Identification of Intraday False Data Injection Attack on DER Dispatch Signals

Jip Kim[1,3], Siddharth Bhela[2], James Anderson[1], Gil Zussman[1]

[1]Department of Electrical Engineering, Columbia University
[2]SIEMENS Technology
[3]Korea Institute of Energy Technology

Background

- **Deployment** of various DERs in Power Grids:
  - Rapid deployment of **distributed energy resources (DERs)**
  - **Power system operation** heavily relies on **information communication technologies (ICT)** → <u>Increase vulnerability</u>

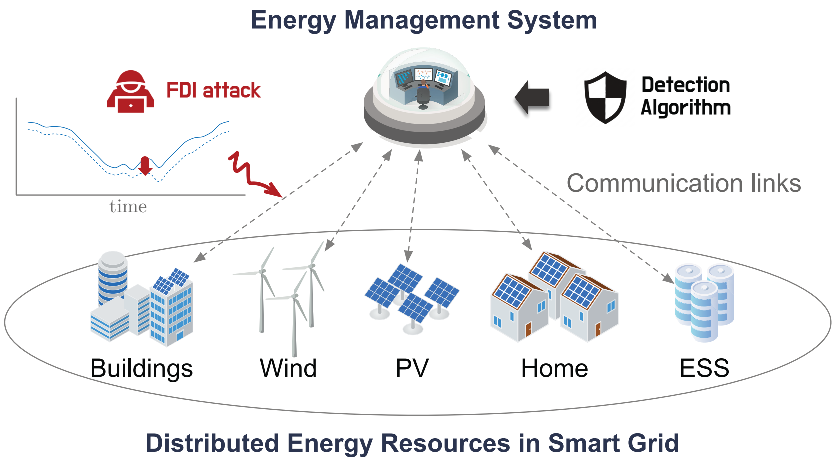
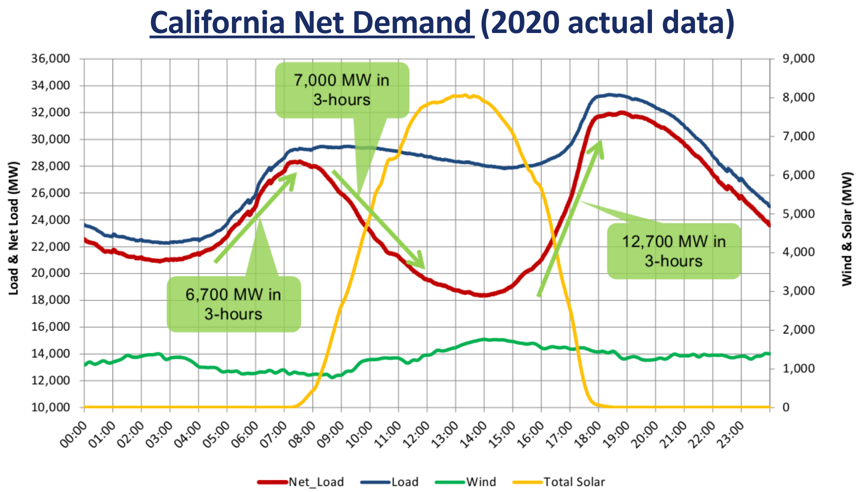
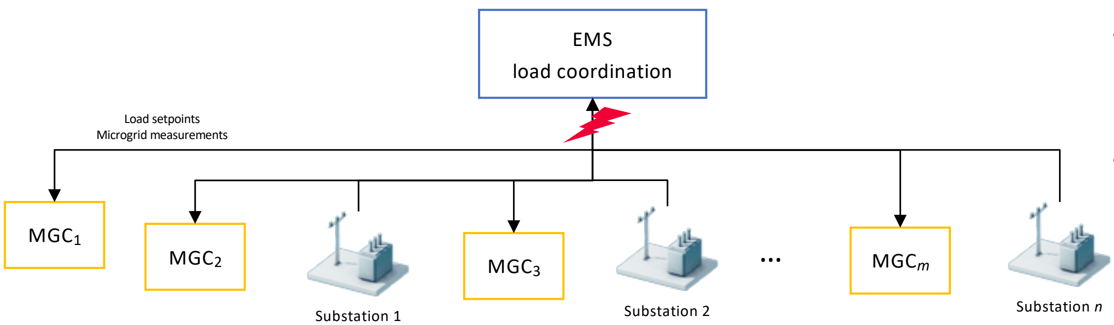


Fig. 1. Various DERs and connecting communication links to the EMS. Potential intraday FDI attacks targeting the communication links and detection algorithm for EMS are also illustrated.

Vulnerability of Power Grids with High Renewable Penetrations



- California ISO predicted (back in 2013) to have **13GW/3hr** maximum **net load ramp-up** for the year 2020 but it turned out to be around **17GW/3hr**
- Steep net load ramp-up → **increases the system vulnerability** to unexpected events such as cyber-attack



- **False data injection (FDI) attack** – during the evening net demand ramping up period.
- The attacker can manipulate the setpoints to the microgrid controller (MGC) so that the **power output of the various DERs** within the microgrid deviate from the original setpoints.

Outline

I. Vulnerability Analysis: Intraday FDI Attack on DER Dispatch Signals
  - Intraday FDI Attack Model
    - Dispatch prediction model
    - FDI attack scenarios

II. Kernel SVR-based Detection
  - Kernel Support Vector Regression (Kernel-SVR)
  - Identification of Intraday FDI Attack w/ Kernel-SVR

Intraday FDI Attack Scenarios

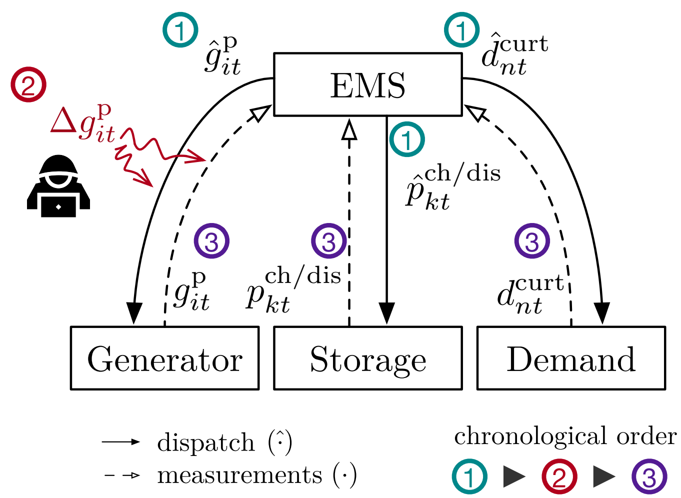


Fig. Dispatch & measurement between EMS & DER

Step 1. Learn about system topology & characteristics
→ Mid-/Long-term observations

Step 2. Predict demand / generations

Step 3. Develop falsification strategies

Name	Attack Model
Network information:	
topology $(\mathcal{N}, \mathcal{L})$	✓
line impedance (R_l, X_l)	✓
line thermal limit (F_l)	✓

Intraday FDI Attack Model Summary

Dispatch <u>Prediction</u> Model (SOCP)	Objective	- Minimize <u>the total operation cost</u>
	Decisions	- **Predictions** on {Generation dispatch, Demand response, Allocated reserve}
	Constraints	- Power flow equations - Generation / Line / Demand response limits - Energy storage operations - Given solar/demand prediction
Dispatch <u>Falsification</u> Model (QP)	Objective	- Minimize the <u>magnitude of falsification</u> signals and temporal changes
	Decisions	- DER dispatch **falsification signals**
	Constraints	- Individual falsification size limits proportional to the predicted signals - Supply-demand deviation exceeding the predicted reserve

Vulnerability analysis (Scenario A - generation setpoints)

- Attacker manipulates the generation dispatch signal from/to EMS
- From EMS's perspective, **the monitoring signal** is consistent with the **original signal** as the attack falsify the both directions.

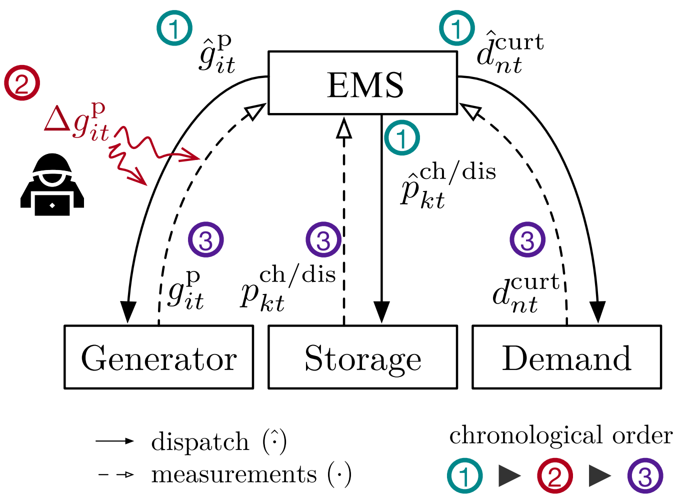


Fig. Summary of dispatch and measurement signal flow – falsified signals are marked with red color

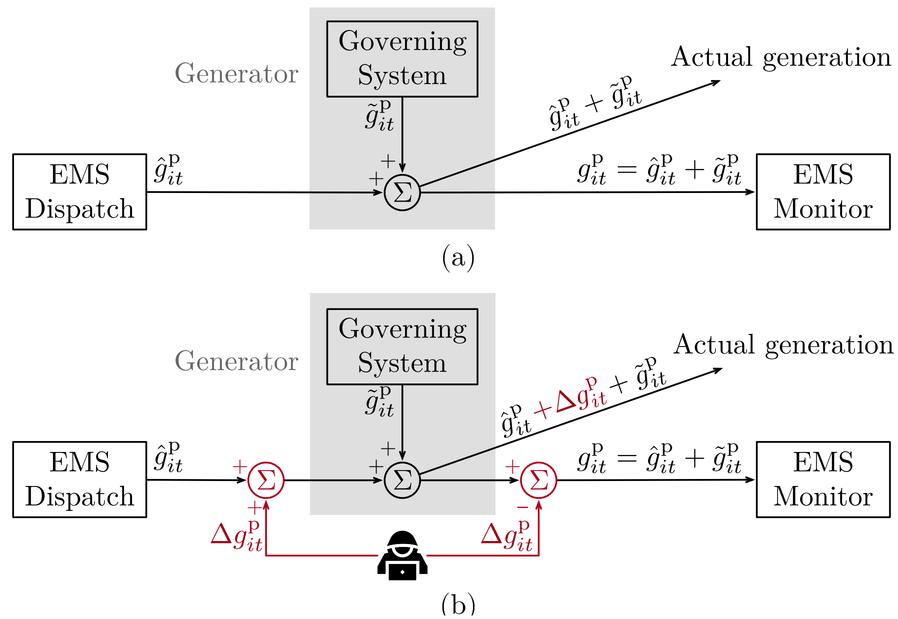


Fig. Dispatch signal flow of (a) normal operation case and (b) operation under FDI attack on generation dispatch.

Generation profile of the HCE test system

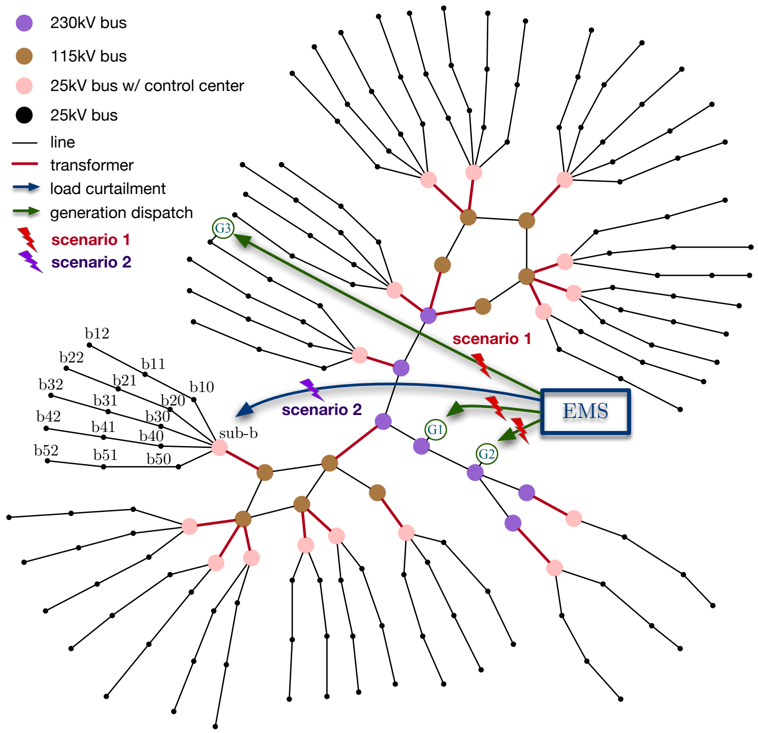


Fig. HCE HV network

Day 110 (April 21)

peak demand : 40 ~ 80MW

ramping capability : 13.5MW/3h

storage : (10MWh, 1MW)

** Solar generation has been scaled up 6 times (overall resulting solar penetration is around 15%, similar to the current practice in the state of California)

- **Test system & DERs**
  - Network, Demand, Generation profile provided from HCE
  - High solar penetration assumed (around 15%)
  - Three controllable generators
- **Software platforms**
  - **Optimization models** implemented using Julia/JuMP packages with Gurobi/Ipopt
  - **Kernel SVR model** implemented using Scikit-learn library
- **Detection model**
  - 6-hour of monitoring window, 2-hour of prediction window

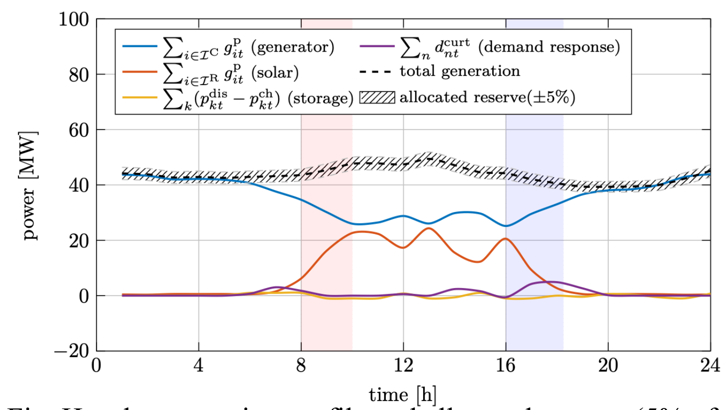


Fig. Hourly generation profile and allocated reserve (5% of the system load) of Day 110 under FDI attack on generation setpoints.

Vulnerability analysis (Scenario A - generation setpoints)

- Attacker manipulates the generation dispatch signal from/to EMS
- As a result, the reduced total generation (dashed lines) exceeding the system security margin (shaded area).

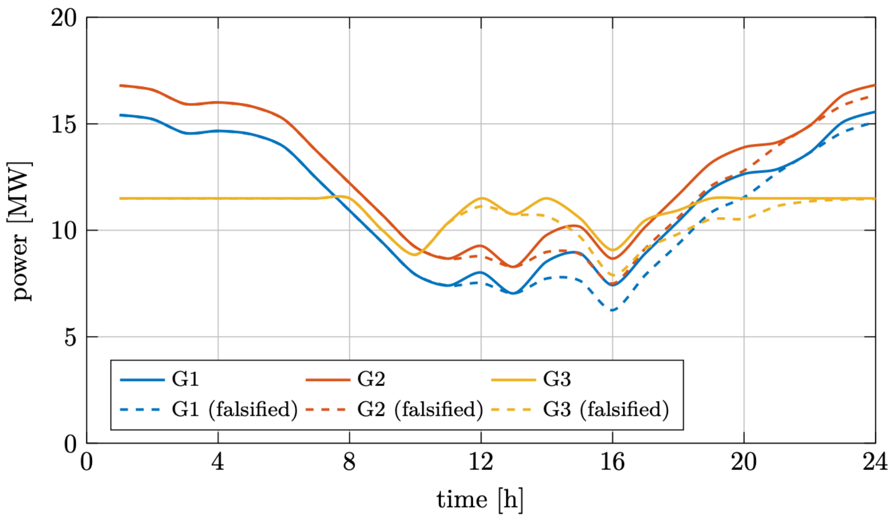
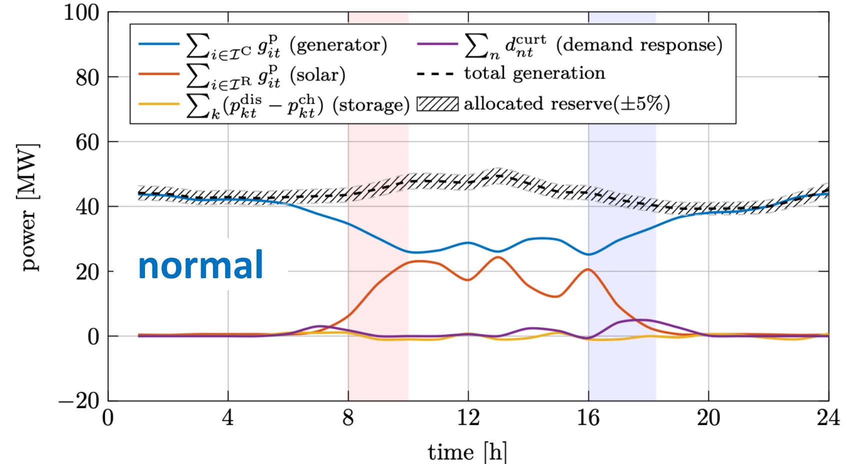
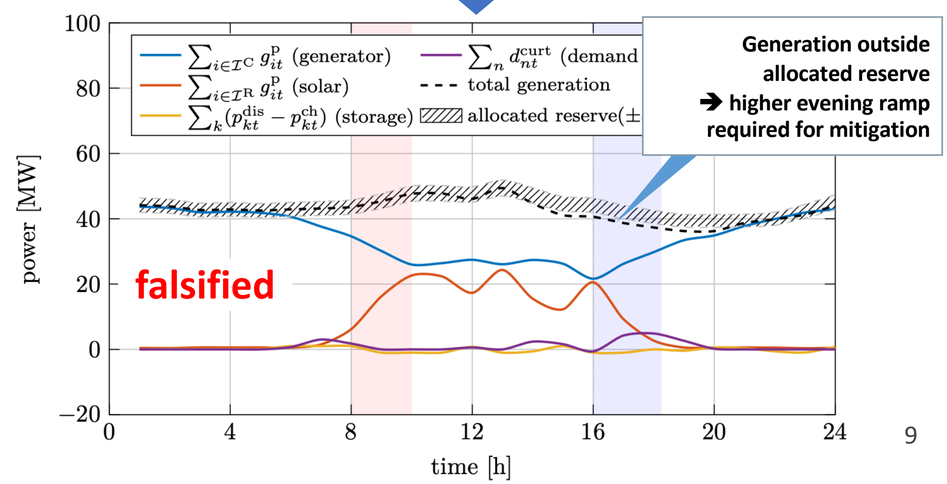


Fig. Generation setpoints: original dispatch (solid lines) and falsified signals (dashed lines)





Vulnerability analysis (Scenario B - load curtailment setpoints)

- Attacker manipulates the load curtailment dispatch from/to EMS
- From EMS's perspective, **<u>the monitoring signal</u>** <u>is consistent with the</u> **<u>original signal</u>** as the attack falsify the both directions.

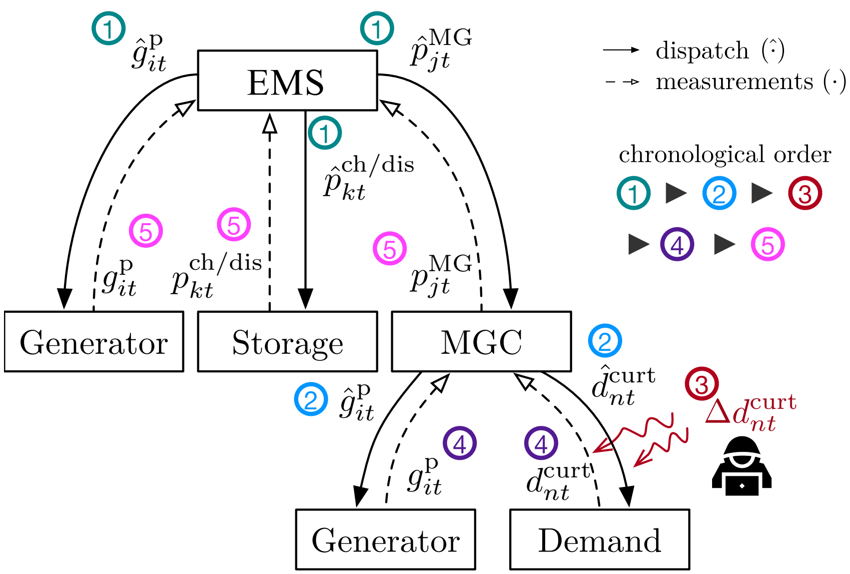


Fig. Summary of dispatch and measurement signal flow – falsified signals are marked with red color

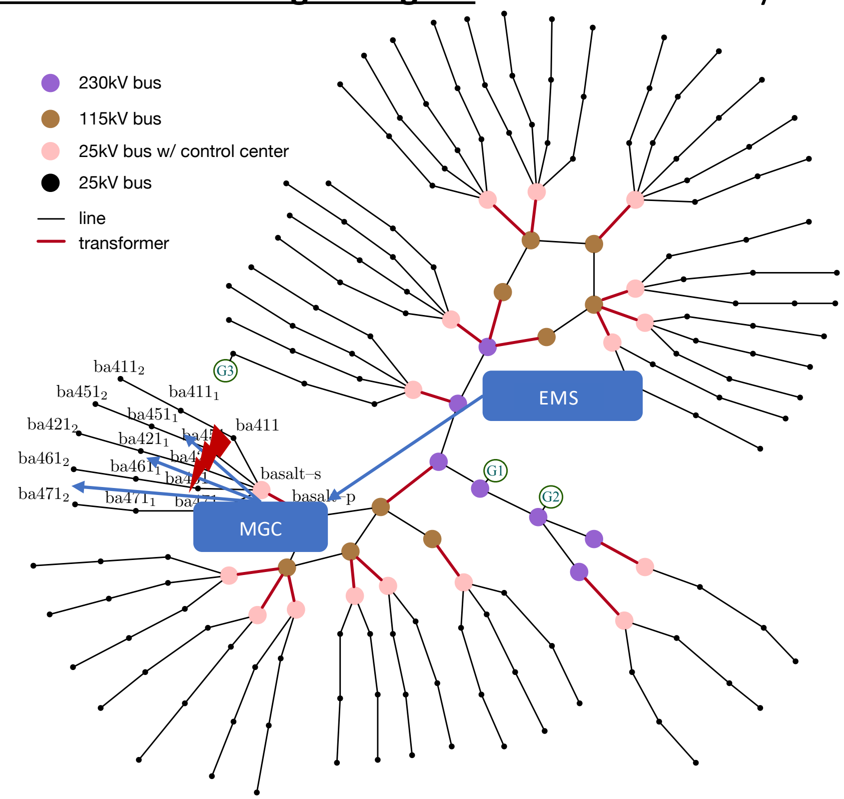


Fig. 4. HCE 187-bus test system.

Vulnerability analysis (Scenario B - load curtailment setpoints)

- Attacker manipulates the load curtailment dispatch from/to EMS
- As a result, the reduced total generation (dashed lines) exceeding the system security margin (shaded area).

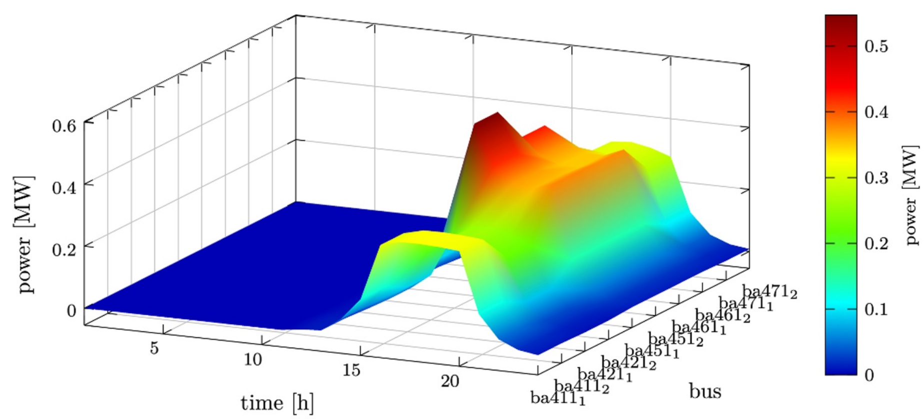
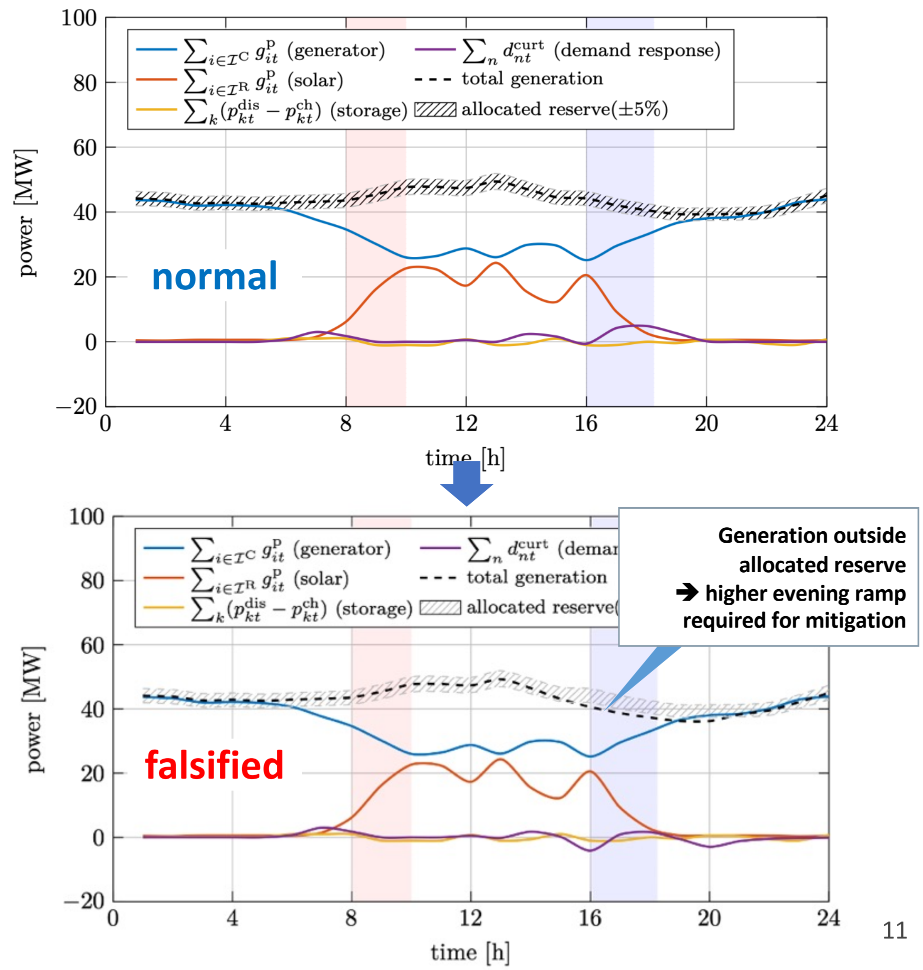


Fig. Falsified load curtailment signal in basalt-s area



Outline

I. Vulnerability Analysis: Intraday FDI Attack on DER Dispatch Signals
- Intraday FDI Attack Model
  - Dispatch prediction model
  - FDI attack scenarios

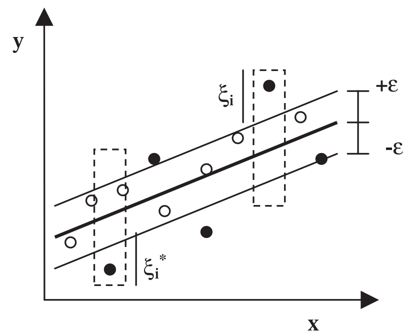
II. Kernel SVR-based Detection
- Kernel Support Vector Regression (Kernel-SVR)
- Identification of Intraday FDI Attack w/ Kernel-SVR

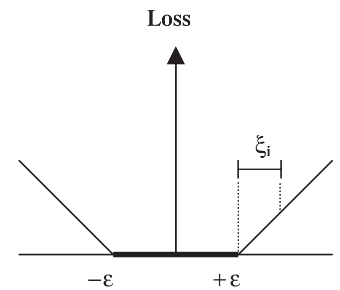
Identification of Intraday False Data Injection Attack

- Supervised Learning-based Detection: **Kernel Support Vector Regression (Kernel-SVR)**
  - Kernel-SVR for time series forecasting (e.g., wind speed prediction, load prediction)
  - Kernel-SVR for multi-temporal correlation in the system status & dispatch signals and identifying time-series FDI attacks

$$
\begin{aligned}
\min_{w,b,\xi_i^+,\xi_i^-} \quad & \|w\| + C\sum_{i=1}^{l}(\xi_i^+ + \xi_i^-) \\
\text{s.t.} \quad & (w^\top x^i + b) - y_i \le \epsilon + \xi_i^+, \quad \forall i \\
& y_i - (w^\top x^i + b) \le \epsilon + \xi_i^-, \quad \forall i \\
& \xi_i^+, \xi_i^- \ge 0, \quad \forall i
\end{aligned}
$$

User-defined parameters
C: Weight, ϵ: Insensitive zone





Identification of Intraday False Data Injection Attack

- Supervised Learning-based Detection: **Kernel Support Vector Regression (Kernel-SVR)**

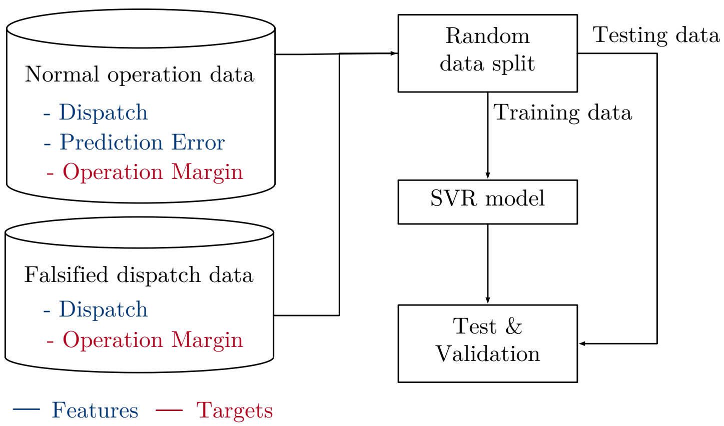


Fig. 1. Flow of data for training and testing the proposed SVR model.

TABLE I. Assumptions for the Attack and Detection Models

Name	Attack Model	Detection Model
Network information:		
topology $(\mathcal{N}, \mathcal{L})$	✓	✓
line impedance (R_l, X_l)	✓	✓
line thermal limit (F_l)	✓	✓
Dispatch signals:		
generation output (g_{it}^{p})	–	P & A
load curtailment (d_{nt}^{curt})	–	P & A
storage dispatch $(p_{kt}^{\mathrm{ch/dis}})$	–	P & A
reserve (r_{it})	–	P & A
Measurements:		
nodal demand (D_{nt}^{p})	P	P & A
nodal voltage (v_{nt}, θ_{nt})	–	✓

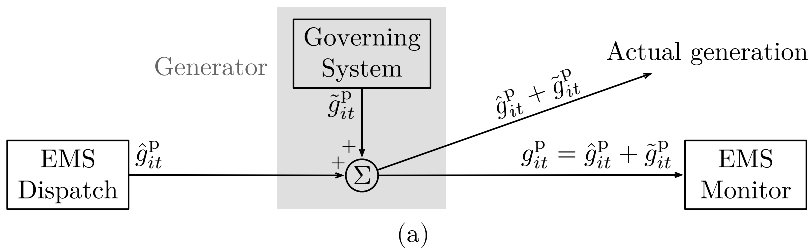
* P: prediction, A: actual, ✓: assumed to be known, –: unknown

Identification of Intraday False Data Injection Attack

$$\min\{r_t^{\mathrm{up}} - \sum_i (x_{it} - \hat{x}_{it}),\ r_t^{\mathrm{dn}} - \sum_i (\hat{x}_{it} - x_{it})\}.$$

System-wide margin

- Supervised Learning-based Detection: **Kernel Support Vector Regression (Kernel-SVR)**



$$x^i = \begin{bmatrix} [\hat{g}_{it}^{\mathrm{p}}]_{\forall i \in \mathcal{I}, t \in \mathcal{T}^{\mathrm{M}}} \\ [g_{it}^{\mathrm{p}}]_{\forall i \in \mathcal{I}, t \in \mathcal{T}^{\mathrm{M}}} \\ [\hat{d}_{nt}^{\mathrm{curt}}]_{\forall n \in \mathcal{N}, t \in \mathcal{T}^{\mathrm{M}}} \\ [d_{nt}^{\mathrm{curt}}]_{\forall n \in \mathcal{N}, t \in \mathcal{T}^{\mathrm{M}}} \\ [\hat{p}_{kt}^{\mathrm{ch/dis}}]_{\forall k \in \mathcal{K}, t \in \mathcal{T}^{\mathrm{M}}} \\ [p_{kt}^{\mathrm{ch/dis}}]_{\forall k \in \mathcal{K}, t \in \mathcal{T}^{\mathrm{M}}} \\ [v_{nt}]_{\forall n \in \mathcal{N}^{\mathrm{s}}, t \in \mathcal{T}^{\mathrm{M}}} \\ [\theta_{nt}]_{\forall n \in \mathcal{N}^{\mathrm{s}}, t \in \mathcal{T}^{\mathrm{M}}} \end{bmatrix} \in \mathbb{R}^m,\ y_i \in \mathbb{R}$$

where $m = (2|\mathcal{I}| + 2|\mathcal{N}| + 2|\mathcal{K}| + 2|\mathcal{N}^{\mathrm{s}}|)|\mathcal{T}^{\mathrm{M}}|$

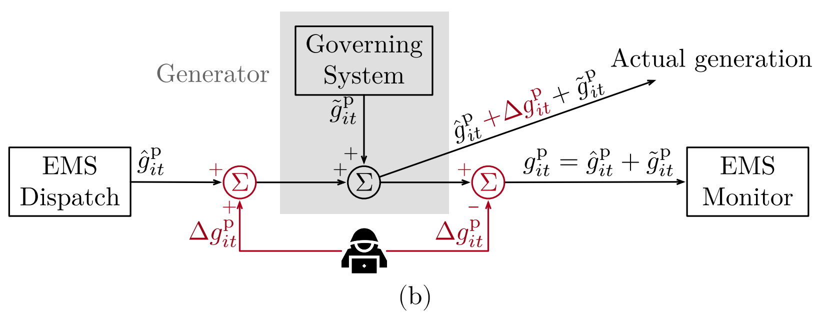


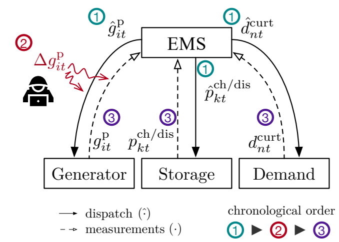
Fig. 2. Dispatch signal flow of (a) normal operation case and (b) operation under FDI attack on generation dispatch.

$$x^i = \begin{bmatrix} [\hat{g}_{it}^{\mathrm{p}}]_{\forall i \in \mathcal{I}, t \in \mathcal{T}^{\mathrm{M}}} \\ [g_{it}^{\mathrm{p}}]_{\forall i \in \mathcal{I}, t \in \mathcal{T}^{\mathrm{M}}} \\ [\hat{d}_{nt}^{\mathrm{curt}}]_{\forall n \in \mathcal{N}, t \in \mathcal{T}^{\mathrm{M}}} \\ [d_{nt}^{\mathrm{curt}}]_{\forall n \in \mathcal{N}, t \in \mathcal{T}^{\mathrm{M}}} \\ [\hat{p}_{kt}^{\mathrm{ch/dis}}]_{\forall k \in \mathcal{K}, t \in \mathcal{T}^{\mathrm{M}}} \\ [p_{kt}^{\mathrm{ch/dis}}]_{\forall k \in \mathcal{K}, t \in \mathcal{T}^{\mathrm{M}}} \\ [v_{nt}]_{\forall n \in \mathcal{N}^{\mathrm{s}}, t \in \mathcal{T}^{\mathrm{M}}} \\ [\theta_{nt}]_{\forall n \in \mathcal{N}^{\mathrm{s}}, t \in \mathcal{T}^{\mathrm{M}}} \end{bmatrix} \in \mathbb{R}^m,\ y_i \in \mathbb{R}$$

Different values under attack

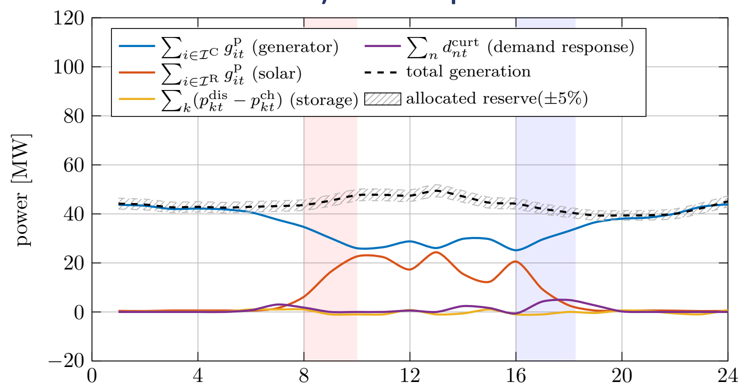
where $m = (2|\mathcal{I}| + 2|\mathcal{N}| + 2|\mathcal{K}| + 2|\mathcal{N}^{\mathrm{s}}|)|\mathcal{T}^{\mathrm{M}}|$

Identification of Intraday FDI Attack – Scenario A

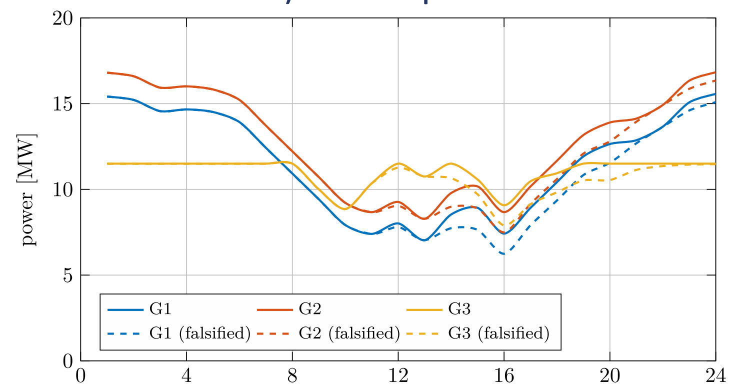


- Supervised Learning-based Detection: **Kernel Support Vector Regression (Kernel-SVR)**

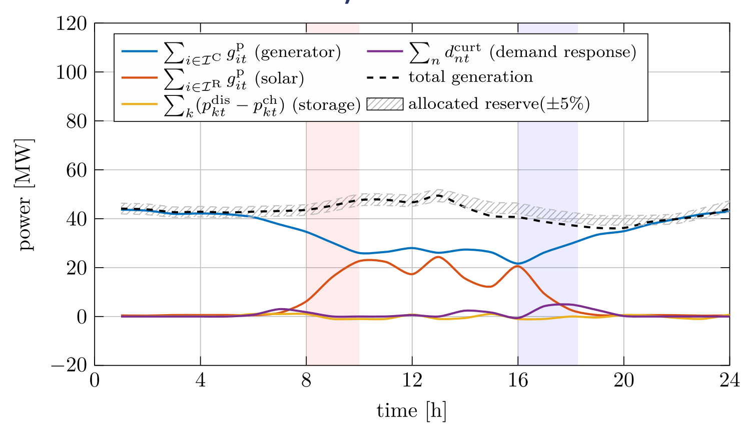
1) Normal dispatch



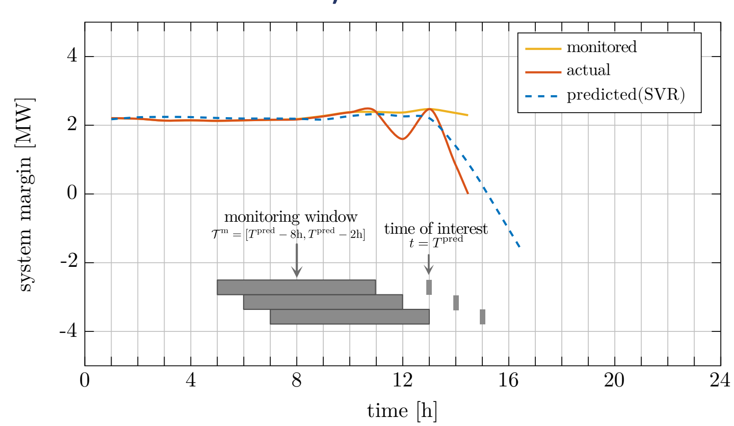
2) Falsified dispatch



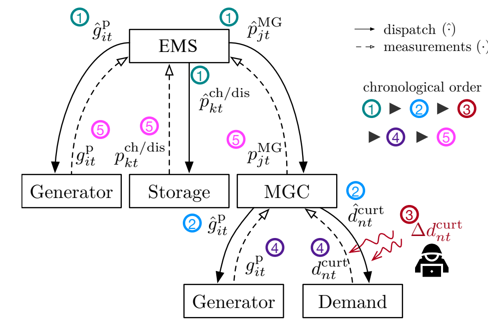
3) Falsification



4) Detection

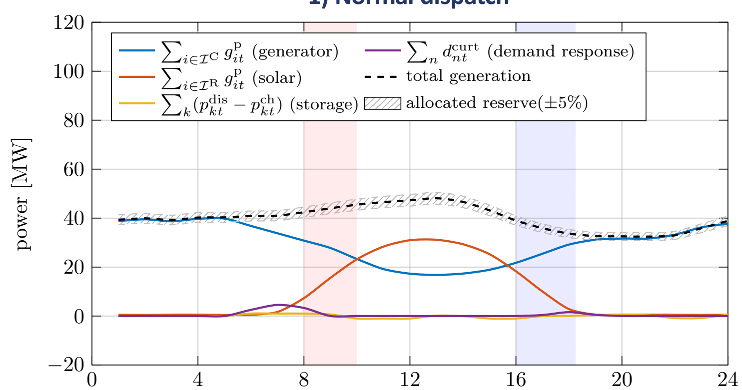


Identification of Intraday FDI Attack – Scenario B

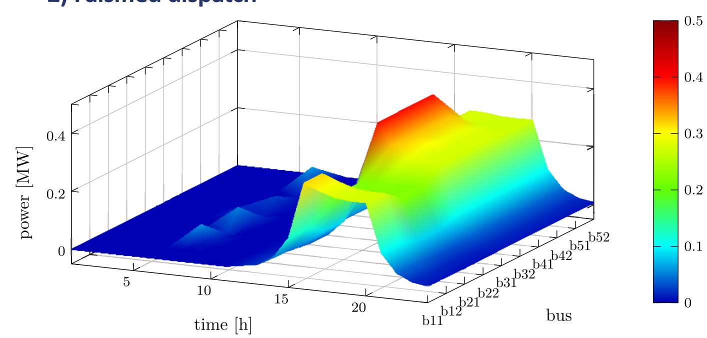


- Supervised Learning-based Detection: **Kernel Support Vector Regression (Kernel-SVR)**

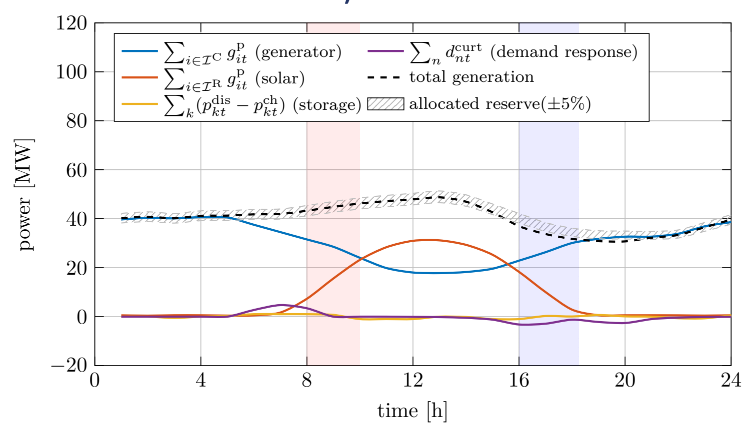
1) Normal dispatch



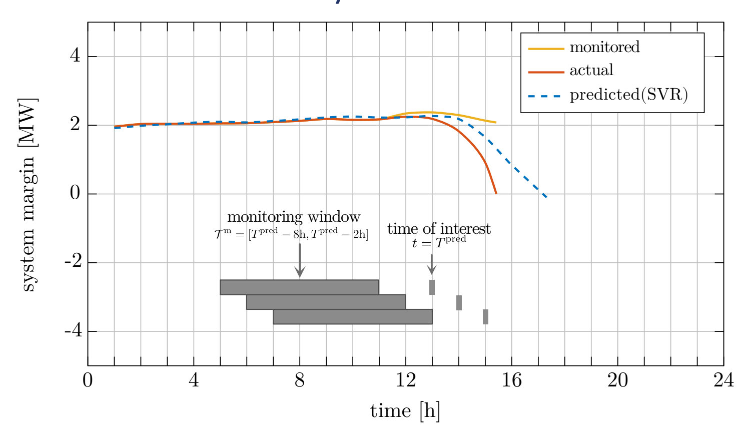
2) Falsified dispatch



3) Falsification



4) Detection



Conclusion

Summary

- We analyzed **the vulnerability of power grids with high PV penetration against a*<u>n intraday FDI attack</u>*** that falsifies DER dispatch and monitoring signals.
- Based upon the dispatch prediction and dispatch falsification models, we **illustrated how <u>gradual manipulation of DER outputs</u> can cause a power imbalance** which exceeds the system reliability margin.
- To enhance the power grid reliability against the attack scenario, we also **<u>proposed a detection model utilizing a kernel SVR</u>** which allows a power grid operator to predict the reduction in the system margin ahead of time.
- The numerical experiments demonstrate the attack scenarios and the performance of the detection model on **<u>the HCE test system</u>**, which is based on **real-world demand and generation profile data** provided from a power utility in Colorado.

Acknowledgement

- We would like to thank Bruno Leao and Ulrich Muenz at Siemens Technology for helpful discussions regarding the attack scenarios. We thank Chris Bilby at Holy Cross Energy for sharing relevant datasets.

Thank you!

jipkim@kentech.ac.kr

ArXiv paper available: https://arxiv.org/abs/2207.03667

* This material is based upon work supported in part by NSF grants CNS- 2148128 and EPCN-2144634, and by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy under the Solar Energy Technology Office Award Number DE-EE0008769. The views expressed herein do not necessarily represent the views of the U.S. Department of Energy or the United States