$N-1$ Reliability Makes It Difficult for False Data Injection Attacks to Cause Physical Consequences

Zhigang Chu, Jiazi Zhang, Oliver Kosut, and Lalitha Sankar

***Abstract*—This paper demonstrates that false data injection (FDI) attacks are extremely limited in their ability to cause physical consequences on $N-1$ reliable power systems operating with real-time contingency analysis (RTCA) and security constrained economic dispatch (SCED). Prior work has shown that FDI attacks can be designed via an attacker-defender bi-level linear program (ADBLP) to cause physical overflows after re-dispatch using DCOPF. In this paper, it is shown that attacks designed using DCOPF fail to cause overflows on $N-1$ reliable systems because the system response modeled is inaccurate. An ADBLP that accurately models the system response is proposed to find the worst-case physical consequences, thereby modeling a strong attacker with system level knowledge. Simulation results on the synthetic Texas system with 2000 buses show that even with the new enhanced attacks, for systems operated conservatively due to $N-1$ constraints, the designed attacks only lead to post-contingency overflows. Moreover, the attacker must control a large portion of measurements and physically create a contingency in the system to cause consequences. Therefore, it is conceivable but requires an extremely sophisticated attacker to cause physical consequences on $N-1$ reliable power systems operated with RTCA and SCED.**

***Index Terms*—False data injection attack, cyber-security, vulnerability of $N-1$ reliable power systems, bi-level optimization.**

I. Introduction

The efficiency and intelligence of modern electric power systems are increasing rapidly with integration of real-time monitoring, sensing, communication and data processing. This integration is accomplished via a cyber layer consisting of the supervisory control and data acquisition (SCADA) system in conjunction with the energy management system (EMS). SCADA monitors the physical system, collects measurements, and sends them to the control center. In the EMS, state estimation (SE) estimates the voltage magnitudes and angles from measurements. This estimate along with the subsequent data processing, optimization and communication, specifically real-time contingency analysis (RTCA) [1] and security constrained economic dispatch (SCED) [2], allow for real-time control of the power systems.

However, the integration of the cyber layer also increases the threat of cyber-attacks on power systems that could lead to severe physical consequences, as illustrated by the recent cyber-attack in Ukraine (see [3]). Therefore, it is crucial to develop techniques to detect and thwart potential attacks, which requires evaluating system vulnerability to credible attacks. Assessing consequences of possible attacks is extremely instructive for system operators, and is important for secure power system operations.

Related work: This paper focuses on unobservable false data injection (FDI) attacks, wherein a malicious attacker replaces a subset of SCADA measurements (power flows and injections) with counterfeits. A wealth of research effort has been undertaken on FDI attacks, showing that they can be designed to target system states [4]–[6], system topology [7], [8], and energy markets [9]. They can bypass the bad data detector (BDD) embedded within SE, and change the load data used for re-dispatch, which in turn cause physical and/or economic consequences. Note that there are "detectable" attacks that can change the SE results despite failing the BDD as shown in [10], but in this paper we focus on worst-case scenarios by assuming powerful attackers that can launch unobservable attacks. Many existing work evaluating the worst-case attack consequences involve solving attacker-defender bi-level linear programs (ADBLPs), wherein the first level models the attacker's objective and limitations (*e.g.*, number of measurements to change), while the second level models the system response under attack via DC optimal power flow (OPF). Examples include attacks that cause line overflows [11], locational marginal price (LMP) changes [12], operating cost increases [13] and sequential outages [14]. The authors of [15] analyzes the physical consequences when the attacker only has limited information, and [16] and [17] focus on cyber-physical coordinated attacks. The authors of [18] propose an ADBLP to find FDI attacks that add or drop contingency pairs with minimum attack effort, and analyze the economic effect of such attacks on LMPs. Rahman *et al.* [19] demonstrate several case studies to showcase the impact of FDI attacks on contingency analysis, but their approach is not optimization-based, which means that it does not consider worst-case scenarios. Both [18] and [19] consider simplified SCED as system response, but the only addition of their SCED to DCOPF is the contingency case line power flow constraints modeled using DC line outage distribution factors (LODFs), while other SCED constraints such as reserve and ramp rate constraints are not considered.

Despite this prior research, there remains a need to evaluate physical consequences of FDI attacks that take into account *detailed* models for the system response, including RTCA and SCED. To understand the worst case consequences, we evaluate the vulnerability of $N-1$ reliable power systems, by modeling a powerful attacker that has system level knowledge

This work is supported by theNational Science Foundation (NSF) under grants number CNS- 1449080 and OAC-1934766. Paper no. TPWRS-00170-2020.

The authors are with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287, US (e-mail: zchu2@asu.edu; jzhan188@asu.edu; okosut@asu.edu; lalithasankar@asu.edu).

and capabilities. In particular, we focus on unobservable FDI attacks that aim to maximize the power flow on a target line after re-dispatch [11]. The authors of [11] design such attacks by solving an ADBLP modeling DCOPF as the system response and demonstrate that they can cause physical overflows. However, we found in our experiments that these attacks fail to cause overflows on systems operating with RTCA and SCED. This observation leads to another question: can attacks designed with complete knowledge of operations lead to more consequences? Note that answering this question inherently focuses on very strong attackers, as in general there is no universally adopted formulation of SCED, and we assume the attacker knows the SCED formulation for the particular system that it is attacking. Our goal of modeling such strong attackers is to understand if the grid is resilient to such worst-case attacks. Solving the attack design ADBLP is challenging due to its non-convexity, especially on large-scale systems. In [20], we introduce a modified Benders' decomposition (MBD) algorithm that can efficiently solve any ADBLP regardless of the problem size. The MBD algorithm converts an ADBLP into a single level problem using duality theory, and then decomposes the single level problem into a main problem and a sub problem (both are linear programs (LPs)) and solve them iteratively until convergence to obtain the solution to the ADBLP. Note that [20] proposes the MBD algorithm from optimization perspective and showcases its effectiveness by solving the ADBLP in [11], while this paper focuses on the effect of $N-1$ reliability on the physical consequences of the attacks. The resulting attacks are tested on the synthetic Texas system [21] with 2000 buses to demonstrate the difficulty of causing physical consequences. Our results show that $N-1$ reliability achieved by RTCA and SCED leads to more conservative operation, making it hard for FDI attacks to cause any pre-contingency overflows, even for the above-mentioned strong attacker. The attacks may still cause post-contingency overflows, but this requires the attacker to perform a cyber-physical coordinated attack by physically creating a contingency. Furthermore, as we show later, these sophisticated attacks also require the attacker to control measurements in a large portion of the system, which is again difficult to achieve in practice.

To summarize, the key contributions of this paper are as follows:

1. We showcase that attacks designed without considering EMS operations including RTCA and SCED do not cause the physical consequences intended by the attacker.

2. Given this observation, we propose an ADBLP modeling SCED as the system response, assuming an extremely strong attacker who has perfect knowledge of EMS operations including RTCA and SCED.

3. We provide simulation results on the synthetic Texas system with 2000 buses. We find that the resulting attacks can only cause post-contingency overflows.

4. We highlight that even the aforementioned powerful attacker must control a large number of measurements and physically create a contingency to cause overflows.

The remainder of this paper is organized as follows. Sec. II describes the power system measurement model and unobservable attack model. Sec. III demonstrates that attacks designed with DCOPF are extremely limited in their ability to have expected consequences if the system re-dispatch using RTCA and SCED. Sec. IV details the knowledge and capabilities of the worst-case attacker, and introduces an ADBLP modeling SCED as system response to find worst-case attacks. Sec. V illustrates the simulation results on the synthetic Texas system. Concluding remarks and future work are presented in Sec. VI.

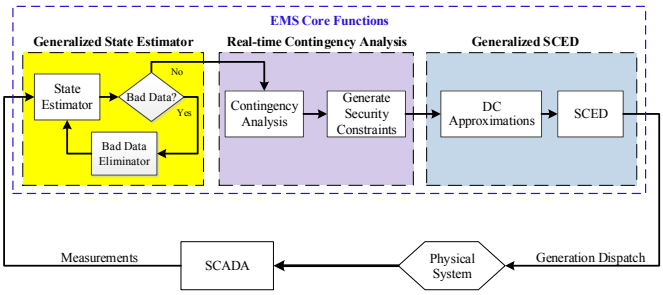

Fig. 1: EMS operation with SE, RTCA, and SCED.

II. System and Attack Model

A. EMS Operation

In this paper, we consider an EMS with three core functions operating in the order of SE, RTCA, and SCED. The EMS operating structure is illustrated in Fig. 1. Power system measurement data collected by SCADA are sent to SE, which estimates the complex voltages after eliminating noise and bad measurements. Given the generator set points, load values are estimated based on SE results. Modern power systems typically require $N-1$ reliability, *i.e.,* the system must operate with no violations if a contingency occurs (one of the system components, generators or branches, is out of service). RTCA simulates one power flow under each contingency k. We say a branch has a *warning* if its power flow is above a threshold τ but less than its limit, while a branch has a *violation* if its power flow exceeds its limit. Both "warning" and "violation" branches are denoted *critical branches*. For post-contingency critical branches, their corresponding contingencies are called *critical contingencies*. Note that in base case, the branch limits are the long-term ratings, while in contingency case they are the short-term ratings. RTCA generates one security constraint to be modeled in SCED for each warning and violation. SCED takes all security constraints, along with other common constraints including reserve and ramp rates, to solves an optimization problem to determine the most economic generation dispatch that ensures $N-1$ reliability.

B. Measurement Model

We model the power system with n_b buses, n_g generators, and n_m measurements. The SCADA system measurement model is given by

$$z = h(x) + e \tag{1}$$

where z is the $n_m \times 1$ measurement vector; x is the $2n_b \times 1$ vector of bus voltage magnitudes and angles (states); $h(\cdot)$ is the non-linear relationship between measurements and states; e is the $n_m \times 1$ vector of measurement noise, whose entries

are assumed to be jointly distributed as $\mathcal{N}(0,R)$ where $R = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_{n_m}^2)$.

C. Unobservable Attack Model

An $n_m \times 1$ false measurement vector $\bar{z}$ is defined to be *unobservable* to the traditional residual-based BDD if

$$\bar{z} = h(x + c) + e, \tag{2}$$

where c is the attack vector [6]. These false measurements cannot be distinguished from the true measurements if the true states are $x+c$, and hence, cannot be detected by the traditional BDD. Given c, an attack subgraph $\mathcal{S}$ can be constructed as in [6], such that the non-zero entries of $\bar{z} - z$ are all within $\mathcal{S}$. Launching such an attack requires the attacker to gain control of all measurements within $\mathcal{S}$. By modifying measurements in $\mathcal{S}$, the attacker can arbitrarily spoof the states of center buses (load buses corresponding to non-zero entries of c) without detection. The attack causes the system estimated loads to re-distribute between load buses within $\mathcal{S}$, while the total load remain unchanged.

III. Consequences of Attacks Designed with DCOPF on $N-1$ Reliable System

In this section, we demonstrate that attacks designed without considering RTCA and SCED (as in many existing literatures) do not cause expected physical consequences on systems operated as outlined in Fig. 1. The attacker's capability assumptions and the attack design ADBLP are adopted from [11]. The purpose of the attacker is to maximize the physical power flow on a target line after re-dispatch, and possibly cause overflow. The attacker is assumed to have knowledge of: (i) the complete network topology (including line parameters and ratings) and load information, and (ii) the cost, capacity, and operational status of all generators in the system. The formulation of this ADBLP is given by

$$\underset{c}{\text{maximize}} \quad P_l - \sigma \left\| c \right\|_1 \tag{3a}$$

subject to

$$\left\| c \right\|_1 \leq N_1 \tag{3b}$$

$$-L_S P_D \leq Hc \leq L_S P_D \tag{3c}$$

$$P_l = \text{PTDF}^l (G_B P_G^* - P_D) \tag{3d}$$

$$\{P_G^*\} = \arg \left\{ \min_{P_G} C_G \left(P_G \right) \right\} \tag{3e}$$

subject to

$$\textstyle\sum_{g=1}^{n_g} P_{Gg} = \sum_{i=1}^{n_b} P_{Di} \tag{3f}$$

$$-P_{\max} \leq \text{PTDF}(G_B P_G - P_D + Hc) \leq P_{\max} \tag{3g}$$

$$P_{G,\min} \leq P_G \leq P_{G,\max} \tag{3h}$$

where the variables are:

c	attack vector, $n_b \times 1$;
P_l	physical power flow on target line l;
P_G	power output of generators, $n_g \times 1$;

and the parameters are:

σ	penalty of the l_1-norm of attack vector c;
G_B	generators to buses connectivity matrix, $n_b \times n_g$;
N_1	attack vector l_1-norm limit;
L_S	load shift factor, in percentage;
H	dependency matrix between power injection measurements and states, $n_b \times n_b$;
P_D	vector of real loads, $n_b \times 1$;
C_G	generation cost vector, $n_g \times 1$;
PTDF	power transfer distribution factor matrix;
PTDF^l	the l^{th} row of PTDF matrix;
$P_{\max}$	vector of base case line limits;
$P_{G,\min}$	generation lower limits vector, $n_g \times 1$;
$P_{G,\max}$	generation upper limits vector, $n_g \times 1$.

In DCOPF, the voltage magnitudes are all considered to be 1 p.u., and hence, c is an $n_b \times 1$ attack vector on the voltage angles. The objective function (3a) is to maximize the physical power flow on target line l, and the second term penalizes the l_1-norm of attack vector c, such that if there exists multiple optimal solutions, the one with the smallest $\|c\|_1$ will be selected. Constraint (3b) limits the attacker's resources. Ideally, this should be characterized by the number of states that can be changed by the attacker, which is the l_0-norm of c. However, l_0-norm is non-convex and intractable, here we use l_1-norm as a proxy. (3d) calculates physical power flows from the optimal generation dispatch under attack. (3c) characterize the detectability of the attacks in terms of load shift, because loads that deviating too much from their true values are easily detectable. Note that Hc is a DC approximation of the injection measurement changes caused by the attack, because the AC relationship $h(\cdot)$ is non-convex. (3e)-(3h) are DCOPF under attack.

The ADBLP (3) is solved using the MBD algorithm that we introduced in [20]. It is briefly summarized as follows:

1. Rewrite the defender's problem as its primal-dual optimality conditions. According to duality theory, the primal objective of an LP equals to its dual objective only at optimal. Thus, the primal problem constraints, the dual problem constraints, along with "primal objective = dual objective" define the primal-dual optimality conditions of the defender's problem. Once this is done the ADBLP is converted to a single level problem.

2. Decompose the single level problem into a main problem and a sub problem. The main problem contains everything that only related to the attack vector c, and its objective is $\alpha - \sigma\|c\|_1$, where α is a variable introduced to represent P_l. The objective of the sub problem is P_l, and its decision variables include P_G, as well as the dual variables of the defender's problem. The attacker vector c in the sub problem is treated as constant, so that both the main problem and the sub problem are LPs.

3. Solve the sub problem with $c = 0$ to add an optimality cut or a feasibility cut to the main problem, and then solve the main problem to obtain a new c to plug in the sub problem. Perform this process iteratively until convergence, *i.e.,* α from the main problem equals P_l from the sub problem.

It is illustrated in [11] that the attacks obtained by solving (3) can cause physical overflows if the system re-dispatches using DCOPF. However, modern EMSs typically operate as outlined in Fig. 1. Thus, if the attacker merely solves (3), it will not accurately predict the system response, and the re-

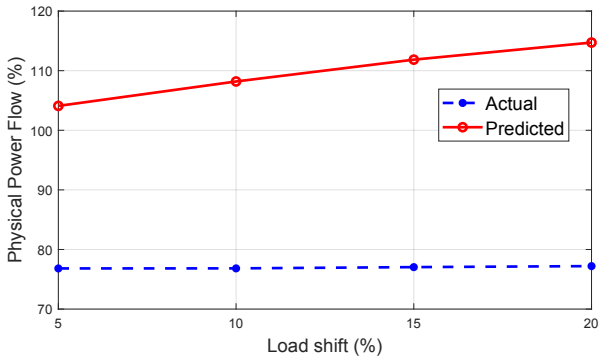

Fig. 2: Consequence of attacks designed with DCOPF on $N-1$ reliable synthetic Texas system, $N_1 = 2$.

dispatch after attack may not cause expected consequences. We have found in our experiments that attacks designed with DCOPF cannot cause any overflows on the synthetic Texas system operating with RTCA and SCED *even in the peak load scenario*. To illustrate this, consider the following example from our experiments.

The attacker continuously monitors the system operating status, and at the peak load hour, it observes that the most critical branch is transformer "tx-3083-3082" with a power flow of 76.72%. It selects this branch as the target and uses (3) (that is, modeling the system response via DCOPF) to obtain the attack vector c as well as the predicted physical power flow. It finds that the predicted flow exceeds the rating. Hence, it creates false measurements $\bar{z} = h(\hat{x} + c)$ to launch an attack. The system estimates loads from $\bar{z}$, and performs RTCA and SCED to find the optimal generation dispatch (details of RTCA and SCED are given in Sec. IV-B). Applying the new dispatch on the real loads yields the actual physical power flows. Fig. 2 illustrates a comparison between the attacker's predicted physical power flows and the actual flows on this target branch as a function of load shifts L_S, with $N_1 = 2$.

From this figure, we can see that the attacker predicted power flows exceed the rating of the branch for every load shift, but the actual flows are not affected. This is because in the pre-attack DCOPF solution, the target branch is congested. The attack redistributes the loads in the system, making it appear that the flow on this branch is reduced. The higher the load shift, the more the reduction on the flow. Thus, DCOPF will re-dispatch the generations to increase the flow on this branch, making it congested again. This will overload the branch in the physical system, since the real loads are not changed. However, SCED models more constraints than DCOPF does, and this branch is congested in neither base case nor contingency cases. The load redistribution caused by the attack does not affect any binding constraints in SCED, and hence, has no effect on the re-dispatch. We have experimented on the 5 branches with highest base case flows, and observed similar consequences.

IV. Worst-case Attacks

A. Attacker Assumptions

The observations illustrated in the previous section lead to the following new question: if the attacker knows the system operation details, can it cause physical consequences through FDI attacks? To this end, we model the worst-case attacker who has knowledge of system EMS operations. In other words, the attacker is able to perform the same RTCA and SCED as the system does, and hence, can design attacks that maximize the consequences. This is a very strong assumption, because in addition to having access to the database of the control center, now the attacker further knows the algorithms and assumptions used by the system. While this assumption may be impractical, we aim to understand whether $N-1$ reliable system is resilient against such strong adversaries through this worst-case approach.

In order to accurately predict the system response under attack, the attacker needs to know all the constraints modeled in SCED. This requires the attacker to gain knowledge of the power flow algorithm used in RTCA to get the same post-contingency flows on all branches, as well as the threshold τ as described in Sec. II, to determine the security constraints to be included in SCED. In addition to these security constraints, the attacker must know the detailed modeling of other SCED constraints, as different systems may have different SCED implementations. We assume the attacker has full knowledge of RTCA and SCED implementations, in particular:

1) Contingency ratings of the branches;
2) Loss handling method;
3) Ramp rates and reserve costs of all generators;
4) Reserve policy and requirements;
5) Criteria to determine which base case line limits are to be modeled. This can be the same threshold as τ in post-contingency case, but can also be different;
6) Branch flow calculation method in both base case and contingency case;
7) Load shedding policy and costs.

Although it is not entirely impossible to have this level of knowledges [22], since such complex attacks often involve sophisticated (even nation-state) attackers that can exploit or have access to insider knowledge [3], [23], it is still extremely hard to have such a strong attacker in practice. However, this is the worst-case assumptions from the optimization perspective, because the attacker can most accurately predict the system response by modeling exactly same SCED. Modeling less constraints or relaxing any of those constraints will increase the feasible region of SCED, and hence, exaggerate the attacker predicted consequences. Understanding the vulnerability of power system to such worst-case attacks can serve as an upper bound on risks to system operations.

B. ADBLP to Find Worst-case Attacks

The worst-case line overflow attacks can be found using an ADBLP similar to (3). The first level models the attacker's objective and limitations, while the second level models the system response via SCED. The security constraints are generated by an RTCA that simulates branch contingencies, excluding radial branches. Contingency k indicates that branch k is out of service. The attacker can choose critical branches in either base case or contingency case as target branch. Without

loss of generality, we assume the flow on l is positive; if it is not the case, its absolute value can be maximized.

The ADBLP takes the following form:

$$\underset{c}{\text{maximize}} \quad P_l - \sigma \|c\|_1 \quad \text{or} \quad P_{l,k_t} - \sigma\|c\|_1 \tag{4a}$$

subject to

$$\|c\|_1 \le N_1 \tag{4b}$$

$$-L_S P_D \le Hc \le L_S P_D \tag{4c}$$

$$P_l = \text{PTDF}^l(G_B P_G^* - P_D) \tag{4d}$$

$$P_{l,k_t} = \text{OTDF}_{k_t}^l(G_B P_G^* - P_D) \tag{4e}$$

$$\{P_G^*\} = \arg\left\{\min_{P_G,R_G,\bar{P},\bar{P}_k} C_G(P_G) + C_R R_G\right\} \tag{4f}$$

subject to

$$\textstyle\sum_{g=1}^{n_g} P_{Gg} = \sum_{i=1}^{n_b} P_{Di} \tag{4g}$$

$$\bar{P} = P_0 + \text{PTDF}(G_B(P_G - P_{G0}) + Hc) \tag{4h}$$

$$\bar{P}_k = P_{k0} + \text{OTDF}_k(G_B(P_G - P_{G0}) + Hc) + \text{LODF}_k \cdot \text{PTDF}^k \cdot Hc, \forall k \tag{4i}$$

$$-P_{\max} \le \bar{P} \le P_{\max} \tag{4j}$$

$$-P_{k,\max} \le \bar{P}_k \le P_{k,\max}, \forall k \tag{4k}$$

$$P_G \ge \max\{P_{G0} - M_G T_h, P_{G,\min}\} \tag{4l}$$

$$P_G \le \min\{P_{G0} + M_G T_h, P_{G,\max}\} \tag{4m}$$

$$0 \le R_G \le M_G T_r \tag{4n}$$

$$P_G + R_G \le P_{G,\max} \tag{4o}$$

$$\textstyle\sum_{g=1}^{n_g} R_{Gg} \ge P_{Gg} + R_{Gg} \ , \forall g \tag{4p}$$

In addition to the variables and parameters introduced in (3), the new notations are listed as follows.

$\bar{P}, \bar{P}_k$	vectors of monitored line cyber power flows in base case and under contingency k, respectively;
P_{l,k_t}	physical power flow on target line l under target contingency k_t;
R_G	spinning reserve of the generators, $n_g \times 1$;
OTDF_k	outage transfer distribution factor matrix under contingency k;
OTDF_k^l	l^{th} row of OTDF_k;
C_R	reserve cost vector, $n_g \times 1$;
P_0, P_{k0}	vectors of pre-SCED monitored line power flows in base case and under contingency k, respectively;
P_{G0}	pre-SCED generator outputs, $n_g \times 1$;
PTDF^k	k^{th} row of PTDF;
LODF_k	line outage distribution factors of monitored lines under contingency k;
$P_{k,\max}$	vector of line limits under contingency k;
M_G	ramp rates of all generators, $n_g \times 1$;
T_h	look-ahead time for one period SCED;
T_r	time for spinning reserve requirement.

The attacker's limitations (4b)-(4c) are the same as those in (3). (4d) and (4e) are the physical power flows on line l under base case and under target contingency k_t, respectively. The OTDF_k matrix is computed using the same method as computing the PTDF matrix by setting branch k out of service. The second level SCED (4f)-(4p) models the system response to the attack. The SCED is a linearized approximation of ACOPF, and we model it in a "hot start" fashion to reduce the AC-DC discrepancy. The variables of the SCED problem are generation dispatch P_G, reserve schedules R_G, base case power flows $\bar{P}$, and contingency case power flows $\bar{P}_k$, where $\bar{P}$ and $\bar{P}_k$ are functions of P_G. The reserve is modeled to prepare for unexpected load increase and/or generator contingencies. The objective of the operator (4f) is to minimize the total cost, consisting of generation cost and reserve cost; constraint (4g) is the power balance equation; (4h) is the cyber power flow of the base case monitored lines. Here P_0 is the vector of base case pre-SCED branch flows obtained from RTCA, and is non-linearly related to the pre-SCED generation P_{G0}. Therefore, only the change in base case branch flows, $\bar{P} - P_0$, are linearly related to the generation change $P_G - P_{G0}$, and the AC-DC discrepancy is less than that if this constraint is modeled as (3g). Note that this constraint is only modeled for critical lines whose pre-SCED power flow is greater than the threshold τ, *i.e.,* $|P_0/P_{\max}| \ge \tau$. This is under the assumption that the line flows will not change dramatically after the SCED re-dispatch, due to the ramping constraints of the generators. Similarly, (4i) is the cyber power flows on monitored lines under each contingency k, where P_{k0} is obtained from RTCA and $|P_{k0}/P_{k,\max}| \ge \tau$. Here we assume the base case and contingency case monitoring thresholds are the same. In the right hand side of (4i), the first term is the pre-SCED post-contingency flows; the second term is the change of the flows as a result of re-dispatch and false loads; the third term quantifies the amount of power on the monitored lines resulting from the effect of false loads on the contingency line k, which is not considered in P_{k0}. Constraints (4j) and (4k) are the line limits in base case and contingency case, respectively. The active power limits in both base case and contingency cases, $P_{\max}$ and $P_{k,\max}$, are approximated from the MVA ratings and reactive flows on the branches by

$$P_{\max} = \sqrt{S_{\max}^2 - [\max(Q_{\text{from}}, Q_{\text{to}})]^2} \tag{5}$$

$$P_{k,\max} = \sqrt{S_{k,\max}^2 - [\max(Q_{k,\text{from}}, Q_{k,\text{to}})]^2} \tag{6}$$

where $S_{\max}$ and $S_{k,\max}$ are branch long-term and short-term ratings, respectively; Q_{from} and Q_{to} are the base case reactive branch flows at the "from" end and "to" end, respectively; $Q_{k,\text{from}}$ and $Q_{k,\text{to}}$ are those flows in contingency cases. This is an additional approach to reduce the AC-DC discrepancy. Constraints (4l) and (4m) are the ramp rate limits. They limit the generation change within one SCED period. Typically, the ramping capabilities of generators are limited, as they need time to change their output. For example, a generator with a 2 MW/min ramp rate can increase/decrease its generation at most 30 MW in 15 minutes. We assume the ramp up and down rates are the same for all generators. However, the output of a generator must be no less than its minimum generation limit $P_{G,\min}$, and no more than its capacity $P_{G,\max}$. (4n) is the reserve limit, which models how much reserve can the generators provide within the spinning reserve requirement time T_r. The reserves need to be activated in situations such as a demand increase or generator contingency. The generators whose reserves are activated must increase their output in order

to meet the demand within the reserve requirement time T_r. Hence, the amount of reserve that a generator can provide is also ramp-limited. (4o) is the generation-reserve limit that constrains the total energy and reserve provided by a generator to be less than its capacity. Though the RTCA does not simulate generator contingencies, in SCED it is required that when a generator is out, the reserves of all other generators are sufficient to cover the output of the lost generator. This system reserve requirement is captured in (4p).

With knowledge of system RTCA and SCED, the attacker can wisely select the target branch, so that the constraints associated with this branch is binding or nearly binding in the pre-attack SCED solution. Therefore, the false loads can mislead the SCED to re-dispatch the generation to increase the flow on the target branch, and possibly cause overflow. Solving the ADBLP (4) provides the attack vector c and resulting physical power flows, which allow for evaluating the vulnerability.

C. Attack Implementation

Fig. 3 illustrates the implementation of the attack and the vulnerability assessment approach. We assume the attacker aims to cause post-contingency overflows, and the real loads remain unchanged during the attack period. The physical system behavior and the SCADA measurement collection are simulated by solving an AC power flow. The true measurements z_1 from the power flow solution are acquired by the attacker to estimate the states (denoted $\hat{x}_1$). It then performs RTCA to achieve the security constraints and solves the attack design ADBLP to find the attack vector c. Recall that the second level of the ADBLP is a SCED in response to the attack, and by solving it the attacker obtains the predicted maximal physical power flow on the target branch, which is the optimal objective P^*_{l,k_t}. To implement the designed attack, the attacker then constructs false measurements $\bar{z}_1 = h(\hat{x}_1 + c)$ and injects $\bar{z}_1$ to the system SE instead of the true measurements z_1. Again, only the measurements in the attack subgraph $\mathcal{S}$ are changed. Since the generator outputs are known to the system, the false measurements will cause the SE to estimate a set of false loads. RTCA and SCED are then performed by the system to determine the new optimal generation dispatch P^*_G in response to the false loads. Once the generators re-dispatch, the true measurements changed to z_2. To sustain the attack on the system, the attacker again acquires z_2, and estimates the new states $\hat{x}_2$. It then sends $\bar{z}_2 = h(\hat{x}_2 + c)$ to the system SE to estimate new false loads. The system operator again runs RTCA with the new false loads and observes the cyber power flow $\bar{P}_{l,k_t}$. However, the new dispatch applied on the physical system, will maximize the physical power flow on target line l under target contingency k_t, and possibly cause overflow. The true physical power flow, P_{l,k_t}, is obtained by running RTCA with the new dispatch on real loads.

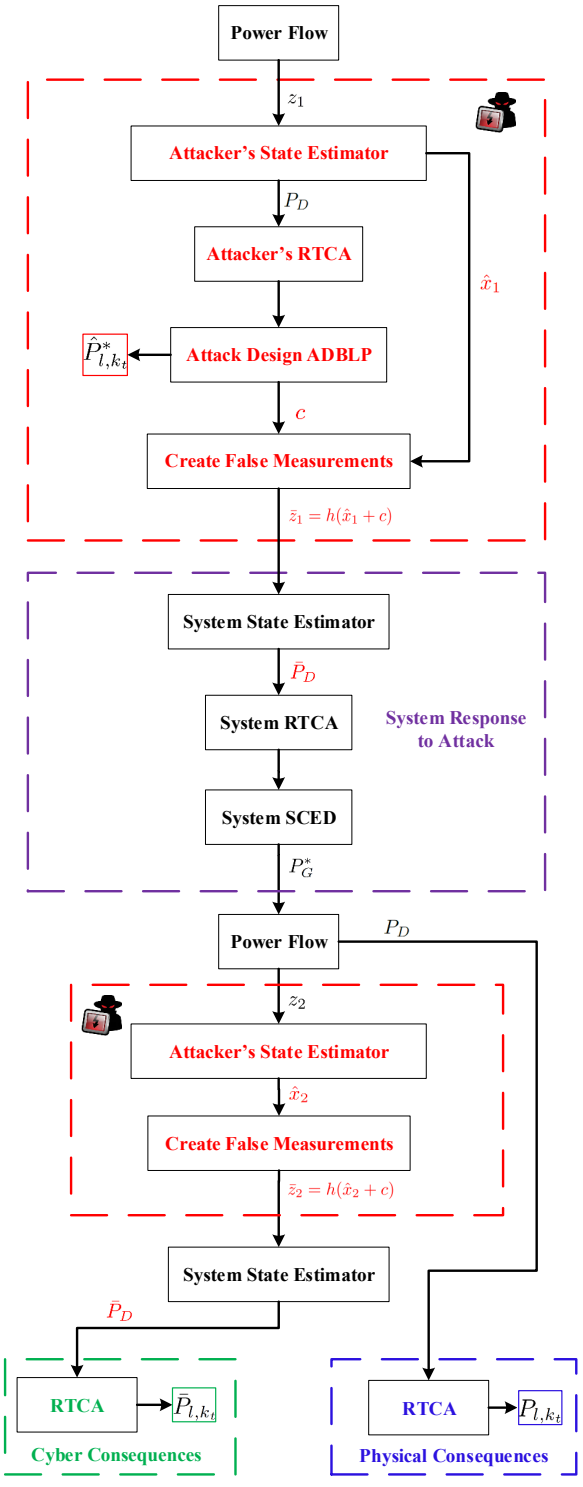

Fig. 3: Attack implementation and system vulnerability assessment approach.

V. Simulation Results and Discussions

In this section, we present physical consequences through simulations of the attacks designed using ADBLP (4). We use the synthetic Texas system with 2000 buses, 3210 branches,

and 432 generators [21]. The base case power flow and RTCA are performed using OpenPA [24], a Java-based EMS simulation platform that we developed in collaboration with our industry partners IncSys [25] and PowerData [26]. Without attack, the system is operating at steady-state, which means that SCED does not change the generation dispatch between each EMS loop. In the base case power flow solution, the total losses among the system is 2% of the net load. We assume the SCED handles losses by uniformly increasing all loads by this percentage. RTCA simulates contingencies of all branches whose end bus voltages are both at least 100 kV, except radial branches. The short-term branch limit is assumed to be 115% of the long-term limit, *i.e.,* $S_{k,\max} = 115\% \times S_{\max}$; SCED look ahead time $T_h = 15$ minutes; spinning reserve time $T_r = 10$ minutes. The ADBLP is solved using a Modified Benders' decomposition (MBD) algorithm that we introduced in [20], which can efficiently solve large-scale ADBLPs. The ADBLP and the MBD algorithm are implemented in Matlab and solved with solver Gurobi. With a warning threshold $\tau = 90\%$, RTCA reports no base case critical branches, and 25 post-contingency critical branches before attack. We exhaustively design attacks targeting each of those 25 branches for post-contingency overflows with load shift $L_S = [10\%, 20\%]$, and l_1-norm constraint $N_1 = [0.2, 2]$ in steps of 0.2. All simulations are conducted on a 3.4 GHz PC with 32 GB RAM.

A. Results on Maximal Physical Power Flows

Fig. 4 compares physical power flow P^*_{l,k_t} predicted by the attacker, the true power flow P_{l,k_t} in the physical system, as well as the power flow (cyber) seen by the system operator $\bar{P}_{l,k_t}$, as a function of the l_1-norm constraint N_1. These power flows are plotted as percentage values relative to the active power limit $P_{l,k,\max}$ calculated using (6). The attacker's goal is to maximize the power flow on line 'ln-2025-2055' when line 'ln-2054-5236' is out of service.

The results indicate that the attacks cause post-contingency overflows. When the load shift $L_S = 10\%$, P^*_{l,k_t} and P_{l,k_t} increase as N_1 increases. When $L_S = 20\%$, similar results are observed, but P^*_{l,k_t} and P_{l,k_t} are not monotonically increasing as N_1 increases. This suggests that the MBD algorithm provides sub-optimal solutions, because as N_1 increases, the constraints are relaxed, and the optimal solution for a larger N_1 should be at least that of a smaller N_1. As expected, maximal physical power flow is higher when a larger load shift is allowed.

The true physical power flow P_{l,k_t} is slightly lower than the attacker predicted physical power flow P^*_{l,k_t}. One possible reason for this phenomenon is that the attacker is solving a DC approximation of an AC system, and the reactive power flow may change after attack. This could result in a difference in $P_{l,k,\max}$ before and after attack. Another possible reason is that the false measurements $\bar{z}_1$ injected by the attacker generate a different set of security constraints than those result from true measurements z. The attacker uses security constraints generated by pre-attack RTCA to solve the attack design ADBLP, but those constraints used in system SCED are based on the false measurements after attack. As a result, the

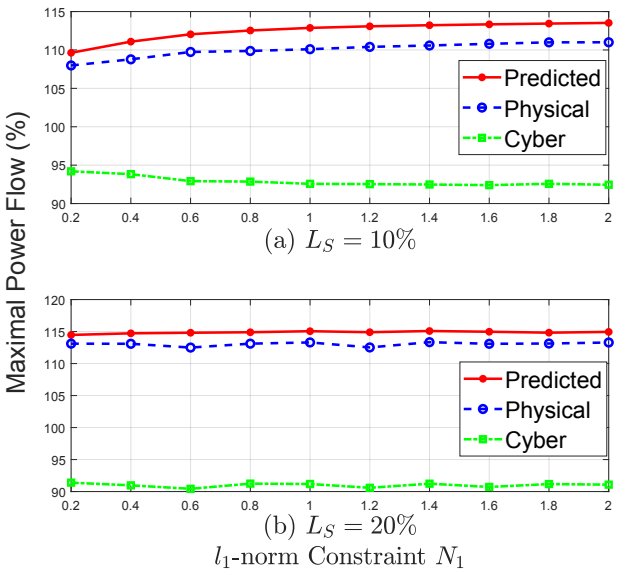

Fig. 4: Comparison of attacker predicted, physical, and cyber power flows on line 'ln-2025-2055' under contingency 'ln-2054-5236', (a) $L_S = 10\%$; (b) $L_S = 20\%$.

system SCED solution is different than the attacker predicted re-dispatch. One approach for the attacker to prevent this situation is to run its own RTCA using the false measurements and include any newly appeared security constraints into the attack design ADBLP, until there are no more new security constraints. However, this approach has no convergence guarantee, and could be too time-consuming.

Note that in order for the attacks to actually cause post-contingency violations requires a particular contingency to occur. Thus, the attacker has to physically create the target contingency itself, otherwise it has to wait for the contingency to occur. If the attacker is sufficiently powerful to physically cause contingencies, it may trip multiple lines to shut down the system, and there is no need for cyber-attacks in this situation. As far as we know, the probability of line failure is pretty low in practice. Thus, even though the attacks can cause post-contingency overflows, they can only put the system into an insecure state rather than cause physical damages, because of the difficulty in creating contingencies.

In the synthetic Texas system, there is no branch whose base case power flow is higher than τ prior to the attack. Thus, to cause base case overflow, the attacker has to shift a tremendous amount of load that may easily trigger an alarm at the control center. We have attempted to design a base case attack targeting top 5 branches with the highest base case power flow in percentage, but no overflow can be found even with $L_S = 90\%$ and $N_1 = 20$. This indicates that RTCA and SCED push the system to operate conservatively, which in turn decreases the vulnerability to line overflow attacks.

B. Results on Attack Resources

Fig. 5 illustrates the relationship between maximal power flow and l_0-norm of the attack vector (*i.e.* the number of center buses in the attack) versus the l_1-norm constraint N_1 for target

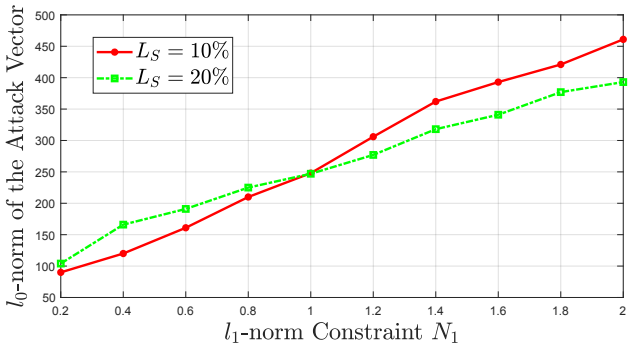

Fig. 5: Comparison of the l_0-norm of the attack vector for target line 'ln-2025-2055' under contingency 'ln-2054-5236'.

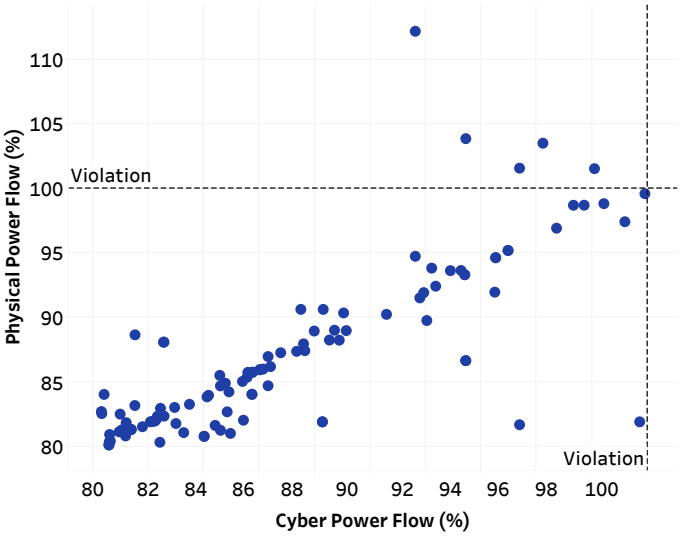

Fig. 6: Physical and cyber RTCA results after re-dispatch.

line 'ln-2025-2055' under contingency 'ln-2054-5236', with different load shift constraints. As N_1 increases, so does the l_0-norm of the attack, indicating that l_1-norm is a valid proxy for l_0-norm for our problem. If a larger load shift is allowed, the maximal power flow on target line increases, but the resulting l_0-norm may decrease for the same N_1. This indicates a trade-off between load shift and attacker's resources: as the attacker attempts to avoid detection by minimizing load changes, it will require control over a larger portion of the system to launch a comparable attack. These results also indicate that the attacker needs a tremendous amount of resources to launch the attacks. For example, with $L_S = 10\%$ and $N_1 = 1$, the attacker needs to change the state of 250 load buses. The corresponding attack subgraph $\mathcal{S}$ contains more than 800 buses, which is almost half the system. Thus, the attacker must control measurements in almost half of the system to successfully launch this attack, which is extremely hard to achieve.

C. Comparison of Physical and Cyber RTCA results

Fig. 6 compares the physical and cyber RTCA results after the re-dispatch resulting from an attack on target line 'ln-2025-2055' under contingency 'ln-2054-5236' with load shift $L_S = 10\%, N_1 = 2$. The cyber post-contingency power flows on the x-axis represent what the system operator observes, while the y-axis represents the post-contingency power flows in the physical system. There is no point beyond 100% of the x-axis, which indicates that the system operator sees no post-contingency violation after the attack. Therefore, the attack successfully spoofed the operator that the system is in a secure state, while in reality, the target line has a 112.2% post-contingency overflow. In addition, there are four post-contingency violations that are caused by the same attack, even though they are not the attacker's targets, but the overflow percentage are less. This observation indicates that the attack does put the system into an insecure state, and the system is no longer $N-1$ reliable under attack. However, overflows can only occur under contingencies.

D. Statistical Results on Attack Consequences

As mentioned at the beginning of Sec. V, we exhaustively tested attacks targeting the 25 branches with post-contingency warnings. The designed attacks successfully cause overflows on 8 out of the 25 target branches. Table I gives the statistical results on attack consequences of these 8 branches. We derived attacks using l_1-norm constraints in the range from $N_1 = 0.2$ to $N_1 = 2$. The table shows the resulting ranges in maximal power flow and l_0-norm of the attack vector c across this range. The load shift constraint $L_S = 10\%$. The prefix 'ln' indicates a transmission line and 'tx' indicates a transformer. From the maximal power flow range, we can see that some branches are more vulnerable than others, but most of the overflows are within 10%. Thus, even when the contingencies occur and the target branch becomes overloaded, it still takes time for it to heat up and trip. During this time, the contingencies may be eliminated by the system, and no physical damage can be dealt. Besides, the system operators can identify critical lines and critical contingencies for attack protection purposes. For example, they can artificially reduce the line limit to keep the attack from being successful. Measurements around vulnerable branches can be encrypted to prevent them from being modified. In our ADBLP, the load shift constraint characterizes the detectability of the attack, indicating that load abnormally detectors can help system operators distinguish between natural load changes and possible cyber attacks based on load redistribution.

TABLE I: Statistical Results on Maximal Physical Power Flow and l_0-norm of the Attack Vector with $N_1 \in [0.2, 2]$

Target	Contingency	Max PF (%)		$\|c\|_0$	
		N_1=0.2	N_1=2	N_1=0.2	N_1=2
ln-6188-7305	ln-7058-7095	101.92	105.08	133	442
ln-6240-6287	ln-6141-6239	102.43	106.76	137	314
ln-7233-7251	tx-6063-6062	105.41	107.90	156	485
ln-1003-1055	ln-3046-3078	102.80	102.94	163	520
ln-2025-2055	ln-2054-5236	107.98	111.00	90	461
ln-2070-5237	ln-2054-5236	101.35	104.35	90	461
ln-1003-1055	ln-1004-3133	102.43	102.56	160	513
ln-7059-7407	ln-7058-7406	100.38	102.24	154	488

VI. Conclusion

We have demonstrated that FDI attacks are extremely limited in their ability to cause physical consequences on power systems operated by EMSs consisting of SE, RTCA, and SCED to ensure $N-1$ reliability. For such systems, we showed that attacks designed with only DCOPF as the system response do not cause expected physical consequences. We then designed attacks by modeling the worst case attacker that can mimic the EMS operations including RTCA and SCED, and tested them on the synthetic Texas system. For this system, we showed that even for the above-mentioned strong attacker, the attacks still cannot cause base case overflows, because the system is pushed to operate conservatively with $N-1$ reliability requirement. The designed attacks can successfully cause post-contingency overflows on target branches, but it requires a specific contingency to occur to deal physical damage to the system. Moreover, the amount of resources required to launch such attacks is tremendous, and the contingencies can be fixed before the overloaded lines trip. Therefore, we argue that it is extremely hard for FDI attacks to cause physical damages to $N-1$ reliable systems. Future work will include investigating consequences of FDI attacks other than the line overflow attacks, studying dynamic impacts of FDI attacks, as well as designing countermeasures to detect, identify, and mitigate such attacks. Tracking the system dynamics may help detect attacks.

Acknowledgment

The authors would like to thank the following at ASU: Mr. Andrea Pinceti for creating the base case, Mr. Roozbeh Khodadadeh for help with the test platform, and Prof. Kory Hedman and his team for their support with RTCA and SCED. The authors also thank Dr. Robin Podmore (IncSys) and Mr. Christopher Mosier (Powerdata) for the OpenPA software.

References

[1] A. Mittal, J. Hazra, N. Jain, V. Goyal, D. P. Seetharam, and Y. Sabharwal, "Real time contingency analysis for power grids," in *ICPP 2011*, 2011.

[2] FERC, "Security constrained economic dispatch: definition, practices, issues and recommendations," Federal Energy Regulatory Commission, Tech. Rep., 2006. [Online]. Available: https://www.ferc.gov/industries/electric/indus-act/joint-boards/final-cong-rpt.pdf

[3] K. Zetter, "Inside the cunning, unprecedented hack of Ukraine's power grid," March 2016. [Online]. Available: http://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/

[4] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of the 16th ACM Conference on Computer and Communications Security*, ser. CCS '09, Chicago, Illinois, USA, 2009, pp. 21–32.

[5] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645–658, 2011.

[6] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.

[7] J. Zhang and L. Sankar, "Physical system consequences of unobservable state-and-topology cyber-physical attacks," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2016–2025, July 2016.

[8] X. Liu and Z. Li, "Local topology attacks in smart grids," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2617–2626, Nov 2017.

[9] R. Moslemi, A. Mesbahi, and J. M. Velni, "Design of robust profitable false data injection attacks in multi-settlement electricity markets," *IET Generation, Transmission Distribution*, vol. 12, no. 6, 2018.

[10] D. Deka, R. Baldick, and S. Vishwanath, "Optimal data attacks on power grids: Leveraging detection measurement jamming," in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2015, pp. 392–397.

[11] J. Liang, L. Sankar, and O. Kosut, "Vulnerability analysis and consequences of false data injection attack on power system state estimation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3864–3872, Sept 2016.

[12] L. Jia, J. Kim, R. J. Thomas, and L. Tong, "Impact of data quality on real-time locational marginal price," *IEEE Trans. Power Systems*, vol. 29, no. 2, pp. 627–636, 2014.

[13] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *Smart Grid, IEEE Transactions on*, vol. 2, no. 2, pp. 382–390, June 2011.

[14] L. Che, X. Liu, Z. Li, and Y. Wen, "False data injection attacks induced sequential outages in power systems," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1513–1523, March 2019.

[15] J. Zhang, Z. Chu, L. Sankar, and O. Kosut, "Can attackers with limited information exploit historical data to mount successful false data injection attacks on power systems?" *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4775–4786, Sep. 2018.

[16] H. Chung, W. Li, C. Yuen, W. Chung, Y. Zhang, and C. Wen, "Local cyber-physical attack for masking line outage and topology attack in smart grid," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4577–4588, July 2019.

[17] Y. Xiang, L. Wang, and N. Liu, "Coordinated attacks on electric power systems in a cyber-physical environment," *Electric Power Systems Research*, vol. 149, pp. 156 – 168, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378779617301700

[18] J. Kang, I. Joo, and D. Choi, "False data injection attacks on contingency analysis: Attack strategies and impact assessment," *IEEE Access*, vol. 6, pp. 8841–8851, 2018.

[19] M. A. Rahman, M. H. Shahriar, M. Jafari, and R. Masum, "Novel attacks against contingency analysis in power grids," 2019. [Online]. Available: arXiv:1911.00928

[20] Z. Chu, J. Zhang, O. Kosut, and L. Sankar, "Vulnerability assessment of large-scale power systems to false data injection attacks," in *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2020.

[21] "ACTIVSg2000: 2000-bus synthetic grid on footprint of Texas," Sep. 2017. [Online]. Available: https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg2000/

[22] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: Implications for false data injection attacks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3317–3318, 2017.

[23] "The stuxnet worm: A cyber-missile aimed at Iran," The Economist, Tech. Rep., 24 September 2010.

[24] "OpenPA." [Online]. Available: https://powerdata.com/openpa/

[25] "IncSys." [Online]. Available: http://www.incsys.com/

[26] "PowerData." [Online]. Available: https://powerdata.com/