{bmatrix}\dot{\boldsymbol{I}}^{hk}_{\mathrm{L},a}\\ \dot{\boldsymbol{I}}^{hk}_{\mathrm{L},b}\\ \dot{\boldsymbol{I}}^{hk}_{\mathrm{L},c}\end{bmatrix}^{*} + \dot{\varepsilon}_{\mathrm{L}}, \quad \begin{cases}k\in h\in\mathcal{N}\\ h\neq k\end{cases} \tag{1}$$

$$\begin{bmatrix}\dot{\boldsymbol{S}}^{h}_{\mathrm{IN},a}\\ \dot{\boldsymbol{S}}^{h}_{\mathrm{IN},b}\\ \dot{\boldsymbol{S}}^{h}_{\mathrm{IN},c}\end{bmatrix} = \begin{bmatrix}\dot{\boldsymbol{V}}^{h}_{a}\\ \dot{\boldsymbol{V}}^{h}_{b}\\ \dot{\boldsymbol{V}}^{h}_{c}\end{bmatrix} \begin{bmatrix}\dot{\boldsymbol{I}}^{h}_{\mathrm{IN},a}\\ \dot{\boldsymbol{I}}^{h}_{\mathrm{IN},b}\\ \dot{\boldsymbol{I}}^{h}_{\mathrm{IN},c}\end{bmatrix}^{*} + \dot{\varepsilon}_{\mathrm{IN}}, \quad h\in\mathcal{N} \tag{2}$$

where $[\dot{\boldsymbol{V}}^{h}_{a}, \dot{\boldsymbol{V}}^{h}_{b}, \dot{\boldsymbol{V}}^{h}_{c}]^{\mathrm{T}}$ is the three-phase voltage phasor at node h in the previous normal operation scenario. The branch and injection current phasors $[\dot{\boldsymbol{I}}^{hk}_{\mathrm{L},a}, \dot{\boldsymbol{I}}^{hk}_{\mathrm{L},b}, \dot{\boldsymbol{I}}^{hk}_{\mathrm{L},c}]^{\mathrm{T}}$ and $[\dot{\boldsymbol{I}}^{h}_{\mathrm{IN},a}, \dot{\boldsymbol{I}}^{h}_{\mathrm{IN},b}, \dot{\boldsymbol{I}}^{h}_{\mathrm{IN},c}]^{\mathrm{T}}$ are given in (3) and (4). $\dot{\varepsilon}_{\mathrm{L}}$ and $\dot{\varepsilon}_{\mathrm{IN}}$ are the measurement noise phasors. $\mathcal{N}$ is the set of system nodes. $(k{\in}h{\in}\mathcal{N}){\cap}(k{\neq}h)$ means any node k connected to node h. $\dot{\boldsymbol{Z}}$ is the three-phase impedance matrix of branch $h{\to}k$. $\dot{\boldsymbol{Y}}$ is the three-phase shunt admittance matrix at node h. $\lambda_h{=}{+}0.5$ if the branch current is measured from node h and $\lambda_h{=}{-}0.5$ if the branch current is measured from node k. The symbol of '$[\cdot]^*$' denotes the conjugate matrix. Eqs. (1) and (2) can be generalized as:

$$[\dot{\boldsymbol{S}}_{\mathrm{L},a}, \dot{\boldsymbol{S}}_{\mathrm{L},b}, \dot{\boldsymbol{S}}_{\mathrm{L},c}]^{\mathrm{T}} = \mathbf{f}_{\mathrm{L}}(\dot{\boldsymbol{V}}_{a}, \dot{\boldsymbol{V}}_{b}, \dot{\boldsymbol{V}}_{c}) + \dot{\varepsilon}_{\mathrm{L}} \tag{5}$$

$$[\dot{\boldsymbol{S}}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}_{\mathrm{IN},c}]^{\mathrm{T}} = \mathbf{f}_{\mathrm{IN}}(\dot{\boldsymbol{V}}_{a}, \dot{\boldsymbol{V}}_{b}, \dot{\boldsymbol{V}}_{c}) + \dot{\varepsilon}_{\mathrm{IN}} \tag{6}$$

In the current cyberattack scenario with the RP matrix $\Delta\mathbb{X}$, the compromised measurements of branch and injection power phasors $[\dot{\boldsymbol{S}}'^{hk}_{\mathrm{L},a}, \dot{\boldsymbol{S}}'^{hk}_{\mathrm{L},b}, \dot{\boldsymbol{S}}'^{hk}_{\mathrm{L},c}]^{\mathrm{T}}$ at branch $h \to k$ and $[\dot{\boldsymbol{S}}'^{h}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'^{h}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}'^{h}_{\mathrm{IN},c}]^{\mathrm{T}}$ at node h are given by:

$$\begin{bmatrix}\dot{\boldsymbol{S}}'^{hk}_{\mathrm{L},a}\\ \dot{\boldsymbol{S}}'^{hk}_{\mathrm{L},b}\\ \dot{\boldsymbol{S}}'^{hk}_{\mathrm{L},c}\end{bmatrix} = \begin{bmatrix}\dot{\boldsymbol{U}}^{h}_{a}\\ \dot{\boldsymbol{U}}^{h}_{b}\\ \dot{\boldsymbol{U}}^{h}_{c}\end{bmatrix} \begin{bmatrix}\dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},a}\\ \dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},b}\\ \dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},c}\end{bmatrix}^{*} + \dot{\epsilon}_{\mathrm{L}}, \quad \begin{cases}k\in h\in\mathcal{N}\\ h\neq k\end{cases} \tag{7}$$

$$\begin{bmatrix}\dot{\boldsymbol{S}}'^{h}_{\mathrm{IN},a}\\ \dot{\boldsymbol{S}}'^{h}_{\mathrm{IN},b}\\ \dot{\boldsymbol{S}}'^{h}_{\mathrm{IN},c}\end{bmatrix} = \begin{bmatrix}\dot{\boldsymbol{U}}^{h}_{a}\\ \dot{\boldsymbol{U}}^{h}_{b}\\ \dot{\boldsymbol{U}}^{h}_{c}\end{bmatrix} \begin{bmatrix}\dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},a}\\ \dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},b}\\ \dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},c}\end{bmatrix}^{*} + \dot{\epsilon}_{\mathrm{IN}}, \quad h\in\mathcal{N} \tag{8}$$

where the compromised branch and injection current phasors $[\dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},a}, \dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},b}, \dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},c}]^{\mathrm{T}}$ and $[\dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},a}, \dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},b}, \dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},c}]^{\mathrm{T}}$ are given in (9) and (10). $[\dot{\boldsymbol{U}}^{h}_{a}, \dot{\boldsymbol{U}}^{h}_{b}, \dot{\boldsymbol{U}}^{h}_{c}]^{\mathrm{T}}$ is the three-phase voltage phasor of node h in the current cyberattack scenario. $\dot{\epsilon}_{\mathrm{L}}$ and $\dot{\epsilon}_{\mathrm{IN}}$ are the measurement noise phasors. Eqs. (7) and (8) can be generalized as:

$$[\dot{\boldsymbol{S}}'_{\mathrm{L},a}, \dot{\boldsymbol{S}}'_{\mathrm{L},b}, \dot{\boldsymbol{S}}'_{\mathrm{L},c}]^{\mathrm{T}} = \mathbf{h}_{\mathrm{L}}(\dot{\boldsymbol{U}}_{a}, \dot{\boldsymbol{U}}_{b}, \dot{\boldsymbol{U}}_{c}, \Delta\mathbb{X}) + \dot{\epsilon}_{\mathrm{L}} \tag{11}$$

$$[\dot{\boldsymbol{S}}'_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'_{\mathrm{IN},b}, \dot{\boldsymbol{S}}'_{\mathrm{IN},c}]^{\mathrm{T}} = \mathbf{h}_{\mathrm{IN}}(\dot{\boldsymbol{U}}_{a}, \dot{\boldsymbol{U}}_{b}, \dot{\boldsymbol{U}}_{c}, \Delta\mathbb{X}) + \dot{\epsilon}_{\mathrm{IN}} \tag{12}$$

To maintain the system voltage stability, the voltage changes should be sufficiently small between voltage magnitude variables in the normal scenario and those in the cyberattack scenario, given by:

$$\begin{bmatrix}0\\0\\0\end{bmatrix} = \begin{bmatrix}|\dot{\boldsymbol{U}}^{h}_{a}|\\ |\dot{\boldsymbol{U}}^{h}_{b}|\\ |\dot{\boldsymbol{U}}^{h}_{c}|\end{bmatrix} - \begin{bmatrix}|\dot{\boldsymbol{V}}^{h}_{a}|\\ |\dot{\boldsymbol{V}}^{h}_{b}|\\ |\dot{\boldsymbol{V}}^{h}_{c}|\end{bmatrix} + \dot{\varepsilon}_{\mathrm{V}}, \ h\in\mathcal{N} \tag{13}$$

where $|\cdot|$ is the magnitude of a phasor variable. $\dot{\varepsilon}_{\mathrm{V}}$ is the measurement noise phasor. It can be generalized as:

$$[0,0,0]^{\mathrm{T}} = \mathbf{f}_{\mathrm{V}}(\dot{\boldsymbol{V}}_{a}, \dot{\boldsymbol{V}}_{b}, \dot{\boldsymbol{V}}_{c}, \dot{\boldsymbol{U}}_{a}, \dot{\boldsymbol{U}}_{b}, \dot{\boldsymbol{U}}_{c}) + \dot{\varepsilon}_{\mathrm{V}} \tag{14}$$

By integrating (1)–(14) into a matrix form, we can get:

$$\begin{bmatrix}\left[\dot{\boldsymbol{S}}_{\mathrm{L},a}, \dot{\boldsymbol{S}}_{\mathrm{L},b}, \dot{\boldsymbol{S}}_{\mathrm{L},c}\right]^{\mathrm{T}}\\ \left[\dot{\boldsymbol{S}}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}_{\mathrm{IN},c}\right]^{\mathrm{T}}\\ \left[\dot{\boldsymbol{S}}'_{\mathrm{L},a}, \dot{\boldsymbol{S}}'_{\mathrm{L},b}, \dot{\boldsymbol{S}}'_{\mathrm{L},c}\right]^{\mathrm{T}}\\ \left[\dot{\boldsymbol{S}}'_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'_{\mathrm{IN},b}, \dot{\boldsymbol{S}}'_{\mathrm{IN},c}\right]^{\mathrm{T}}\\ [0,0,0]^{\mathrm{T}}\end{bmatrix} = \begin{bmatrix}\mathbf{f}_{\mathrm{L}}\left(\dot{\boldsymbol{V}}_{a}, \dot{\boldsymbol{V}}_{b}, \dot{\boldsymbol{V}}_{c}\right)\\ \mathbf{f}_{\mathrm{IN}}\left(\dot{\boldsymbol{V}}_{a}, \dot{\boldsymbol{V}}_{b}, \dot{\boldsymbol{V}}_{c}\right)\\ \mathbf{h}_{\mathrm{L}}\left(\dot{\boldsymbol{U}}_{a}, \dot{\boldsymbol{U}}_{b}, \dot{\boldsymbol{U}}_{c}, \Delta\mathbb{X}\right)\\ \mathbf{h}_{\mathrm{IN}}\left(\dot{\boldsymbol{U}}_{a}, \dot{\boldsymbol{U}}_{b}, \dot{\boldsymbol{U}}_{c}, \Delta\mathbb{X}\right)\\ \mathbf{f}_{\mathrm{V}}\left(\dot{\boldsymbol{V}}_{a}, \dot{\boldsymbol{V}}_{b}, \dot{\boldsymbol{V}}_{c}, \dot{\boldsymbol{U}}_{a}, \dot{\boldsymbol{U}}_{b}, \dot{\boldsymbol{U}}_{c}\right)\end{bmatrix} + \begin{bmatrix}\dot{\varepsilon}_{\mathrm{L}}\\ \dot{\varepsilon}_{\mathrm{IN}}\\ \dot{\epsilon}_{\mathrm{L}}\\ \dot{\epsilon}_{\mathrm{IN}}\\ \dot{\varepsilon}_{\mathrm{V}}\end{bmatrix} \tag{15}$$

where $\mathbf{f}_{\mathrm{L}}$, $\mathbf{f}_{\mathrm{IN}}$, $\mathbf{f}_{\mathrm{V}}$, $\mathbf{h}_{\mathrm{L}}$, and $\mathbf{h}_{\mathrm{IN}}$ are generalized function vectors of state variables and measurement equations.

Finally, we can obtain the generalized mathematical MTD model $\mathbf{G}$ of the three-phase distribution system, given by:

$$\mathbf{Y} = \mathbf{G}(\mathbf{X}) + \varepsilon \tag{16}$$

where $\mathbf{X}$ is the set of estimated variables including state variables and the RP matrix. $\mathbf{Y}$ is the set of measurements. ε is the set of measurement noises. $\mathbf{X}$ and $\mathbf{Y}$ are given by:

$$\begin{aligned}\mathbf{X} = \Big[&\dot{\boldsymbol{V}}^{1}_{a}, \dot{\boldsymbol{V}}^{1}_{b}, \dot{\boldsymbol{V}}^{1}_{c}, \dot{\boldsymbol{V}}^{2}_{a}, \dot{\boldsymbol{V}}^{2}_{b}, \dot{\boldsymbol{V}}^{2}_{c}, \cdots, \dot{\boldsymbol{V}}^{N}_{a}, \dot{\boldsymbol{V}}^{N}_{b}, \dot{\boldsymbol{V}}^{N}_{c};\\ &\dot{\boldsymbol{U}}^{1}_{a}, \dot{\boldsymbol{U}}^{1}_{b}, \dot{\boldsymbol{U}}^{1}_{c}, \dot{\boldsymbol{U}}^{2}_{a}, \dot{\boldsymbol{U}}^{2}_{b}, \dot{\boldsymbol{U}}^{2}_{c}, \cdots, \dot{\boldsymbol{U}}^{N}_{a}, \dot{\boldsymbol{U}}^{N}_{b}, \dot{\boldsymbol{U}}^{N}_{c}; \Delta\mathbb{X}\Big]\end{aligned} \tag{17}$$

$$\begin{aligned}\mathbf{Y} = \Big[&\dot{\boldsymbol{S}}^{1}_{\mathrm{L},a}, \dot{\boldsymbol{S}}^{1}_{\mathrm{L},b}, \dot{\boldsymbol{S}}^{1}_{\mathrm{L},c}, \dot{\boldsymbol{S}}^{2}_{\mathrm{L},a}, \dot{\boldsymbol{S}}^{2}_{\mathrm{L},b}, \dot{\boldsymbol{S}}^{2}_{\mathrm{L},c}, \cdots, \dot{\boldsymbol{S}}^{M}_{\mathrm{L},a}, \dot{\boldsymbol{S}}^{M}_{\mathrm{L},b}, \dot{\boldsymbol{S}}^{M}_{\mathrm{L},c};\\ &\dot{\boldsymbol{S}}^{1}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}^{1}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}^{1}_{\mathrm{IN},c}, \dot{\boldsymbol{S}}^{2}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}^{2}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}^{2}_{\mathrm{IN},c}, \cdots, \dot{\boldsymbol{S}}^{N}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}^{N}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}^{N}_{\mathrm{IN},c};\\ &0^{1}_{a}, 0^{1}_{b}, 0^{1}_{c}, 0^{2}_{a}, 0^{2}_{b}, 0^{2}_{c}, \cdots, 0^{N}_{a}, 0^{N}_{b}, 0^{N}_{c}; \dot{\boldsymbol{S}}'^{1}_{\mathrm{L},a}, \dot{\boldsymbol{S}}'^{1}_{\mathrm{L},b}, \dot{\boldsymbol{S}}'^{1}_{\mathrm{L},c},\\ &\dot{\boldsymbol{S}}'^{2}_{\mathrm{L},a}, \dot{\boldsymbol{S}}'^{2}_{\mathrm{L},b}, \dot{\boldsymbol{S}}'^{2}_{\mathrm{L},c}, \cdots, \dot{\boldsymbol{S}}'^{M}_{\mathrm{L},a}, \dot{\boldsymbol{S}}'^{M}_{\mathrm{L},b}, \dot{\boldsymbol{S}}'^{M}_{\mathrm{L},c}; \dot{\boldsymbol{S}}'^{1}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'^{1}_{\mathrm{IN},b},\\ &\dot{\boldsymbol{S}}'^{1}_{\mathrm{IN},c}, \dot{\boldsymbol{S}}'^{2}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'^{2}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}'^{2}_{\mathrm{IN},c}, \cdots, \dot{\boldsymbol{S}}'^{N}_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'^{N}_{\mathrm{IN},b}, \dot{\boldsymbol{S}}'^{N}_{\mathrm{IN},c}\Big]\end{aligned} \tag{18}$$

where N is the total number of system nodes. M is the total number of system branches.

As non-Gaussian noises are assumed beforehand, the conventional weighted least square (WLS) method is not applicable to the aforementioned DH-MTD model. Thus, the nonlinear least square (NLS) method is used to find the minimum errors between estimated measurements and original attacked measurements. The objective function $\mathbf{f}(\mathbf{X})$ of DH-MDT is given by:

$$\widetilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \mathbf{f}(\mathbf{X}) = \arg\min_{\mathbf{X}} [\mathbf{Y} - \mathbf{G}(\mathbf{X})]^{\mathrm{T}} \cdot [\mathbf{Y} - \mathbf{G}(\mathbf{X})] \tag{19}$$

$$\begin{bmatrix}\dot{\boldsymbol{I}}_{\mathrm{L},a}^{hk}\\ \dot{\boldsymbol{I}}_{\mathrm{L},b}^{hk}\\ \dot{\boldsymbol{I}}_{\mathrm{L},c}^{hk}\end{bmatrix}=\underbrace{\begin{bmatrix}\dot{\boldsymbol{z}}_{aa'}^{hk} & \dot{\boldsymbol{z}}_{ab'}^{hk} & \dot{\boldsymbol{z}}_{ac'}^{hk}\\ \dot{\boldsymbol{z}}_{ba'}^{hk} & \dot{\boldsymbol{z}}_{bb'}^{hk} & \dot{\boldsymbol{z}}_{bc'}^{hk}\\ \dot{\boldsymbol{z}}_{ca'}^{hk} & \dot{\boldsymbol{z}}_{cb'}^{hk} & \dot{\boldsymbol{z}}_{cc'}^{hk}\end{bmatrix}}_{\dot{\boldsymbol{Z}}}^{-1}\left(\begin{bmatrix}\dot{\boldsymbol{V}}_{a}^{h}\\ \dot{\boldsymbol{V}}_{b}^{h}\\ \dot{\boldsymbol{V}}_{c}^{h}\end{bmatrix}-\begin{bmatrix}\dot{\boldsymbol{V}}_{a'}^{k}\\ \dot{\boldsymbol{V}}_{b'}^{k}\\ \dot{\boldsymbol{V}}_{c'}^{k}\end{bmatrix}\right)+\lambda_h\underbrace{\begin{bmatrix}\dot{\boldsymbol{y}}_{aa'}^{h} & \dot{\boldsymbol{y}}_{ab'}^{h} & \dot{\boldsymbol{y}}_{ac'}^{h}\\ \dot{\boldsymbol{y}}_{ba'}^{h} & \dot{\boldsymbol{y}}_{bb'}^{h} & \dot{\boldsymbol{y}}_{bc'}^{h}\\ \dot{\boldsymbol{y}}_{ca'}^{h} & \dot{\boldsymbol{y}}_{cb'}^{h} & \dot{\boldsymbol{y}}_{cc'}^{h}\end{bmatrix}}_{\dot{\boldsymbol{Y}}}\begin{bmatrix}\dot{\boldsymbol{V}}_{a}^{h}\\ \dot{\boldsymbol{V}}_{b}^{h}\\ \dot{\boldsymbol{V}}_{c}^{h}\end{bmatrix},\ \begin{cases}k\in h\in\mathcal{N}\\ h\neq k\end{cases}\tag{3}$$

$$\begin{bmatrix}\dot{\boldsymbol{I}}_{\mathrm{IN},a}^{h}\\ \dot{\boldsymbol{I}}_{\mathrm{IN},b}^{h}\\ \dot{\boldsymbol{I}}_{\mathrm{IN},c}^{h}\end{bmatrix}=\sum_{k\in h}\left\{\lambda_h\underbrace{\begin{bmatrix}\dot{\boldsymbol{y}}_{aa'}^{h} & \dot{\boldsymbol{y}}_{ab'}^{h} & \dot{\boldsymbol{y}}_{ac'}^{h}\\ \dot{\boldsymbol{y}}_{ba'}^{h} & \dot{\boldsymbol{y}}_{bb'}^{h} & \dot{\boldsymbol{y}}_{bc'}^{h}\\ \dot{\boldsymbol{y}}_{ca'}^{h} & \dot{\boldsymbol{y}}_{cb'}^{h} & \dot{\boldsymbol{y}}_{cc'}^{h}\end{bmatrix}}_{\dot{\boldsymbol{Y}}}\begin{bmatrix}\dot{\boldsymbol{V}}_{a}^{h}\\ \dot{\boldsymbol{V}}_{b}^{h}\\ \dot{\boldsymbol{V}}_{c}^{h}\end{bmatrix}+\underbrace{\begin{bmatrix}\dot{\boldsymbol{z}}_{aa'}^{hk} & \dot{\boldsymbol{z}}_{ab'}^{hk} & \dot{\boldsymbol{z}}_{ac'}^{hk}\\ \dot{\boldsymbol{z}}_{ba'}^{hk} & \dot{\boldsymbol{z}}_{bb'}^{hk} & \dot{\boldsymbol{z}}_{bc'}^{hk}\\ \dot{\boldsymbol{z}}_{ca'}^{hk} & \dot{\boldsymbol{z}}_{cb'}^{hk} & \dot{\boldsymbol{z}}_{cc'}^{hk}\end{bmatrix}}_{\dot{\boldsymbol{Z}}}^{-1}\left(\begin{bmatrix}\dot{\boldsymbol{V}}_{a}^{h}\\ \dot{\boldsymbol{V}}_{b}^{h}\\ \dot{\boldsymbol{V}}_{c}^{h}\end{bmatrix}-\begin{bmatrix}\dot{\boldsymbol{V}}_{a'}^{k}\\ \dot{\boldsymbol{V}}_{b'}^{k}\\ \dot{\boldsymbol{V}}_{c'}^{k}\end{bmatrix}\right)\right\},\ \begin{cases}k\in h\in\mathcal{N}\\ h\neq k\end{cases}\tag{4}$$

$$\begin{bmatrix}\dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},a}\\ \dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},b}\\ \dot{\boldsymbol{I}}'^{hk}_{\mathrm{L},c}\end{bmatrix}=\left\{\underbrace{\begin{bmatrix}\dot{\boldsymbol{z}}_{aa'}^{hk} & \dot{\boldsymbol{z}}_{ab'}^{hk} & \dot{\boldsymbol{z}}_{ac'}^{hk}\\ \dot{\boldsymbol{z}}_{ba'}^{hk} & \dot{\boldsymbol{z}}_{bb'}^{hk} & \dot{\boldsymbol{z}}_{bc'}^{hk}\\ \dot{\boldsymbol{z}}_{ca'}^{hk} & \dot{\boldsymbol{z}}_{cb'}^{hk} & \dot{\boldsymbol{z}}_{cc'}^{hk}\end{bmatrix}}_{\dot{\boldsymbol{Z}}}+j\Delta\mathbb{X}\right\}^{-1}\left(\begin{bmatrix}\dot{\boldsymbol{U}}_{a}^{h}\\ \dot{\boldsymbol{U}}_{b}^{h}\\ \dot{\boldsymbol{U}}_{c}^{h}\end{bmatrix}-\begin{bmatrix}\dot{\boldsymbol{U}}_{a'}^{k}\\ \dot{\boldsymbol{U}}_{b'}^{k}\\ \dot{\boldsymbol{U}}_{c'}^{k}\end{bmatrix}\right)+\lambda_h\underbrace{\begin{bmatrix}\dot{\boldsymbol{y}}_{aa'}^{h} & \dot{\boldsymbol{y}}_{ab'}^{h} & \dot{\boldsymbol{y}}_{ac'}^{h}\\ \dot{\boldsymbol{y}}_{ba'}^{h} & \dot{\boldsymbol{y}}_{bb'}^{h} & \dot{\boldsymbol{y}}_{bc'}^{h}\\ \dot{\boldsymbol{y}}_{ca'}^{h} & \dot{\boldsymbol{y}}_{cb'}^{h} & \dot{\boldsymbol{y}}_{cc'}^{h}\end{bmatrix}}_{\dot{\boldsymbol{Y}}}\begin{bmatrix}\dot{\boldsymbol{U}}_{a}^{h}\\ \dot{\boldsymbol{U}}_{b}^{h}\\ \dot{\boldsymbol{U}}_{c}^{h}\end{bmatrix},\ \begin{cases}k\in h\in\mathcal{N}\\ h\neq k\end{cases}\tag{9}$$

$$\begin{bmatrix}\dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},a}\\ \dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},b}\\ \dot{\boldsymbol{I}}'^{h}_{\mathrm{IN},c}\end{bmatrix}=\sum_{k\in h}\left\{\lambda_h\underbrace{\begin{bmatrix}\dot{\boldsymbol{y}}_{aa'}^{h} & \dot{\boldsymbol{y}}_{ab'}^{h} & \dot{\boldsymbol{y}}_{ac'}^{h}\\ \dot{\boldsymbol{y}}_{ba'}^{h} & \dot{\boldsymbol{y}}_{bb'}^{h} & \dot{\boldsymbol{y}}_{bc'}^{h}\\ \dot{\boldsymbol{y}}_{ca'}^{h} & \dot{\boldsymbol{y}}_{cb'}^{h} & \dot{\boldsymbol{y}}_{cc'}^{h}\end{bmatrix}}_{\dot{\boldsymbol{Y}}}\begin{bmatrix}\dot{\boldsymbol{U}}_{a}^{h}\\ \dot{\boldsymbol{U}}_{b}^{h}\\ \dot{\boldsymbol{U}}_{c}^{h}\end{bmatrix}+\left\{\underbrace{\begin{bmatrix}\dot{\boldsymbol{z}}_{aa'}^{hk} & \dot{\boldsymbol{z}}_{ab'}^{hk} & \dot{\boldsymbol{z}}_{ac'}^{hk}\\ \dot{\boldsymbol{z}}_{ba'}^{hk} & \dot{\boldsymbol{z}}_{bb'}^{hk} & \dot{\boldsymbol{z}}_{bc'}^{hk}\\ \dot{\boldsymbol{z}}_{ca'}^{hk} & \dot{\boldsymbol{z}}_{cb'}^{hk} & \dot{\boldsymbol{z}}_{cc'}^{hk}\end{bmatrix}}_{\dot{\boldsymbol{Z}}}+j\Delta\mathbb{X}\right\}^{-1}\left(\begin{bmatrix}\dot{\boldsymbol{U}}_{a}^{h}\\ \dot{\boldsymbol{U}}_{b}^{h}\\ \dot{\boldsymbol{U}}_{c}^{h}\end{bmatrix}-\begin{bmatrix}\dot{\boldsymbol{U}}_{a'}^{k}\\ \dot{\boldsymbol{U}}_{b'}^{k}\\ \dot{\boldsymbol{U}}_{c'}^{k}\end{bmatrix}\right)\right\},\ \begin{cases}k\in h\in\mathcal{N}\\ h\neq k\end{cases}\tag{10}$$

B. Discussion of RP Matrix $\Delta\mathbb{X}$ for Comparing the Difference of MTD Strategies

For the conventional MTD, each element of RP matrix $\Delta\mathbb{X}$ is up to $\pm(10\%{\sim}20\%)$ of the line reactance X [7], [8] with a uniform but arbitrary coefficient β_{H}. For the proposed DH-MTD that can deeply hide reactance changes, we elaborate a set of six separate and independent RP coefficients (i.e., $\{\beta_1,\cdots,\beta_6\}$) by multiplying the self and mutual reactance at each branch installed with D-FACTS. As the impedance matrix $\dot{\boldsymbol{Z}}$ is symmetric, the RP matrix $\Delta\mathbb{X}$ is also set as a symmetric matrix to deeply perturb cyberattackers. Taking one branch $h{\to}k$ installed with D-FACTS as an example, the RP matrix $\Delta\mathbb{X}$ is a $3{\times}3$ matrix. We only need to solve 6 elements of the upper (or lower) triangular matrix in the $3{\times}3$ RP matrix $\Delta\mathbb{X}$. Thus, $\Delta\mathbb{X}$ can be expressed as:

$$\Delta\mathbb{X}{=}\begin{bmatrix}\Delta X_{aa'}^{hk} & \Delta X_{ab'}^{hk} & \Delta X_{ac'}^{hk}\\ \Delta X_{ba'}^{hk} & \Delta X_{bb'}^{hk} & \Delta X_{bc'}^{hk}\\ \Delta X_{ca'}^{hk} & \Delta X_{cb'}^{hk} & \Delta X_{cc'}^{hk}\end{bmatrix}{=}\begin{bmatrix}\beta_1{\cdot}X_{aa'}^{hk} & \beta_2{\cdot}X_{ab'}^{hk} & \beta_3{\cdot}X_{ac'}^{hk}\\ \beta_2{\cdot}X_{ba'}^{hk} & \beta_4{\cdot}X_{bb'}^{hk} & \beta_5{\cdot}X_{bc'}^{hk}\\ \beta_3{\cdot}X_{ca'}^{hk} & \beta_5{\cdot}X_{cb'}^{hk} & \beta_6{\cdot}X_{cc'}^{hk}\end{bmatrix}\tag{20}$$

The 6 elements of the RP matrix $\Delta\mathbb{X}$ estimated are $\Delta X_{aa'}^{hk}$, $\Delta X_{ab'}^{hk}(\Delta X_{ba'}^{hk})$, $\Delta X_{ac'}^{hk}(\Delta X_{ca'}^{hk})$, $\Delta X_{bb'}^{hk}$, $\Delta X_{bc'}^{hk}(\Delta X_{cb'}^{hk})$, and $\Delta X_{cc'}^{hk}$. Equivalently, we need to solve $\{\beta_1,\cdots,\beta_6\}$. Let the number of system nodes be N, the number of D-FACTS devices be N_{D}, and the voltage magnitude of the node connected with the main grid be fixed as 1.0 p.u. Thus, the total number of estimated variables should be $(2N{-}1)\times 6+N_{\mathrm{D}}\times 6$. The variables in (17) can be completely expressed as:

$$\begin{aligned}\mathbf{X}{=}\Big[&\dot{\boldsymbol{V}}_{a}^{1},\dot{\boldsymbol{V}}_{b}^{1},\dot{\boldsymbol{V}}_{c}^{1},\dot{\boldsymbol{V}}_{a}^{2},\dot{\boldsymbol{V}}_{b}^{2},\dot{\boldsymbol{V}}_{c}^{2},\cdots,\dot{\boldsymbol{V}}_{a}^{N},\dot{\boldsymbol{V}}_{b}^{N},\dot{\boldsymbol{V}}_{c}^{N};\\ &\dot{\boldsymbol{U}}_{a}^{1},\dot{\boldsymbol{U}}_{b}^{1},\dot{\boldsymbol{U}}_{c}^{1},\dot{\boldsymbol{U}}_{a}^{2},\dot{\boldsymbol{U}}_{b}^{2},\dot{\boldsymbol{U}}_{c}^{2},\cdots,\dot{\boldsymbol{U}}_{a}^{N},\dot{\boldsymbol{U}}_{b}^{N},\dot{\boldsymbol{U}}_{c}^{N};\\ &\Delta X_{aa'}^{hk},\Delta X_{ab'}^{hk},\Delta X_{ac'}^{hk},\Delta X_{bb'}^{hk},\Delta X_{bc'}^{hk},\Delta X_{cc'}^{hk}\Big]\end{aligned}\tag{21}$$

For the sake of comparison, two types of shallowly-hidden MTD strategies are additionally designed as: (i) only Shallowly-Hiding self reactance (SH1-MTD) and (ii) only Shallowly-Hiding mutual reactance (SH2-MTD). To vividly show the distinction among AH-, SH1-, SH2-, and DH-MTD, Table II illustrates the differentiated RP matrix $\Delta\mathbb{X}$ (see the second row) and the total number of estimated variables (see the third row). Basically, AH-MTD can arbitrarily hide N_{D} coefficients; SH1- and SH2-MTD can shallowly hide $3N_{\mathrm{D}}$ coefficients; and the proposed DH-MTD can deeply hide $6N_{\mathrm{D}}$ coefficients.

C. DH-MTD Model Solving: Trust-Region Method

To solve the NLS problem in (19), we use the trust-region method, which is a simple yet powerful concept in optimization. A brief standard pseudocode is given in Algorithm 1, where N_{T} is the total number of iterations.

Definition 1: *The nonlinear objective function* $\mathbf{f}\left(\mathbf{X}\right)$ *in* (19) *is a second-order continuous differentiable function defined in the real coordinate space* $\mathbb{R}^n$ *of* n *dimensions (or variables). The neighborhood of the current point* $\mathbf{X}_i$ *is defined as* $\boldsymbol{\Omega}_i$*:* $\boldsymbol{\Omega}_i{=}\{\mathbf{X}{\in}\mathbb{R}^n|\|\mathbf{X}{-}\mathbf{X}_i\|{\le}\boldsymbol{\zeta}_i\}$*, where* $\boldsymbol{\zeta}_i$ *is the trust region radius.*

TABLE II
DIFFERENCE OF RP MATRIX $\Delta\mathbb{X}$ AMONG AH-MTD, SH1-MTD, SH2-MTD, AND DH-MTD

AH-MTD	SH1-MTD	SH2-MTD	DH-MTD
$\Delta\mathbb{X}=\beta_{\mathrm{H}}\cdot\begin{bmatrix}X_{aa'}^{hk} & X_{ab'}^{hk} & X_{ac'}^{hk}\\ X_{ba'}^{hk} & X_{bb'}^{hk} & X_{bc'}^{hk}\\ X_{ca'}^{hk} & X_{cb'}^{hk} & X_{cc'}^{hk}\end{bmatrix}$	$\Delta\mathbb{X}=\begin{bmatrix}\beta_1\cdot X_{aa'}^{hk} & 0 & 0\\ 0 & \beta_2\cdot X_{bb'}^{hk} & 0\\ 0 & 0 & \beta_3\cdot X_{cc'}^{hk}\end{bmatrix}$	$\Delta\mathbb{X}=\begin{bmatrix}0 & \beta_1\cdot X_{ab'}^{hk} & \beta_2\cdot X_{ac'}^{hk}\\ \beta_1\cdot X_{ba'}^{hk} & 0 & \beta_3\cdot X_{bc'}^{hk}\\ \beta_2\cdot X_{ca'}^{hk} & \beta_3\cdot X_{cb'}^{hk} & 0\end{bmatrix}$	$\Delta\mathbb{X}=\begin{bmatrix}\beta_1\cdot X_{aa'}^{hk} & \beta_2\cdot X_{ab'}^{hk} & \beta_3\cdot X_{ac'}^{hk}\\ \beta_2\cdot X_{ba'}^{hk} & \beta_4\cdot X_{bb'}^{hk} & \beta_5\cdot X_{bc'}^{hk}\\ \beta_3\cdot X_{ca'}^{hk} & \beta_5\cdot X_{cb'}^{hk} & \beta_6\cdot X_{cc'}^{hk}\end{bmatrix}$
$(2N{-}1)\times 6+N_{\mathrm{D}}\times 1$	$(2N{-}1)\times 6+N_{\mathrm{D}}\times 3$	$(2N{-}1)\times 6+N_{\mathrm{D}}\times 3$	$(2N{-}1)\times 6+N_{\mathrm{D}}\times 6$

Algorithm 1: Trust-Region Method-based NLS for Solving DH-MTD Model

```
Input: Initial point $\mathbf{X}_0$, upper bound $\bar{\boldsymbol{\zeta}}$ of $\boldsymbol{\zeta}$:$\boldsymbol{\zeta}_0\in(0,\bar{\boldsymbol{\zeta}})$,
  thresholds $\phi$, $0<\eta_1\leq\eta_2<1$, and $0<\gamma_1<1<\gamma_2$;
for Iteration $i$ from $0$ to $N_T$ do
    Solve trust-region problem in (22) and obtain $\boldsymbol{\Theta}_i$;
    if Impedance matrix $\dot{\boldsymbol{Z}}+j\Delta\mathbb{X}$ is invertible then
        Calculate $\mathbf{f}(\mathbf{X}_i)$, $\mathbf{r}_i$, and $\mathbf{X}_{i+1}$:
            $\begin{cases}\mathbf{f}(\mathbf{X}_i)=[\mathbf{Y}-\mathbf{G}(\mathbf{X}_i)]^{\mathrm{T}}\cdot[\mathbf{Y}-\mathbf{G}(\mathbf{X}_i)]\\ \mathbf{r}_i=\dfrac{[\mathbf{f}(\mathbf{X}_i)-\mathbf{f}(\mathbf{X}_i+\boldsymbol{\Theta}_i)]}{[\mathbf{q}_i(0)-\mathbf{q}_i(\boldsymbol{\Theta}_i)]}\\ \mathbf{X}_{i+1}=\begin{cases}\mathbf{X}_i+\boldsymbol{\Theta}_i, & \text{if } \mathbf{r}_i\geq\eta_1\\ \mathbf{X}_i, & \text{else}\end{cases}\end{cases}$;
        Calibrate the trust-region radius $\boldsymbol{\zeta}_{i+1}$:
            $\begin{cases}\boldsymbol{\zeta}_{i+1}\in(0,\gamma_1\boldsymbol{\zeta}_i], \text{ if } \mathbf{r}_i<\eta_1\\ \boldsymbol{\zeta}_{i+1}\in[\gamma_1\boldsymbol{\zeta}_i,\boldsymbol{\zeta}_i], \text{ if } \mathbf{r}_i\in[\eta_1,\eta_2)\\ \boldsymbol{\zeta}_{i+1}\in\left[\boldsymbol{\zeta}_i,\min\left\{\gamma_2\boldsymbol{\zeta}_i,\bar{\boldsymbol{\zeta}}\right\}\right], \text{ if } \mathbf{r}_i\geq\eta_2\end{cases}$;
        Generate gradient $\mathbf{g}_{i+1}$ and Hessian matrix
          $\mathbf{H}_{i+1}$;
    else
        As impedance matrix $\dot{\boldsymbol{Z}}+j\Delta\mathbb{X}$ is irreversible,
          variable $\mathbf{X}_{i+1}$ is directly updated by:
                $\mathbf{X}_{i+1}=\boldsymbol{\Theta}_i-\mathbf{X}_i$;
    end
    if $\max\left[|\mathbf{X}_i-\mathbf{X}_{i+1}|,|\mathbf{f}(\mathbf{X}_i)-\mathbf{f}(\mathbf{X}_{i+1})|\right]<\phi$ then
        Terminate the iteration;
    end
end
```

Mathematically, the trust-region problem is constituted as:

$$\begin{cases}\min & \mathbf{q}_i(\boldsymbol{\Theta})=\mathbf{f}(\boldsymbol{\Theta})+\mathbf{g}_i^{\mathrm{T}}\boldsymbol{\Theta}+\frac{1}{2}\boldsymbol{\Theta}^{\mathrm{T}}\mathbf{H}_i\boldsymbol{\Theta}\\ s.t. & \|\boldsymbol{\Theta}\|\leq\boldsymbol{\zeta}_i\end{cases}\tag{22}$$

where $\boldsymbol{\Theta}=\mathbf{X}{-}\mathbf{X}_i$. $\mathbf{g}_i$ is the gradient of objective function $\mathbf{f}(\cdot)$ at the current point $\mathbf{X}_i$ and $\mathbf{g}_i=\nabla\mathbf{f}(\mathbf{X}_i)$. $\mathbf{H}_i$ is the Hessian matrix at the current point $\mathbf{X}_i$. $\|\cdot\|$ denotes the 2-norm.

Four steps are briefly introduced in Algorithm 1 as follows. Step i): the initial point $\mathbf{X}_0$, upper bound $\bar{\boldsymbol{\zeta}}$ of the trust region radius $\boldsymbol{\zeta}$, thresholds ϕ, η_1, η_2, γ_1, and γ_2 are prepared; Step ii): in each iteration, the trust-region problem modeled in (22) is solved to obtain the solution $\boldsymbol{\Theta}_i$; Step iii): if the impedance matrix is invertible, the objective function $\mathbf{f}(\mathbf{X}_i)$ and estimated variable $\mathbf{X}_{i+1}$ are calculated while the trust-region radius $\boldsymbol{\zeta}_{i+1}$ is calibrated; Step iv): if the threshold is satisfied, the iteration is terminated.

III. DATA-DRIVEN MTD ALLOCATION

Though we have designed an elaborate DH-MTD model, it is still challenging to choose the exact transmission line to install D-FACTS for the reactance perturbation. This concept is the so-called MTDA problem in this paper. To solve this problem, we propose a novel data-driven MTDA method. First, it is assumed that operators have already learned and known the prior information of the DSSE under historical normal operating conditions. This process can be readily implemented by state-of-the-art DSSE techniques. Then, let $\varepsilon_N=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$ denote the set of estimation errors and $f_N(\varepsilon_N)$ denote the corresponding PDF under the historical normal operating condition, while $\tilde{f}(\tilde{\varepsilon})$ denotes the PDF of estimation errors in cyberattack scenarios. The proposed NPPR index can be defined as:

$$NPPR=\left[f_N(\varepsilon_N=0)-\tilde{f}(\tilde{\varepsilon}=0)\right]/\tilde{f}(\tilde{\varepsilon}=0)\tag{23}$$

Remark 1: *As the ideal value of estimation errors is expected to be zero ($\varepsilon\to 0$) and their probability distribution is unimodal, the PDF peak (PP) should approximate the value that zero corresponds to, i.e, $PP=\lim\limits_{\varepsilon\to 0}f(\varepsilon)\approx f(\varepsilon=0)$.*

To control the smoothness of the resulting density curve of estimation errors, a nonparametric kernel smoothing estimator is used to represent the PDF, i.e., $f(x)=\sum_{k=1}^{n}K\left(\frac{x-\varepsilon_k}{h}\right)/nh$. Considering the definition of NPPR in (23), we can get:

$$NPPR=\frac{\frac{1}{n_N h_N}\sum_{k=1}^{n_N}K\left(\frac{\varepsilon_{N,k}}{h_N}\right)-\frac{1}{\tilde{n}\tilde{h}}\sum_{k=1}^{\tilde{n}}K\left(\frac{\tilde{\varepsilon}_k}{\tilde{h}}\right)}{\frac{1}{\tilde{n}\tilde{h}}\sum_{k=1}^{\tilde{n}}K\left(\frac{\tilde{\varepsilon}_k}{\tilde{h}}\right)}\tag{24}$$

where $K(\cdot)$ is the kernel smoothing function. n is the sample size. h is the bandwidth. Marks of '$\sim$' and 'N' denote the cyberattack scenario and historical normal operating condition.

Based on the NPPR index, we propose a data-driven technique to allocate the exact transmission line for MTD. The optimal branch i_{opt} with the smallest NPPR value, i.e., $\arg\min(NPPR_i)_{i\in\mathcal{M}}$, where $\mathcal{M}$ is the set of branch measurements, is chosen as the final MTDA. This is because the smaller NPPR is, the less impacts to the distribution system cyberattacks may cause. As the NPPR metric has already been learned and known by operators, the minimum NPPR can be used to dynamically select measurements for DSSE. The pseudocode of the proposed data-driven MTDA is given in Algorithm 2. Four steps are briefly introduced as follows. Step i): the dataset of historical state estimation errors is prepared under normal operating conditions; Step ii): the data-driven NPPR metric is calculated in the sth Monte Carlo simulation; Step iii): the measurements with the minimum $NPPR_{i,s}$ for scenarios are selected as a data-driven MTDA and checked

Algorithm 2: Proposed Data-Driven MTDA

```
Input: Historical state estimation errors $\varepsilon_N$ and actual
  measurements in cyberattack scenarios.
for Monte Carlo scenario $s$ from 1 to $N_s$ do
    Calculate NPPR for all branch measurements $\mathcal{M}$.
    Find the optimal branch with minimum NPPR:
        $$i_{\text{opt},s} = \arg \min_{i \in \mathcal{M}} NPPR_{i,s}$$   (25)
    if topology is observable by spanning trees then
        Hold $i_{\text{opt},s}$ and $\mathcal{M}$;
    else
        Remove $i_{\text{opt},s}$: $\mathcal{M}'|_{i_{\text{opt},s} \notin \mathcal{M}'} = \mathcal{M} - i_{\text{opt},s}$ and
          update:
        $$j_{\text{opt},s} = \arg \min_{j \in \mathcal{M}'} NPPR_{j,s}$$   (26)
        $$i_{\text{opt},s} \leftarrow j_{\text{opt},s};\ \mathcal{M} \leftarrow \mathcal{M}'$$   (27)
    end
    Implement DSSE and save estimated errors of
      measurements.
end
```

by a spanning tree for ensuring the topological observability; and Step iv): the measurements selected with MTDA are used for state estimation. The optimal set of measurements is not uniquely determined and can be dynamically updated.

The overall framework of the proposed DH-MTD strategy considering voltage stability is illustrated in Fig. 2. It mainly consists of four major steps: D-FACTS allocation, DH-MTD modeling, DH-MTD model solving, and performance evaluation. The overall procedure of the proposed DH-MTD method mainly consists of four steps, which are briefly introduced as follows.

(i) A data-driven NPPR index is designed to allocate the D-FACTS device.

(ii) A mathematical NLS optimization model is constituted for the DH-MTD strategy, including both the branch and injection power phasor measurement constraints in the normal operation and cyberattack scenarios.

(iii) A trust-region algorithm is utilized to solve the nonlinear DH-MTD model considering non-Gaussian ambient noises.

(iv) Numerical, physical, and visualized metrics are developed to evaluate the performance of multiple MTD strategies, i.e., AH-, SH1-, SH2-, and DH-MTD.

IV. Evaluation Metrics

Multiple metrics are used to evaluate the performance, that is, root mean square error (RMSE), standard deviation (STD), voltage unbalance factor (VUF), true positive rate (TPR), performance index (PI), and performance diagram. RMSE and STD have commonly been used as a measure of how far the estimates are from the real measurements. VUF is an European standard to indicate the degree of unbalance [18]. TPR has widely been used to indicate the detection accuracy. PI is used to analyze the contingency [19], [20]. The performance diagram is used to visualize the estimation skill.

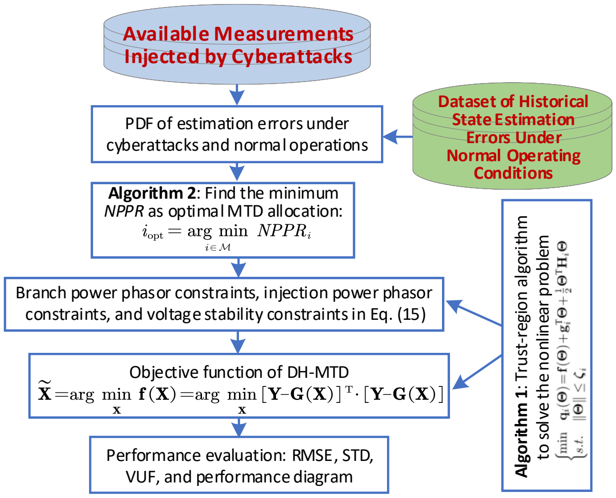

Available Measurements Injected by Cyberattacks
Dataset of Historical State Estimation Errors Under Normal Operating Conditions
PDF of estimation errors under cyberattacks and normal operations
Algorithm 2: Find the minimum NPPR as optimal MTD allocation: $i_{\text{opt}} = \arg \min_{i \in \mathcal{M}} NPPR_i$
Branch power phasor constraints, injection power phasor constraints, and voltage stability constraints in Eq. (15)
Objective function of DH-MTD $\widetilde{\mathbf{X}} = \arg \min_{\mathbf{X}} \mathbf{f}(\mathbf{X}) = \arg \min_{\mathbf{X}} [\mathbf{Y} - \mathbf{G}(\mathbf{X})]^{\mathrm{T}} \cdot [\mathbf{Y} - \mathbf{G}(\mathbf{X})]$
Performance evaluation: RMSE, STD, VUF, and performance diagram
Algorithm 1: Trust-region algorithm to solve the nonlinear problem $\begin{cases} \min & \mathbf{q}_i(\boldsymbol{\Theta}) = \mathbf{f}(\boldsymbol{\Theta}) + \mathbf{g}_i^{\mathrm{T}} \boldsymbol{\Theta} + \frac{1}{2} \boldsymbol{\Theta}^{\mathrm{T}} \mathbf{H}_i \boldsymbol{\Theta} \\ s.t. & \|\boldsymbol{\Theta}\| \le \zeta_i \end{cases}$

Fig. 2. Framework of the proposed DH-MTD strategy considering voltage stability.

A. Metrics I: Numerical and Physical Evaluation

To numerically evaluate the performance of different MTD strategies for cyberattacks, RMSE and STD are used for comparison:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\widetilde{X}_i - X_i \right)^2} \times 100\% \qquad (28)$$

$$STD = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(e_i - \bar{e} \right)^2} \qquad (29)$$

where $\widetilde{X}_i \in \widetilde{\mathbf{X}}$ is the estimated state value and solved by (19). $X_i \in \mathbf{X}$ is the actual state value. $e_i = \widetilde{X}_i - X_i$ is the estimation error. $\bar{e}$ is its corresponding mean value.

As a physical metric, the VUF metric defined as the percentage ratio of the negative sequence voltage ($\dot{\boldsymbol{U}}_{\text{negative}}$) over the positive sequence voltage ($\dot{\boldsymbol{U}}_{\text{positive}}$) has been used in the literature [21]:

$$VUF_i = 100 \times \dot{\boldsymbol{U}}^i_{\text{negative}} / \dot{\boldsymbol{U}}^i_{\text{positive}} \qquad (30)$$

where i denotes the ith system node. To physically validate the effectiveness of the proposed DH-MTD, we use the VUF deviation (VUFD) to design a metric of the average absolute VUFD ($\overline{VUFD}$) for evaluation:

$$\overline{VUFD} = \frac{1}{N} \sum_{i}^{N} |VUFD_i| = \frac{1}{N} \sum_{i}^{N} \left| \widetilde{VUF_i} - VUF_i \right| \qquad (31)$$

where $\widetilde{VUF_i}$ is the estimated VUF at node i using one MTD strategy. VUF_i is the actual VUF at node i. Ideally, $\overline{VUFD}$ should be as close to zero as possible. The smaller $\overline{VUFD}$ is, the better performance one MTD strategy can provide.

To demonstrate the effectiveness regarding the detection accuracy, TPR is defined as the percentage of the number of detected true positive cyberattacks (TP) that are actually observed in the real system measurements over the total number of cyberattacks (N_A), given by:

$$TPR = TP / N_{\text{A}} \times 100\% \qquad (32)$$

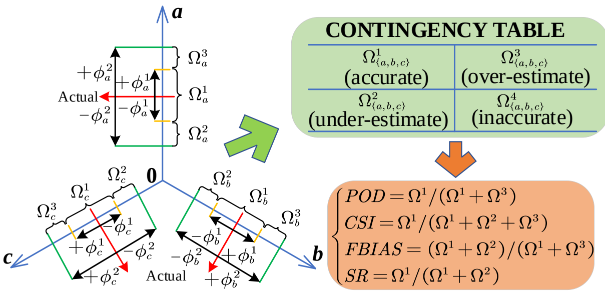

a
b
c
0
Actual
Actual
$+\phi_a^2$
$+\phi_a^1$
$-\phi_a^1$
$-\phi_a^2$
Ω_a^3
Ω_a^1
Ω_a^2
Ω_b^1
Ω_b^2
Ω_b^3
Ω_c^1
Ω_c^2
Ω_c^3
$-\phi_b^1$
$+\phi_b^1$
$-\phi_b^2$
$+\phi_b^2$
$-\phi_c^1$
$+\phi_c^1$
$-\phi_c^2$
$+\phi_c^2$
CONTINGENCY TABLE
$\Omega^1_{\{a,b,c\}}$ (accurate)
$\Omega^3_{\{a,b,c\}}$ (over-estimate)
$\Omega^2_{\{a,b,c\}}$ (under-estimate)
$\Omega^4_{\{a,b,c\}}$ (inaccurate)
$POD = \Omega^1/(\Omega^1+\Omega^3)$
$CSI = \Omega^1/(\Omega^1+\Omega^2+\Omega^3)$
$FBIAS = (\Omega^1+\Omega^2)/(\Omega^1+\Omega^3)$
$SR = \Omega^1/(\Omega^1+\Omega^2)$

Fig. 3. Estimation metrics with tolerances $\{\phi_a^1, \phi_b^1, \phi_c^1\}$ and $\{\phi_a^2, \phi_b^2, \phi_c^2\}$.

To perform the contingency analysis of multiple MTD strategies, the performance index PI is formulated by:

$$PI = \sum_{h\in\mathcal{N}} \frac{W_h}{2n} \left(\frac{\left|\dot{\boldsymbol{V}}_h\right| - \left|\dot{\boldsymbol{V}}_h^{\text{sp}}\right|}{\Delta\dot{\boldsymbol{V}}_h^{\text{Lim}}} \right)^{2n} \tag{33}$$

where $\left|\dot{\boldsymbol{V}}_h\right|$ is the voltage magnitude at bus h. $\left|\dot{\boldsymbol{V}}_h^{\text{sp}}\right|$ is the specified (rated) voltage magnitude at bus h. $\Delta\dot{\boldsymbol{V}}_h^{\text{Lim}}$ is the voltage deviation limit. n is the exponent of penalty function (n=1 preferred). $\mathcal{N}$ is the set of system nodes. W_h is the real non-negative weighting factor (W_h=1 preferred). A smaller PI value indicates a better contingency performance of one MTD strategy.

B. Metrics II: Visualized Estimation Performance

Fig. 3 shows a measure of estimation skill based on a contingency table for comparison. Given sets of Ω_1, Ω_2, and Ω_3 are used in this figure, where the accurate set is $\Omega^1_{\{a,b,c\}}$: $\widetilde{\mathbf{X}} \in \mathbf{X} \pm \phi^1_{\{a,b,c\}}$; the under-estimated set is $\Omega^2_{\{a,b,c\}}$: $(\widetilde{\mathbf{X}} \in \mathbf{X} - \phi^2_{\{a,b,c\}}) - (\widetilde{\mathbf{X}} \in \mathbf{X} - \phi^1_{\{a,b,c\}})$; and the over-estimated set is $\Omega^3_{\{a,b,c\}}$: $(\widetilde{\mathbf{X}} \in \mathbf{X} + \phi^2_{\{a,b,c\}}) - (\widetilde{\mathbf{X}} \in \mathbf{X} + \phi^1_{\{a,b,c\}})$. The performance diagram is visualized by metrics including the probability of detection (POD), critical success index (CSI), frequency bias score (FBIAS), and success ratio (SR). Detailed information about the performance diagram and metrics can be found in [22], [23].

V. Case Studies and Results

The unbalanced IEEE 123-node standard distribution network [21] is used to carry out case studies in this section. This network is a radial distribution system. The presence of single- and two-phase loads causes significant load imbalance. To perform the data-driven D-FACTS allocation, we set 1,000 Monte Carlo simulations. Numerical simulations are performed in the MATLAB R2017a environment [24], [25]. Four strategies are compared to validate the effectiveness of the proposed MTD method, i.e., AH-, SH1-, SH2-, and DH-MTD. The measurement noises are assumed to obey the non-Gaussian distribution[2]. Non-Gaussian noises are generated using the moment-based Hermite transformation model [26]. The NLS solver lsqnonlin in MATLAB [27] is used to solve the aforementioned MTD models. The threshold ϕ of the stopping tolerance is set as 10^{-6}. The maximum number of iterations is set as 400. To validate the effectiveness of the proposed DH-MTD strategy, five basic assumptions are made in this section:

TABLE III
RP Coefficient Estimation Results

MTD	RP Coeff.	Actual	Estimation
AH	β_{H}	0.18	0.1808
SH1	$\{\beta_1, \beta_2, \beta_3\}$	{0.18, 0.16, 0.12}	{0.1798, 0.1598, 0.1196}
SH2	$\{\beta_1, \beta_2, \beta_3\}$	{0.18, 0.16, 0.12}	{0.1805, 0.1613, 0.1211}
DH	$\{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6\}$	{0.18, 0.16, 0.12, 0.15, 0.13, 0.17}	{0.1801, 0.1609, 0.1193, 0.1495, 0.1296; 0.1688}

(i) The information of the historical normal operating condition (i.e., DSSE) has been entirely known by operators in advance, while the information of the exact location and type of the cyberattack is unclear in the current cyberattack scenario.

(ii) Voltage magnitudes are estimated with significantly less fluctuations in the cyberattack scenario. This assumption aims to guarantee the voltage stability [6].

(iii) The proposed DH-MTD-based network reconfiguration is performed in the unbalanced distribution network, which is still an ongoing extension as mentioned in [12].

(iv) The measurement noises are assumed to obey the non-Gaussian distribution, which has seldom been studied in the existing literature.

(v) Based on the available prior knowledge, pseudo measurements are predefined at the buses where no measurements devices are present. The three-phase active and reactive power injection is inferred based on historical data or statistical assumptions.

In the following case studies, first, we seek to validate the effectiveness of the proposed data-driven allocation method for MTD. Second, we seek to validate the effectiveness of the proposed DH-MTD. Then, we try to use it for the analysis of system voltage unbalancing status and cyberattack detection.

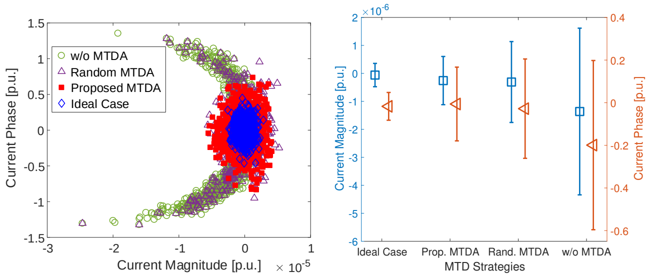
A. MTDA Validation and Analysis Using Monte Carlo

Three MTDA methods are compared to validate the effectiveness of the proposed MTDA, i.e., random MTDA, w/o (without) MTDA, and proposed MTDA. For the random MTDA, measurements are randomly chosen as inputs for the MTD strategy. For the w/o MTDA, no MTD strategies are adopted in cyberattack scenarios. For the ideal case, the measurement data is securely collected without considering any intrusion of cyberattacks.

To validate the effectiveness of the proposed MTDA method, the current magnitude measurement is manipulated by attackers using a multiplicative cyberattack with a coefficient (×10). Multiple Monte Carlo simulations are performed with this coefficient by adding ambient noises. This cyberattack is set at Branch 9-10 (NO. 10), which is taken as an example for a better illustration.

Fig. 4 shows the estimation results of measurements using Monte Carlo simulations with three strategies. Among all the three allocation methods (namely w/o MTDA, random MTDA, and proposed MTDA), the proposed MTDA method shows the

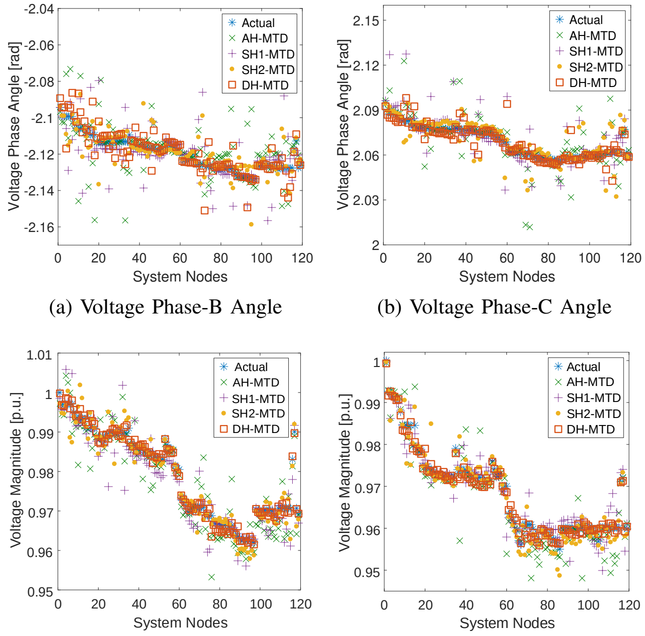
[2] As Gaussian distribution-based noises have already been processed by state-of-the-art literature [6], [7], [10], [14]–[16], we do not seek to repeatedly cope with Gaussian noises in this paper.



(a) Scatter diagram of current magnitude and phase

(b) Error bars of current magnitude and phase

Fig. 4. Estimation results of measurements using Monte Carlo simulations with three MTDA methods.



(a) Voltage Phase-B Angle

(b) Voltage Phase-C Angle

(c) Voltage Phase-B Magnitude

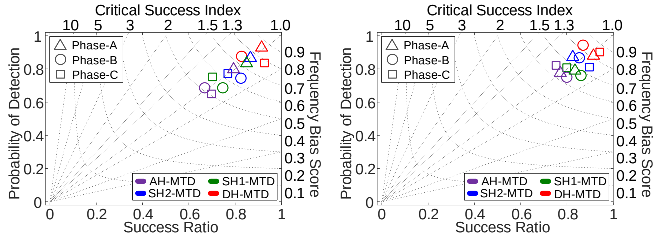
(d) Voltage Phase-C Magnitude

Fig. 5. Estimation results of state variables using four MTD strategies.

sharpest peaks and lightest tails for both current phase and magnitude measurements (see the solid red lines). It shows that the proposed MTDA outperforms the random MTDA and w/o MTDA. Fig. 4a vividly illustrates the scatter diagram of current magnitude and phase using 1,000 Monte Carlo simulations. As shown in this figure, the proposed MTDA can obtain relatively concentrated estimation errors around zero (see the red rectangles). Fig. 4b shows estimation error bars of current magnitude and phase measurements. As can be seen, the proposed MTDA can provide narrower confidence intervals and smaller mean values of estimation errors, compared with the strategy without MTDA and random MTDA methods.

B. Effectiveness Analysis of Proposed DH-MTD

The estimated RP coefficients are illustrated in Table III using AH-, SH1-, SH2-, and DH-MTD strategies. As can be seen, the estimated RP coefficients are very close to the actual values. To further validate the effectiveness of the proposed DH-MTD, Fig. 5 illustrates the estimation results of voltage phase angle and magnitude at each node using four MTD strategies: AH-, SH1-, SH2-, and DH-MTD. As can be seen, the estimation results using our proposed DH-MTD (see the rectangles) are much concentrated to the actual values (see the asterisks), compared with AH-MTD (see the crosses), SH1-MTD (see the pluses), and SH2-MTD (see the yellow dots). Also, it should be noted that some samples are almost under 0.95 p.u. using AH-, SH1-, and SH2-MTD at the bottom of Fig 5d. It means that voltage profiles may violate the normal operating limit of 0.95∼1.05 p.u. This is because those MTD strategies cannot elaborately hide the RP matrix to defend against cyberattacks, which consequently causes the compromised estimation results of voltage magnitude.



(a) Voltage Phase Angle

(b) Voltage Magnitude

Fig. 6. Performance diagram of contingency table for MTD strategies.

TABLE IV
COMPARISON OF RMSE METRICS FOR DIFFERENT MTD STRATEGIES [‰]

Variables	Phase	AH-MTD	SH1-MTD	SH2-MTD	DH-MTD
Angle	A	9.07	6.54	8.68	***5.36***
	B	11.62	10.73	9.67	***7.09***
	C	11.61	10.86	9.73	***5.24***
Magnitude	A	3.95	3.19	2.97	***1.68***
	B	4.35	3.56	3.17	***2.21***
	C	4.07	3.11	2.94	***1.89***

Fig. 6 shows the visualized performance diagram for comparison of AH-, SH1-, SH2-, and DH-MTD strategies, respectively. For a performance diagram shown in Fig. 6, 1) the left axis represents the value of POD; 2) the bottom axis represents SR; 3) the diagonal dashed lines represent FBIAS; and 4) the dashed curves represent CSI. For a better performance, the points should be close toward the top right corner of the performance diagram. Three sets of thresholds $\{\phi_a^1, \phi_b^1, \phi_c^1\}$ and $\{\phi_a^2, \phi_b^2, \phi_c^2\}$ are predefined based on different variable values at each phase. Fig. 6 demonstrates the good state estimation performance of the proposed DH-MTD model compared with the other MTD models. The red marks (estimated by DH-MTD) are the closest to the top right corner than the blue marks (estimated by SH2-MTD), green marks (estimated by SH1-MTD), and purple marks (estimated by AH-MTD). Specifically, for the voltage phase angle, the mean SR value (bottom x-axis) is 0.87 by using DH-MTD. The mean SR values using AH-, SH1-, and SH2-MTD are 0.74, 0.79, and 0.82, respectively. For the voltage magnitude, the mean SR value is 0.93 by using DH-MTD. The mean SR values using AH-, SH1-, and SH2-MTD are 0.79, 0.83, and 0.86, respectively.

To quantitatively evaluate the performance of different MTD strategies, Tables IV and V compare the RMSE and standard deviation metrics for AH-, SH1-, SH2-, and DH-MTD strategies, respectively. As can be seen, the proposed DH-MTD can

TABLE V
COMPARISON OF STANDARD DEVIATION STD METRICS FOR DIFFERENT MTD STRATEGIES [‰]

Variables	Phase	AH-MTD	SH1-MTD	SH2-MTD	DH-MTD
Angle	A	6.86	6.84	4.82	***3.63***
	B	14.02	10.25	9.01	***6.07***
	C	14.23	10.01	6.71	***5.19***
Magnitude	A	3.34	3.19	2.34	***1.59***
	B	4.45	3.38	2.68	***2.64***
	C	4.01	3.04	3.01	***1.69***

TABLE VI
COMPARISON OF GAUSSIAN AND NON-GAUSSIAN NOISE ASSUMPTION USING PROPOSED DH-MTD MODEL [‰]

Variables	Phase	RMSE		STD	
		Gaussian	non-Gaussian	Gaussian	non-Gaussian
Angle	Phase A	6.57	5.36	6.88	3.63
	Phase B	8.65	7.09	8.52	6.07
	Phase C	5.98	5.24	6.57	5.19
Mag.	Phase A	3.54	1.68	3.54	1.59
	Phase B	4.65	2.21	5.62	2.64
	Phase C	3.21	1.89	3.22	1.69

estimate the smallest RMSE and standard deviation values compared with the other strategies. This is mainly because the proposed DH-MTD can elaborately hide both the self and mutual reactance in the RP matrix on the three-phase branch. For example, with respect to the phase-A angle, RMSE using DH-MTD is reduced by 69.22%, 22.01%, and 61.94% in terms of AH-, SH1-, and SH2-MTD. Also, the standard deviation using DH-MTD is reduced by 88.98%, 88.43%, and 32.78%.

Another interesting finding is that the RMSE and standard deviation metrics of voltage magnitude are much smaller than those of voltage phase angle. For instance, the mean RMSE of phase angle using DH-MTD is 5.89‰, while the mean RMSE of magnitude is 1.92‰. The mean standard deviation of phase angle using DH-MTD is 4.96‰, while the mean standard deviation of magnitude is 1.97‰. This is because all of the MTD strategies consider to maintain the voltage stability as formulated in (13) and (14).

To show the effectiveness of processing non-Gaussian noises, Gaussian distribution-based measurement noises are used for a comparison to the validate the improvement of the proposed DH-MTD model. Measurements are added by a Gaussian noise with zero mean and a standard deviation as one thousandth of the corresponding value. Table VI compares RMSE and STD metrics for the proposed DH-MTD model based on the Gaussian and non-Gaussian noise assumption. As can be seen, compared with the Gaussian noise assumption, both RMSE and STD metrics are slightly reduced by using the proposed DH-MTD model for the non-Gaussian noise assumption.

C. Voltage Balancing Status Analysis

The voltage balancing status information is critical to help the distribution systems operate in a good condition. Fig. 7 illustrates the estimated VUF values at each system node using AH-, SH1-, SH2-, and DH-MTD strategies. This figures shows all of the MTD strategies can estimate VUF values within the acceptable limit of 0∼3% as mentioned in [21]. However, using AH-, SH1-, and SH2-MTD may cause some diffused VUF values that are relatively far away from the corresponding actual VUF values. The proposed DH-MTD can estimate the most accurate values that are closest to the actual VUF.

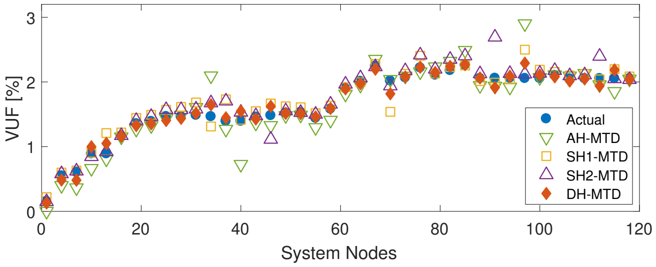


Fig. 7. VUF estimated values at each system node using MTD strategies.

TABLE VII
COMPARISON OF VUFD METRICS FOR DIFFERENT MTD STRATEGIES AT FIVE REPRESENTATIVE NODES [$\times 10^{-4}$]

Metrics	AH-MTD	SH1-MTD	SH2-MTD	DH-MTD
$VUFD_9$	5.72	5.18	2.34	***1.96***
$VUFD_{31}$	12.00	19.58	9.39	***2.65***
$VUFD_{56}$	12.80	10.29	7.83	***4.37***
$VUFD_{86}$	30.37	3.35	11.64	***1.93***
$VUFD_{117}$	13.41	20.21	9.05	***3.69***
$\overline{VUFD}$	13.36	10.63	10.01	***5.91***

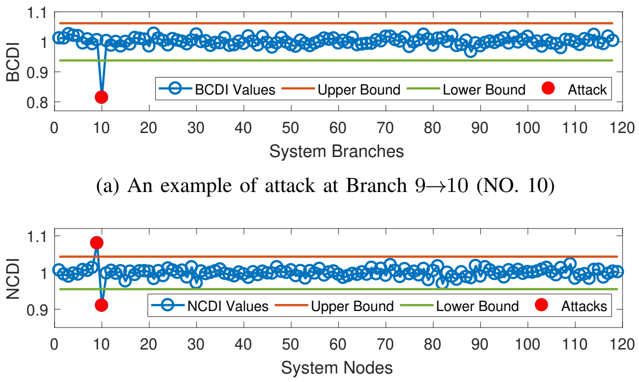
Table VII compares the VUFD metrics for different MTD strategies at five representative nodes (i.e., Nodes 9, 31, 56, 86, and 117). As can be seen, the proposed DH-MTD can provide the smallest VUFD values compared with other MTD strategies. For the $\overline{VUFD}$ metric, DH-MTD can reduce it by 126% [=(13.36-5.91)/5.91], 79.86% [=(10.63-5.91)/5.91], and 69.37% [=(10.01-5.91)/5.91], compared with AH-, SH1-, and SH2-MTD. This is because the proposed DH-MTD strategy can enhance the fitting accuracy of the voltage unbalance with respect to the actual voltage unbalance. However, using AH-, SH1-, and SH2-MTD cannot guarantee the system balancing status as the actual one. This observation validates that the actual balancing status of the unbalanced distribution systems is not affected too much by the proposed DH-MTD strategy in the cyberattack scenario.

D. Cyberattack Detection Analysis

Based on the estimated voltage phasor $\{\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c\}$ in the cyberattack scenario, the normal measurements can be readily estimated as $\mathbf{f}_{\mathrm{L}}(\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c)$ using (5) by distribution system operators[3]. This is because operators have the actual information of both the reactance and RP matrices. However, attackers know nothing about the RP matrix that has already been deeply hidden by the proposed DH-MTD. By comparing the estimated normal measurements with the power phasor measurements $\{\dot{\boldsymbol{S}}'_{\mathrm{L},a}, \dot{\boldsymbol{S}}'_{\mathrm{L},b}, \dot{\boldsymbol{S}}'_{\mathrm{L},c}\}$ in the cyberattack scenario, cyberattacks can be readily detected based on the three-sigma rule ($\mu \pm 3\sigma$). The branch cyberattack detection index (BCDI) is designed:

$$BCDI = \frac{\mathbf{f}_{\mathrm{L}}\left(\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c\right)}{\left[\dot{\boldsymbol{S}}'_{\mathrm{L},a}, \dot{\boldsymbol{S}}'_{\mathrm{L},b}, \dot{\boldsymbol{S}}'_{\mathrm{L},c}\right]^{\mathrm{T}}} \notin [\mu_{\mathrm{L}} - 3\sigma_{\mathrm{L}}, \mu_{\mathrm{L}} + 3\sigma_{\mathrm{L}}] \quad (34)$$

[3] Note that here we must use functions $\mathbf{f}_{\mathrm{L}}$ and $\mathbf{f}_{\mathrm{IN}}$ (instead of $\mathbf{h}_{\mathrm{L}}$ and $\mathbf{h}_{\mathrm{IN}}$) to estimate normal measurements that are not tampered with cyberattacks.



(a) An example of attack at Branch 9→10 (NO. 10)

(b) An example of attacks at Nodes 9 and 10

Fig. 8. Examples of branch and node cyberattacks using BCDI and NCDI.

TABLE VIII
COMPARISON OF MTD STRATEGIES FOR CYBERATTACKS

Methods	AH-MTD	SH1-MTD	SH2-MTD	DH-MTD
TPR [%]	94.8	95.9	96.5	99.2

Likewise, the node cyberattack detection index (NCDI) is designed as:

$$NCDI = \frac{\mathrm{f}_{\mathrm{IN}}\left(\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c\right)}{\left[\dot{\boldsymbol{S}}'_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'_{\mathrm{IN},b}, \dot{\boldsymbol{S}}'_{\mathrm{IN},c}\right]^{\mathrm{T}}} \notin [\mu_{\mathrm{IN}} - 3\sigma_{\mathrm{IN}}, \mu_{\mathrm{IN}} + 3\sigma_{\mathrm{IN}}] \tag{35}$$

Fig. 8 shows two examples of branch and node cyberattacks using BCDI and NCDI, where μ_{L}=0.9997; σ_{L}=0.0207; μ_{IN}=0.999; and σ_{IN}=0.0149. The upper and lower bounds are determined by the three-sigma rule, that is, $[\mu_{\mathrm{L}} - 3\sigma_{\mathrm{L}}, \mu_{\mathrm{L}} + 3\sigma_{\mathrm{L}}]$ and $[\mu_{\mathrm{IN}} - 3\sigma_{\mathrm{IN}}, \mu_{\mathrm{IN}} + 3\sigma_{\mathrm{IN}}]$. As can be seen, cyberattacks can be accurately detected by using BCDI and NCDI based on the proposed DH-MTD.

Table VIII compares different MTD strategies for detecting cyberattacks under 1,000 scenarios in the IEEE 123-node distribution system. As can be seen, the proposed DH-MTD shows the largest TPR metric (∼99%) compared with AH-MTD (∼94%), SH1-MTD (∼96%), and SH2-MTD (∼97%) strategies. This is because the estimation results $\mathrm{f}_{\mathrm{L}}(\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c)$ and $\mathrm{f}_{\mathrm{IN}}(\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c)$ using our proposed DH-MTD are much closer to the actual measurement values compared with AH-MTD, SH1-MTD, and SH2-MTD, which has been demonstrated in Fig. 5. Thus, the difference between estimates $\mathrm{f}_{\mathrm{L}}(\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c)$ & $\mathrm{f}_{\mathrm{IN}}(\dot{\boldsymbol{U}}_a, \dot{\boldsymbol{U}}_b, \dot{\boldsymbol{U}}_c)$ and attacked measurements $[\dot{\boldsymbol{S}}'_{\mathrm{L},a}, \dot{\boldsymbol{S}}'_{\mathrm{L},b}, \dot{\boldsymbol{S}}'_{\mathrm{L},c}]$ & $[\dot{\boldsymbol{S}}'_{\mathrm{IN},a}, \dot{\boldsymbol{S}}'_{\mathrm{IN},b}, \dot{\boldsymbol{S}}'_{\mathrm{IN},c}]$ can be identified more distinctly using the proposed DH-MTD.

E. Computational Efficiency Analysis

To demonstrate the computational efficiency of the proposed method, we compare it with a two-stage benchmark model. In the first stage, the benchmark model estimates the voltages $\dot{\boldsymbol{V}}$ for the previous normal operating condition. After obtaining the estimated $\dot{\boldsymbol{V}}$, the benchmark model estimates the voltages $\dot{\boldsymbol{U}}$ in the cyberattack scenario in the second stage. The comparison results of computational time for four cyberattacks on branches are illustrated in Table IX. As can be seen, the computational time using the proposed DH-MTD model is approximately in the range of 1.5∼1.9 seconds, while the computational time using the two-stage DH-MTD benchmark model is approximately in the range of 2.6∼3.1 seconds. Thus, the proposed model can significantly reduce the computational time and enhance the computational efficiency.

TABLE IX
COMPUTATIONAL TIME OF TWO DH-MTD MODELS FOR FOUR CYBERATTACKS ON BRANCHES [S]

Attacks	two-stage model	proposed model
Branch 9-10	2.6544	1.5364
Branch 26-29	3.0121	1.8352
Branch 58-59	2.8646	1.6645
Branch 99-100	3.0654	1.5323

TABLE X
PI INDEX COMPARISON FOR THE CONTINGENCY ANALYSIS OF MULTIPLE MTD STRATEGIES

Methods	AH-MTD	SH1-MTD	SH2-MTD	DH-MTD
PI	4.6525	1.6411	1.6724	0.2457

To analyze the contingency of multiple MTD strategies, Table X compares the PI index for the contingency analysis of multiple MTD strategies. As can be seen in this table, the proposed DH-MTD provides the smallest PI index (0.2457), while AH-, SH1-, and SH2-MTD strategies obtain much larger PI values. This observation is consistent with the findings in Fig. 5, where more samples are under 0.95 p.u. using AH-, SH1-, and SH2-MTD than using the proposed DH-MTD.

VI. DISCUSSION

To maintain the cybersecure and time-varying RP matrix, we design an architecture of the secure communication between the remote sensors (or telemetered devices) and DMS with cyberattack detection, which is shown in Fig. 9. In this architecture, the communication module is used to transmit the reactance and RP matrix information $\dot{\boldsymbol{Z}} + \Delta\mathbb{X}$ that is shared by the remote sensors (or telemetered devices). Due to the intrusion of cyberattacks on the communication module, the reactance and RP matrix information received by DMS is changed to $\dot{\boldsymbol{Z}}' + \Delta\mathbb{X}'$. We propose to design a cyberattack detection module that can compare the communication information sent by remote sensors $\dot{\boldsymbol{Z}} + \Delta\mathbb{X}$ and received by DMS $\dot{\boldsymbol{Z}}' + \Delta\mathbb{X}'$. A metric ϱ can be used to judge whether it exceeds the threshold ϵ. If $\varrho > \epsilon$, it means the communication module has been manipulated and remedial actions have to be developed by practitioners for remote sensor, communication module, and DMS. Under this circumstance, the reactance and RP matrix information $\dot{\boldsymbol{Z}} + \Delta\mathbb{X}$ has to dynamically vary from time to time. Besides, the reactance and RP matrix information $\dot{\boldsymbol{Z}}' + \Delta\mathbb{X}'$ received by DMS will not be included. Note that this paper does not seek to solve the cyberattack detection problem in the communication system, which has been widely studied in the intrusion detection system (IDS) for decades.

As seen at the bottom of Fig. 9, if the communication module is manipulated, the current reactance and RP matrix information $\dot{\boldsymbol{Z}}_t + \Delta\mathbb{X}_t$ would dynamically change to $\dot{\boldsymbol{Z}}_{t+1} + \Delta\mathbb{X}_{t+1}$ at the next time slot t+1. As cyberattacks have been detected and remedial actions have been taken at time slot t, the reactance and RP matrix $\dot{\boldsymbol{Z}}_{t+1} + \Delta\mathbb{X}_{t+1}$ at time

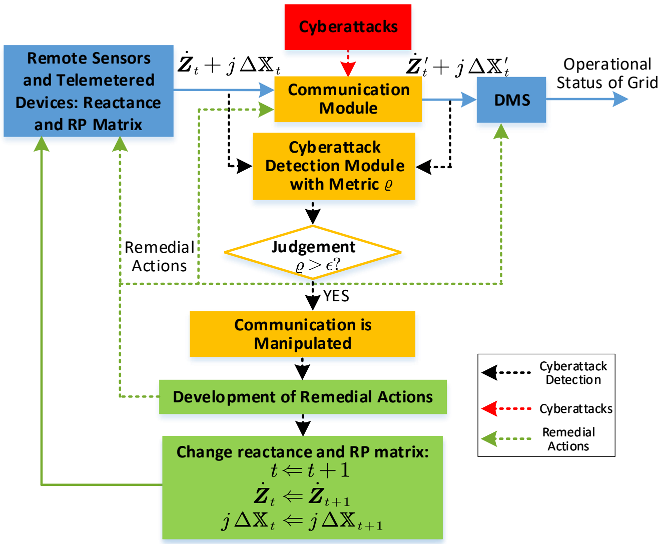

Cyberattacks
Remote Sensors and Telemetered Devices: Reactance and RP Matrix
$\dot{Z}_t + j\Delta\mathbb{X}_t$
$\dot{Z}'_t + j\Delta\mathbb{X}'_t$
Communication Module
DMS
Operational Status of Grid
Cyberattack Detection Module with Metric ϱ
Judgement $\varrho > \epsilon$?
Remedial Actions
YES
Communication is Manipulated
Development of Remedial Actions
Change reactance and RP matrix: $t \Leftarrow t+1$ $\dot{Z}_t \Leftarrow \dot{Z}_{t+1}$ $j\Delta\mathbb{X}_t \Leftarrow j\Delta\mathbb{X}_{t+1}$
Cyberattack Detection
Cyberattacks
Remedial Actions

Fig. 9. An architecture of the secure communication between the remote sensors (or telemetered devices) and DMS with cyberattack detection.

slot t+1 can be securely transmitted from remote sensors (or telemetered devices) to DMS.

Assume that an independent in-house testbed including network architectures and protocols can be developed for the cyberattack detection module (CDM), such as the Internet-Scale Event and Attack Generation Environment (ISEAGE) [28]. This in-house testbed has the highest security level, which means that its information of CDM can only be shared with distribution system operators who have the access authority. Moreover, a scalable Internet environment is independently provided to perform the detection and defense against cyberattacks that manipulate the sensor data packets of CDM.

VII. Conclusion

In this paper, we propose a novel moving-target-defense (MTD) strategy that can elaborately and actively change the self and mutual reactance of the transmission line in the unbalanced AC distribution system. The proposed deeply-hidden MTD (DH-MTD) model is constructed by the branch and injection power phasor measurement functions considering both the cyberattack scenario and normal operating condition. Also, the voltage stability can be ensured by solving the nonlinear least square (NLS) based DH-MTD model. Compared with conventional MTD methods, the proposed DH-MTD can estimate more accurate state variables and voltage balancing status. Also, it can be extended to the detection problem as cyberattacks occur. In future work, the measurement redundancy analysis will be performed based on the proposed MTD strategy, which may identify critical measurements and thereby benefit the MTD mechanism.

References

[1] X. Liu, L. Che, K. Gao, and Z. Li, “Power system intra-interval operational security under false data injection attacks,” *IEEE Trans. Ind. Inform.*, vol. 16, no. 8, pp. 4997–5008, Aug. 2019.

[2] A cyber-attack on an Indian nuclear plant raises worrying questions. [Online]. Available: https://www.economist.com/asia/2019/11/01/a-cyber-attack-on-an-indian-nuclear-plant-raises-worrying-questions

[3] M. Cui, J. Wang, and B. Chen, “Flexible machine learning-based cyberattack detection using spatiotemporal patterns for distribution systems,” *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1805–1808, Mar. 2020.

[4] S. Mousavian, J. Valenzuela, and J. Wang, “A probabilistic risk mitigation model for cyber-attacks to PMU networks,” *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 156–165, Jan. 2014.

[5] A. Ashok, M. Govindarasu, and J. Wang, “Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid,” *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1389–1407, 2017.

[6] J. Tian, R. Tan, X. Guan, and T. Liu, “Enhanced hidden moving target defense in smart grids,” *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2208–2223, March 2019.

[7] C. Liu, J. Wu, C. Long, and D. Kundur, “Reactance perturbation for detecting and identifying FDI attacks in power system state estimation,” *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 4, pp. 763–776, Aug. 2018.

[8] D. M. Divan, W. E. Brumsickle, R. S. Schneider, B. Kranz, R. W. Gascoigne, D. T. Bradshaw, M. R. Ingram, and I. S. Grant, “A distributed static series compensator system for realizing active power flow control on existing power lines,” *IEEE Trans. Power Deliv.*, vol. 22, no. 1, pp. 642–649, Jan. 2007.

[9] Z. Li, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, “Bilevel model for analyzing coordinated cyber-physical attacks on power systems,” *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2260–2272, 2016.

[10] Z. Zhang, R. Deng, D. K. Yau, P. Cheng, and J. Chen, “Analysis of moving target defense against false data injection attacks on power grid,” *IEEE Trans. Inf. Forensic Secur.*, vol. 15, pp. 2320–2335, 2020.

[11] B. Li, G. Xiao, R. Lu, R. Deng, and H. Bao, “On feasibility and limitations of detecting false data injection attacks on power grid state estimation using D-FACTS devices,” *IEEE Trans. Ind. Inform.*, vol. 16, no. 2, pp. 854–864, Feb. 2019.

[12] B. Liu, H. Wu, A. Pahwa, F. Ding, E. Ibrahim, and T. Liu, “Hidden moving target defense against false data injection in distribution network reconfiguration,” in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Portland, OR, USA, 2018, pp. 1–5.

[13] Y. Yao and Z. Li, “MTD-inspired state estimation based on random measurements selection,” in *2016 North American Power Symposium (NAPS)*. IEEE, 2016, pp. 1–6.

[14] S. Lakshminarayana, E. V. Belmega, and H. V. Poor, “Moving-target defense for detecting coordinated cyber-physical attacks in power grids,” in *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, Oct. 2019.

[15] Z. Zhang, R. Deng, D. Yau, P. Cheng, and J. Chen, “On effectiveness of detecting FDI attacks on power grid using moving target defense,” in *2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Feb 2019, pp. 1–5.

[16] S. Lakshminarayana and D. K. Y. Yau, “Cost-benefit analysis of moving-target defense in power grids,” in *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, June 2018, pp. 139–150.

[17] B. Chen, C. Chen, J. Wang, and K. L. Butler-Purry, “Sequential service restoration for unbalanced distribution systems and microgrids,” *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1507–1520, 2018.

[18] A. Von Jouanne and B. Banerjee, “Assessment of voltage unbalance,” *IEEE Trans. Power Deliv.*, vol. 16, no. 4, pp. 782–790, 2001.

[19] G. Ejebe and B. F. Wollenberg, “Automatic contingency selection,” *IEEE Trans. on Power Apparatus and Systems*, no. 1, pp. 97–109, 1979.

[20] C. A. Castro and A. Bose, “Correctability of voltage violations in on-line contingency analysis,” *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1651–1657, 1994.

[21] C. Muscas, S. Sulis, A. Angioni, F. Ponci, and A. Monti, “Impact of different uncertainty sources on a three-phase state estimator for distribution networks,” *IEEE Trans. Instrum. Meas.*, vol. 63, no. 9, pp. 2200–2209, Sep. 2014.

[22] P. J. Roebber, “Visualizing multiple measures of forecast quality,” *Weather and Forecasting*, vol. 24, no. 2, pp. 601–608, 2009.

[23] M. Cui, J. Zhang, A. R. Florita, B.-M. Hodge, D. Ke, and Y. Sun, “An optimized swinging door algorithm for identifying wind ramping events,” *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 150–162, Jan. 2016.

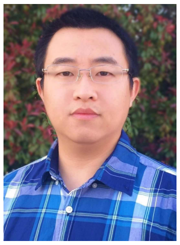
[24] S. Fan, G. He, X. Zhou, and M. Cui, “Online optimization for networked distributed energy resources with time-coupling constraints,” *IEEE Trans. Smart Grid*, 2020, in press.

[25] C. Chen, M. Cui, F. Li, S. Yin, and X. Wang, “Model-free emergency frequency control based on reinforcement learning,” *IEEE Trans. Ind. Inform.*, 2020, in press.

[26] Non-Gaussian process generation. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/52193-non-gaussian-process-generation

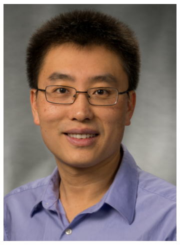
[27] The Mathworks, Inc. Nonlinear least square solver. [Online]. Available: https://www.mathworks.com/help/optim/ug/lsqnonlin.html.

[28] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 847–855, 2013.



Mingjian Cui (S'12–M'16–SM'18) received the B.S. and Ph.D. degrees from Wuhan University, Wuhan, Hubei, China, all in Electrical Engineering and Automation, in 2010 and 2015, respectively.

Currently, he is a Research Assistant Professor at Southern Methodist University, Dallas, Texas, USA. He was also a Visiting Scholar from 2014 to 2015 in the Transmission and Grid Integration Group at the National Renewable Energy Laboratory (NREL), Golden, Colorado, USA. His research interests include renewable energy, power system operation, power system cybersecurity, power system data analytics, and machine learning. He has authored/coauthored over 60 peer-reviewed publications. Dr. Cui serves as an Associate Editor for journals of IET SMART GRID, IEEE ACCESS, IEEE POWER ENGINEERING LETTERS, and IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY (OAJPE). He is also the Best Reviewer of IEEE TRANS. SMART GRID for 2018 and IEEE TRANS. SUSTAINABLE ENERGY for 2019.



Jianhui Wang (M'07–SM'12) received the Ph.D. degree in electrical engineering from Illinois Institute of Technology, Chicago, Illinois, USA, in 2007.

Dr. Jianhui Wang is a Professor with the Department of Electrical and Computer Engineering at Southern Methodist University. Dr. Wang has authored and/or co-authored more than 300 journal and conference publications, which have been cited for more than 20,000 times by his peers with an H-index of 79. He has been invited to give tutorials and keynote speeches at major conferences including IEEE ISGT, IEEE SmartGridComm, IEEE SEGE, IEEE HPSC and IGEC-XI.

Dr. Wang is the past Editor-in-Chief of the IEEE Transactions on Smart Grid and an IEEE PES Distinguished Lecturer. He is also a guest editor of a Proceedings of the IEEE special issue on power grid resilience. He is the recipient of the IEEE PES Power System Operation Committee Prize Paper Award in 2015 and the 2018 Premium Award for Best Paper in IET Cyber-Physical Systems: Theory & Applications. Dr. Wang is a 2018 and 2019 Clarivate Analytics highly cited researcher for production of multiple highly cited papers that rank in the top 1% by citations for field and year in Web of Science.