SAND2019-3590C

GOMACTech-2019

Power Grid Bad Data Injection Attack Modeling in PRESTIGE

27 March 2019

Yu-Cheng Chen
Georgia Institute of Technology
Brandon Eames
Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

- Motivation
- PLADD
- PRESTIGE
- Power Grid Scenario
- PRESTIGE Tool
  - Top Level Diagram
    - Development Process
    - Attack Graph
- Result
- Conclusion

Motivation

- Cyber-physical system such as the power grid, relies on microelectronics-based systems
- Attacks such as bad data injection can cause disruptions that transcend the cyber realm and affect the physical world
- We use the PRESTIGE Tool Chain developed by Sandia National Laboratories to model a bad data injection attack scenario in power grid infrastructure
  - What mitigations make sense?
  - How to optimize the use of resources to minimize the probability of a successful attack?

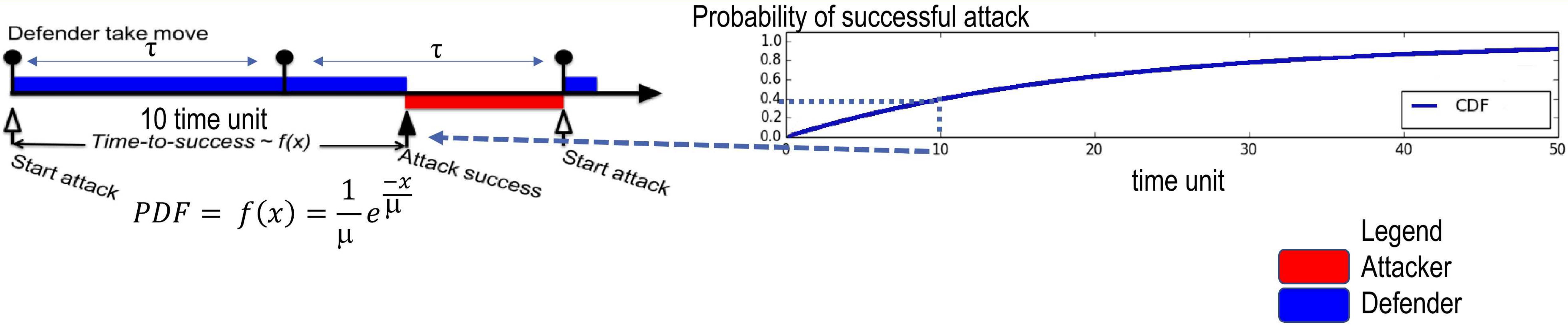
Why Game Theoretic Model?

- Prior attacks on power grid such as the Ukrainian power grid attack in 2015[10] are complex and spans multiple technical field
- Cyber analysts study the vulnerabilities of the power grid
- Domain experts have tools to analyze the reliability of the power grid
- Game theoretic approach focuses on the interaction of attacker and defender

Server containing IP address of devices in the power grid

Unclassified Unlimited Release

PLADD



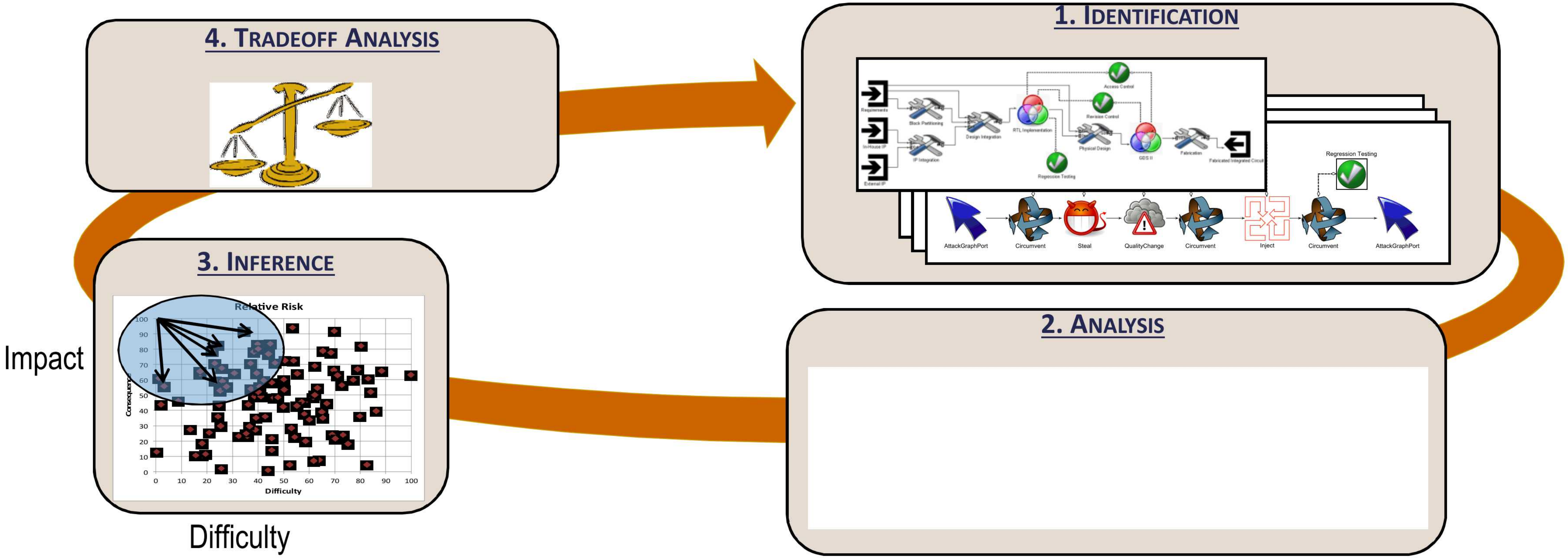
$$PDF = f(x) = \frac{1}{\mu} e^{\frac{-x}{\mu}}$$

Parameters of a single PLADD game	Description
μ	The attacker’s mean-time-to-success
α	The attacker’s fixed cost to initiate a new attack
β	The attacker’s variable cost related to the duration of an attack
c	The defender’s fixed cost for initiating a take move
τ	The defender period or the time between defender moves

Server containing IP address of devices in the power grid

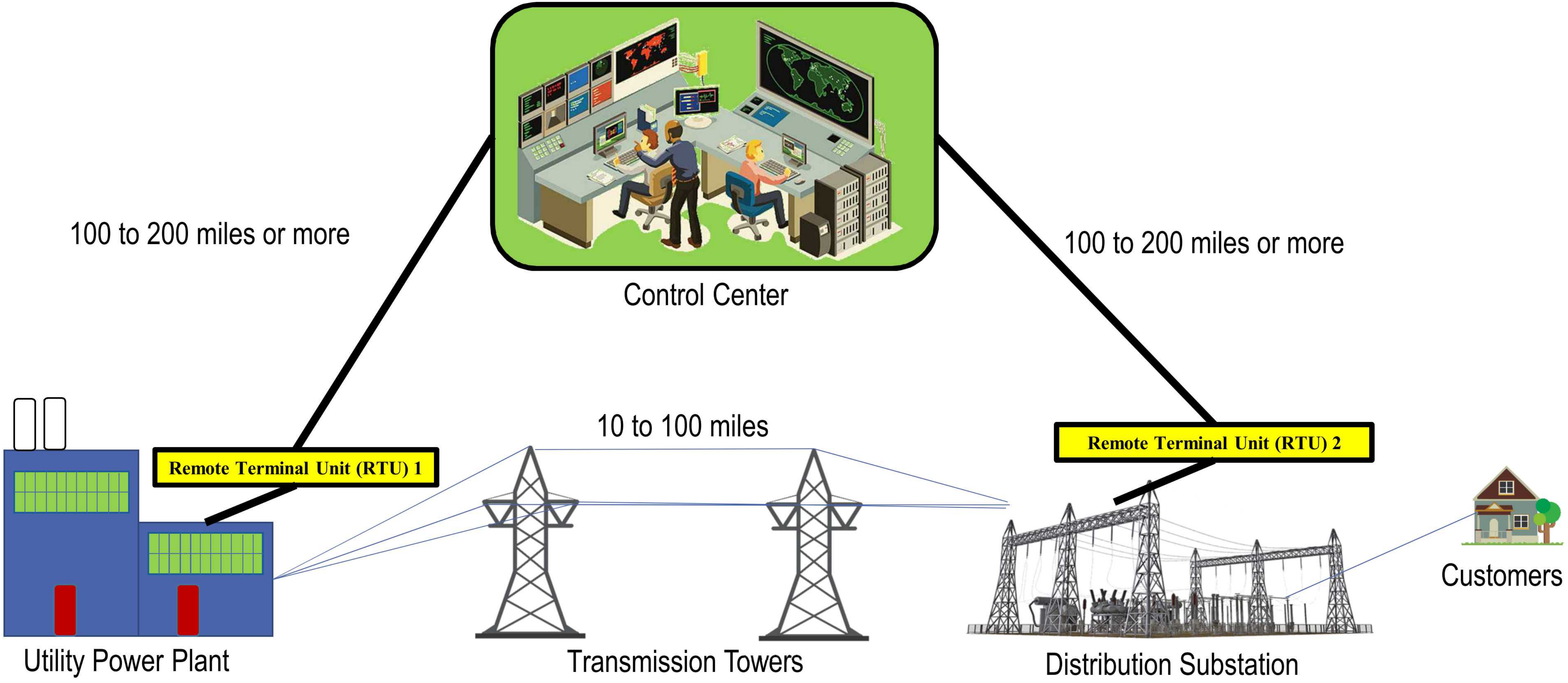
Unclassified Unlimited Release

PRESTIGE: PRactical Evaluation and Synthesis of Trust in Government systEms



Unclassified Unlimited Release

Power Grid Scenario: Geography



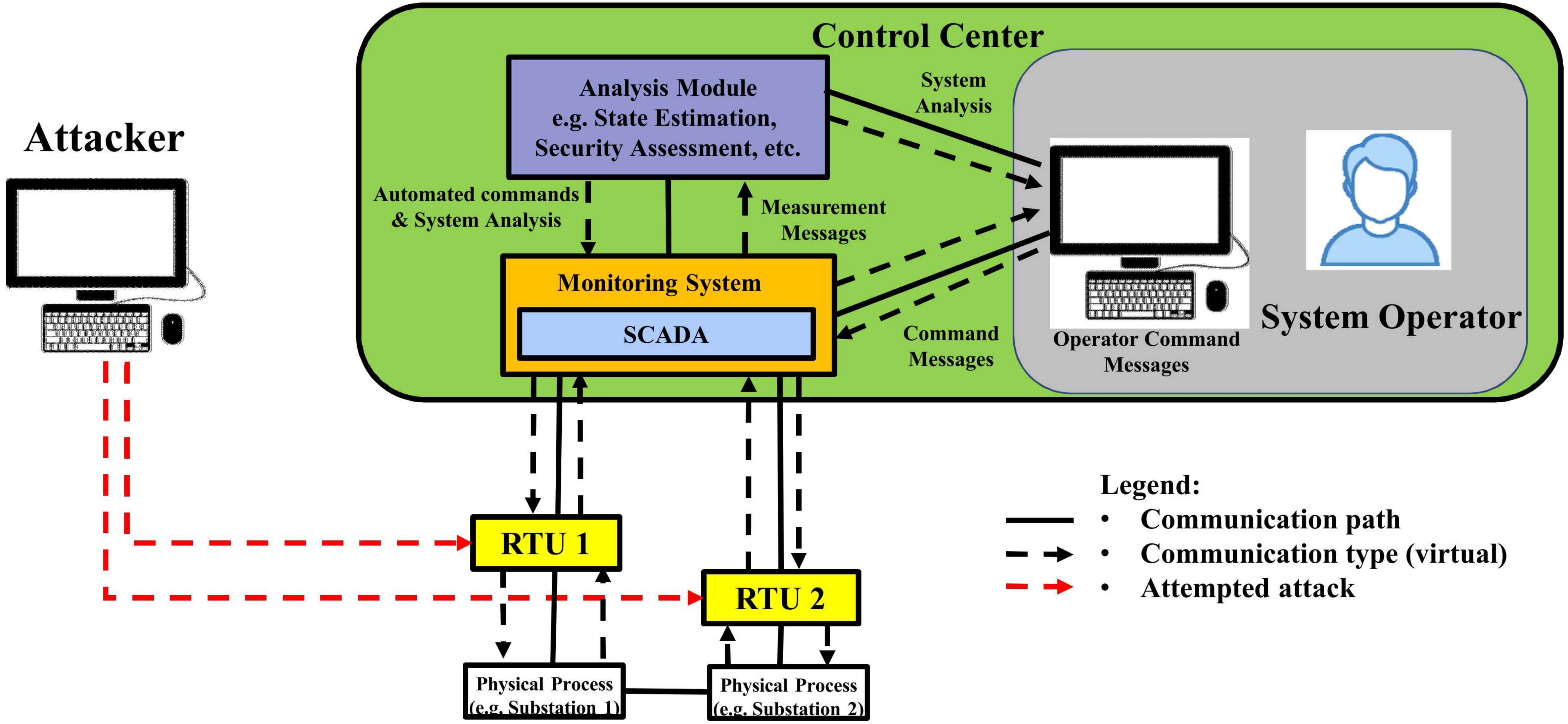
Unclassified Unlimited Release

Electric Substation (Satellite view)

Substation Room

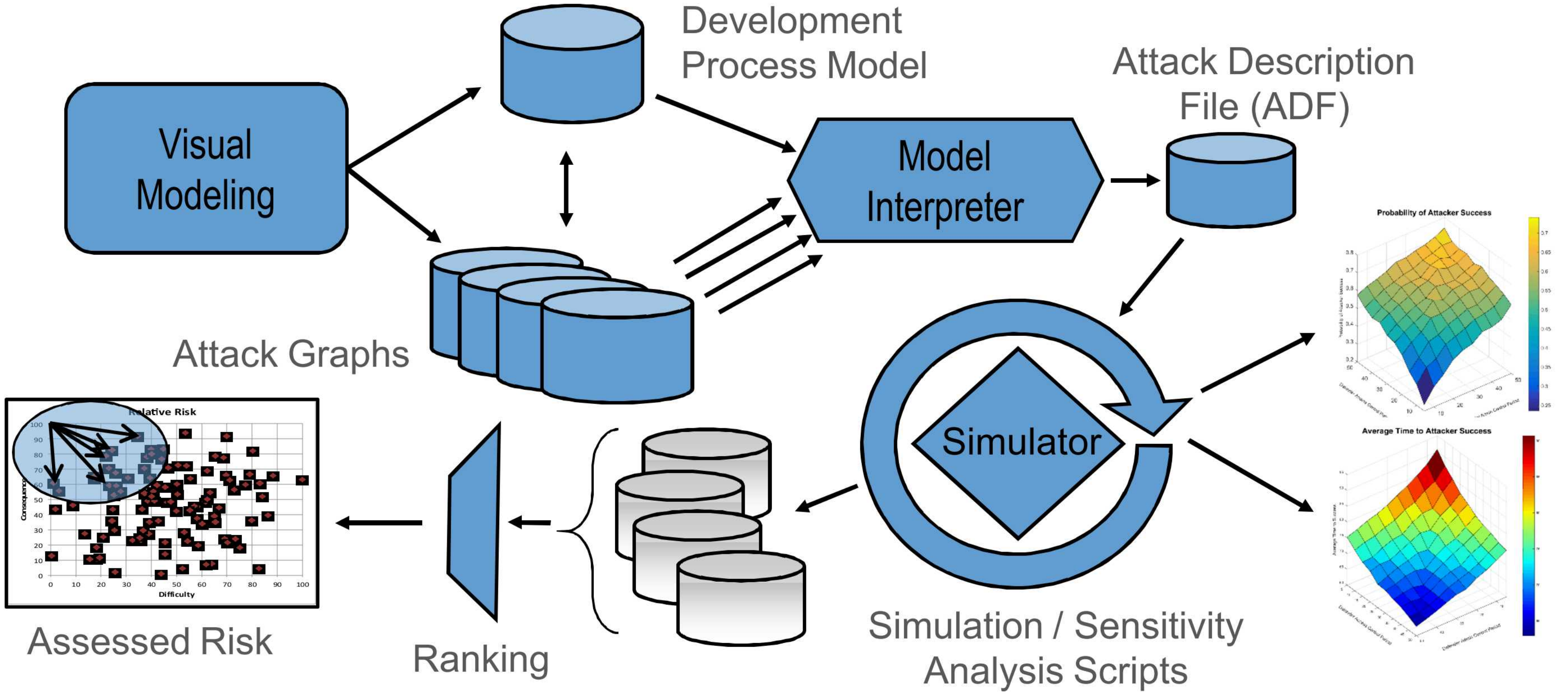
Unclassified Unlimited Release

Power Grid Scenario: Control



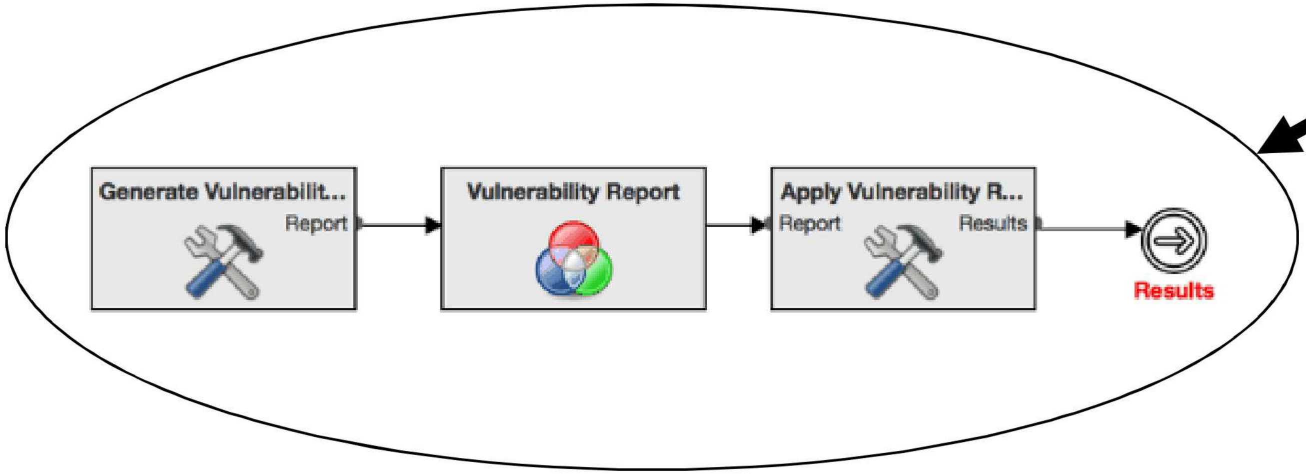
Unclassified Unlimited Release

PRESTIGE Tool Flow

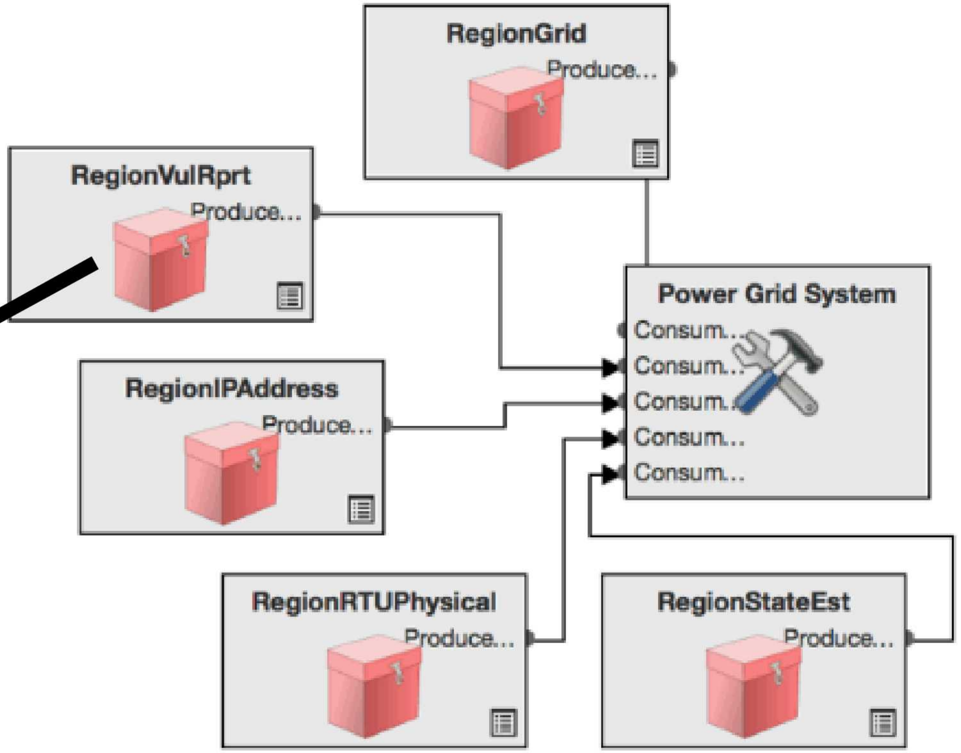


- End-to-end tool flow for modeling, evaluating assurance
  - Visual **modeling tool** for characterizing development processes, attacks
  - **Model interpreter** & **simulator** tools to quantitatively evaluate risk
  - **Ranking & Visualization tools** to navigate risk space and analysis results

PRESTIGE – Development Process



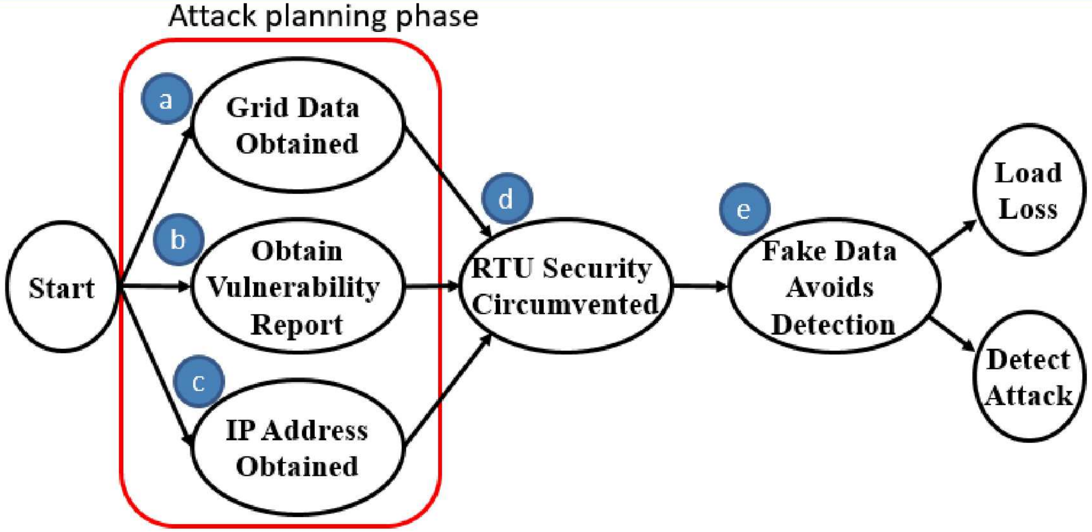
Contents of the RegionVulnRprt block, modeling electric vulnerability report generation and consumption



PRESTIGE Model of Power Grid System

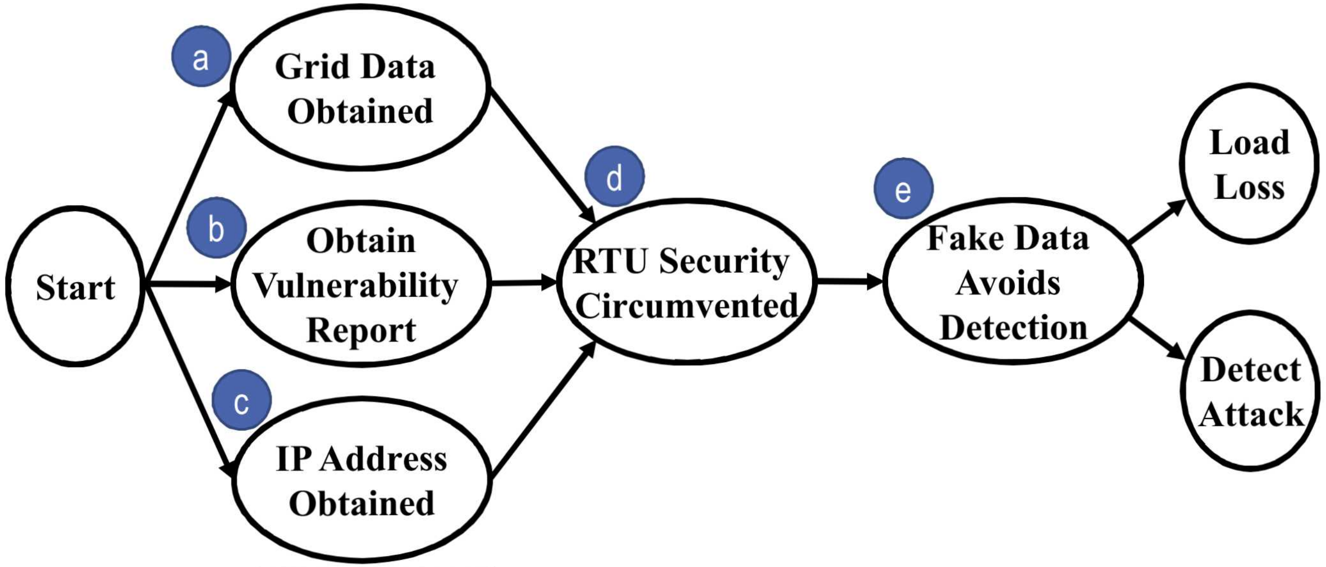
Unclassified Unlimited Release

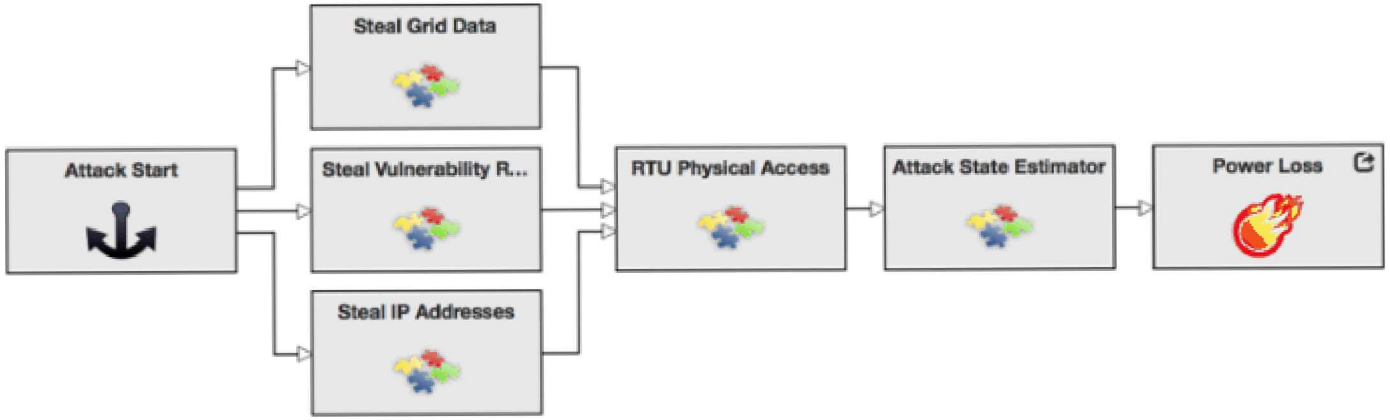
PRESTIGE – Attack Graph



PLADD Node	Attacker	Defender
‘a’	Steals power grid data	Changes the grid topology
‘b’	Steals the vulnerability report	Changes the grid state
‘c’	Steals IP address information	Changes IP address information
‘d’	Circumvents RTU security	Sends a utility engineer to check out the substation
‘e’	Maliciously change the system state	Reset the system state

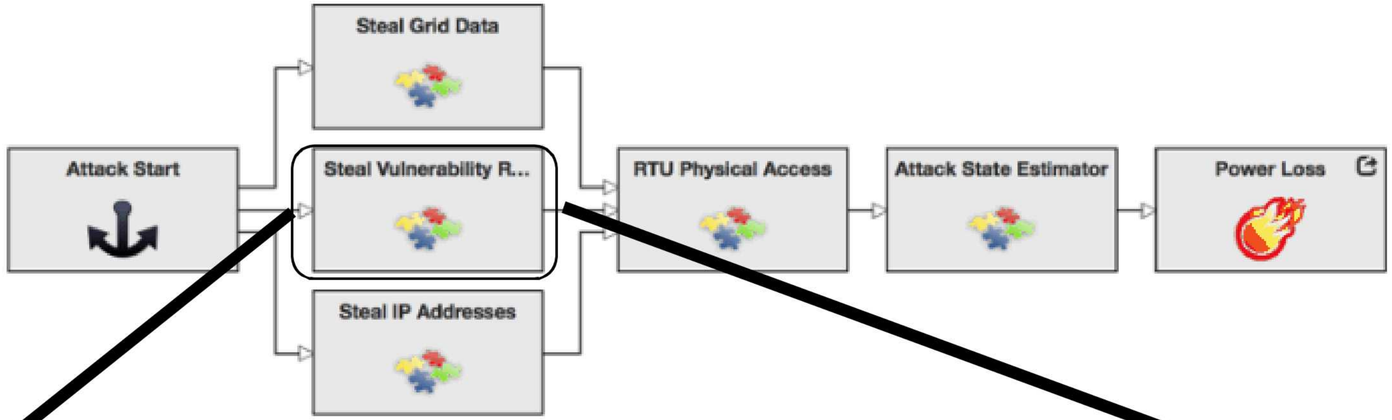
PRESTIGE – Attack Graph



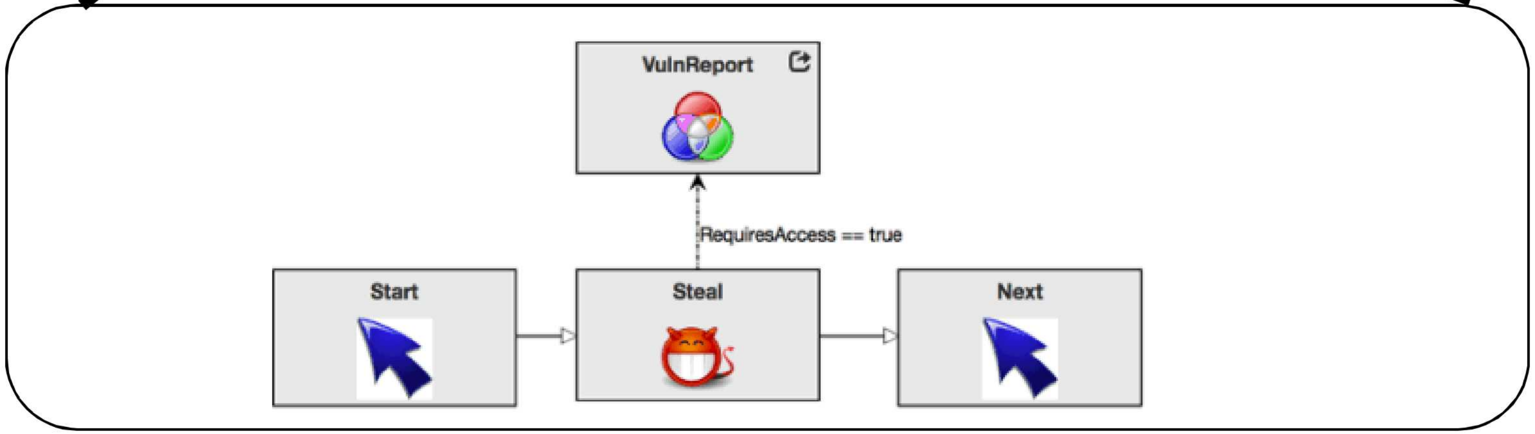


PRESTIGE model of the bad data injection attack

PRESTIGE – Attack Graph



PRESTIGE model of the bad data injection attack



PRESTIGE model of the Steal Electric Vulnerability Report phase of the bad data injection attack

Simulation Parameters

PLADD PARAMETERS IN PRESTIGE

Games	α	β	μ (month)	τ (month)	C
Grid Data	0.2	0.02	.25	60	200
Electric Vulnerability Report	0.4	0.04	.25	60	200
IP Address	0.4	0.04	.25	1	1
RTU Security	0.2	0.02	.00139	6	2
Fake Data Avoids Detection	0.1	0.01	.0171	.00984	2

- Given the current defender strategy, the simulation shows that the IP address information of the devices in the power grid is the least vulnerable point of the attack

Domain expert: Santiago Grijalva

- Georgia Institute of Technology Professor
- Georgia Power Distinguished Professor
- Senior Member of IEEE (Power and Energy, Systems and Control, and Computer Engineering Societies)
- Member of CIGRE USNC
- Atlanta Smart Energy Society, Council Member

Simulation Parameters

PLADD PARAMETERS IN PRESTIGE

Games	α	β	μ (month)	τ (month)	C
Grid Data	0.2	0.02	.25	60	200
Electric Vulnerability Report	0.4	0.04	.25	60	200
IP Address	0.4	0.04	.25	1	1
RTU Security	0.2	0.02	.00139	6	2
Fake Data Avoids Detection	0.1	0.01	.0171	.00984	2

Grid Data

- μ is .25 months as a savvy attacker is capable of stealing grid data passed over a network
- τ is 60 months (5 years) given the topology of the power grid only changes when new developments are created, thus changes are only made approximately every 5 years.
- C is 200 as the changing the topology of the grid data is very expensive

Domain expert: Santiago Grijalva

- Georgia Institute of Technology Professor
- Georgia Power Distinguished Professor
- Senior Member of IEEE (Power and Energy, Systems and Control, and Computer Engineering Societies)
- Member of CIGRE USNC
- Atlanta Smart Energy Society, Council Member

Simulation Parameters

PLADD PARAMETERS IN PRESTIGE

Games	α	β	μ (month)	τ (month)	C
Grid Data	0.2	0.02	.25	60	200
Electric Vulnerability Report	0.4	0.04	.25	60	200
IP Address	0.4	0.04	.25	1	1
RTU Security	0.2	0.02	.00139	6	2
Fake Data Avoids Detection	0.1	0.01	.0171	.00984	2

IP Address

- μ is .25 months as a savvy attacker is capable of stealing the IP address passed over a network
- τ is 1 month as the IP address can be changed more frequently
- C is 1 as the cost to change the IP address is comparably less than changing the topology of the power grid

Domain expert: Santiago Grijalva

- Georgia Institute of Technology Professor
- Georgia Power Distinguished Professor
- Senior Member of IEEE (Power and Energy, Systems and Control, and Computer Engineering Societies)
- Member of CIGRE USNC
- Atlanta Smart Energy Society, Council Member

Result (time)

AVERAGE DURATION OF ATTACK ON NODES IN ATTACK GRAPH

Attack Node	Time spent (month)
Grid Data	5.23
Electric Vulnerability Report	5.2
IP Address	610
RTU Security	0.09
Fake Data Avoids Detect	0.01

- After simulating the attack for 20 repetitions, the following are computed:
  - Attacker's attack success rate is computed by PRESTIGE to be 45%.
    - In those runs where the attacker was successful, the attacker achieved his goals after only 11.67 months.
  - Average time for the defender to complete his goals is computed to be 1105 months.
    - Defender execution time is dominated by the delay time associated with the "consumption" actors.

Result (cost)

- Average cost incurred by the attacker regardless of win or loss is 149,000.
- The average attacker cost when the attacker wins the game is 2830, and when the attacker loses is 269,000.

- Average overall cost is 234,000, and does not change when the defender wins or loses.
  - This independence between defender incurred cost and the state of the game is a representation of the property of PLADD games that the defender is not able to observe the state of the PLADD game.

Conclusion

- The tool computed an average attacker’s time to success to be 11.67 months
- The results can be used to recommend design changes for a power grid
- The tool can compute the tradeoff between increased security versus cost

References

- [1] A. S. Bretas, N. G. Bretas, B. Carvalho, E. Baeyens, and P. P. Khargonekar, "Smart grids cyber-physical security as a malicious data attack: An innovation approach," *Electr. Power Syst. Res.*, vol. 149, pp. 210–219, 2017.
- [2] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *Ccs*, vol. 14, no. 1, pp. 1–33, 2009.
- [3] T. Liu, Y. Gu, D. Wang, Y. Gui, and X. Guan, "A novel method to detect bad data injection attack in smart grid," *2013 Proc. IEEE INFOCOM*, pp. 3423–3428, 2013.
- [4] V. Chukwuka, Y.-C. Cheng, S. Grijalva and V. Mooney, "Bad Data Injection Attack Propagation in Cyber-Physical Power Delivery Systems," Power System Conference, Sep. 2018.
- [5] R. Kinney, P. Crucitti, R. Albert, and V. Latora, "Modeling cascading failures in the North American power grid," *Eur. Phys. J. B*, vol. 46, no. 1, pp. 101–107, 2005.
- [6] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, 2012.
- [7] S. Arianos, E. Bompard, A. Carbone, and F. Xue, "Power grids vulnerability: a complex network approach," Chaos: An Interdisciplinary Journal of Nonlinear Science, Vol. 19, Issue 1, March 2009.
- [8] S. Jones, A. Outkin, J. Gearhart, J. Hobbs, J. Siirola, C. Phillips, S. Verzi, D. Tauritz, S. Mulder, and A. Naugl, "Evaluating Moving Target Defense with PLADD," Sandia National Laboratories, 2015.
- [9] M. Galiardi, E. Vugrin, B. Eames, A. Outkin, G. Wyss, J. Hamlet, R. Helinski, B. Anthony, M. Napier, J. Eldridge, A. Bertels and M. Holmes, "On Modeling Detection for Quantitative Trust Analysis," GOMACTech 2018, Miami FL.
- [10] Robert M. Lee; Michael J. Assante; tim Conway (18 March 2016). Analysis of the Cyber Attack on the Ukrainian Power Grid. Defense Use Case (PDF). E-ISAC.

GOMACTech-2019

Backup Slides

27 March 2019

Yu-Cheng Chen
Georgia Institute of Technology
Brandon Eames
Sandia National Laboratories

Simulation Parameters

PLADD PARAMETERS IN PRESTIGE

Games	α	β	μ (month)	τ (month)	C
Grid Data	0.2	0.02	.25	60	200
Electric Vulnerability Report	0.4	0.04	.25	60	200
IP Address	0.4	0.04	.25	1	1
RTU Security	0.2	0.02	.00139	6	2
Fake Data Avoids Detection	0.1	0.01	.0171	.00984	2

Fake Data Avoids Detection

- μ – Estimated by simulating how many **undetected fake data** are needed to pass successful-attack-criteria
- τ - Estimated by simulating how many **detected fake data** are needed to cause defender to take action

Domain expert: Santiago Grijalva

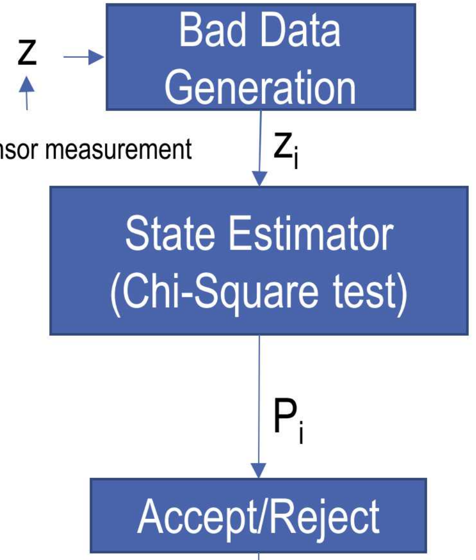
- Georgia Institute of Technology Professor
- Georgia Power Distinguished Professor
- Senior Member of IEEE (Power and Energy, Systems and Control, and Computer Engineering Societies)
- Member of CIGRE USNC
- Atlanta Smart Energy Society, Council Member

Simulation for Attacker's average attack complete time (μ)

$$Z=\begin{bmatrix}V_1\\V_2\\P_{12}\\Q_{21}\\P_2\end{bmatrix}=\begin{bmatrix}144.7\ kV\\119.2\ kV\\463.1\ MV\\-105.0\ Mvar\\404.5\ MW\end{bmatrix}=\begin{bmatrix}1.0485\\0.8623\\4.631\\-1.05\\-4.045\end{bmatrix}pu$$

Base 100 MVA
138 kV

Sign change due to conversion for load



$$Z_i=\begin{bmatrix}V_1\\V_2\\P_{12}\\Q_{21}\\P_2\end{bmatrix}=\begin{bmatrix}1.0485\\0.8623\\4.631*scale\\-1.05\\-4.045*scale\end{bmatrix}pu$$

$$0.1 < scale < 1.9$$

Bad data is generated by altering P_{12} and P_2 by a scale factor

[0, 0, 1, 1, 1, 0, 0, 0, 0, 0, …, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 0,…, 0, 1]

Data frame with 10 data in it

740th index

Attacker's time to success is 740 minutes because the 740th index is first time when the past 10 data contains at least 7 accepted bad data.

Accept/Reject Z_i based on whether P_i is too high

Result vector is indexed from 1 to 1440 which represent one bad data per minute for duration of one day

Simulation for Defender's time-between-action (τ)

$$Z = \begin{bmatrix} V_1 \\ V_2 \\ P_{12} \\ Q_{21} \\ P_2 \end{bmatrix} = \begin{bmatrix} 144.7\ kV \\ 119.2\ kV \\ 463.1\ MV \\ -105.0\ Mvar \\ 404.5\ MW \end{bmatrix} = \begin{bmatrix} 1.0485 \\ 0.8623 \\ 4.631 \\ -1.05 \\ -4.045 \end{bmatrix} pu$$

Base 100 MVA
138 kV

Sign change due to conversion for load

z → Bad Data Generation

Known good sensor measurement

z_i → State Estimator (Chi-Square test)

P_i → Accept/Reject

$$Z_i = \begin{bmatrix} V_1 \\ V_2 \\ P_{12} \\ Q_{21} \\ P_2 \end{bmatrix} = \begin{bmatrix} 1.0485 \\ 0.8623 \\ 4.631 * scale \\ -1.05 \\ -4.045 * scale \end{bmatrix} pu$$

$$0.1 < scale < 1.9$$

Bad data is generated by altering P_{12} and P_2 by a scale factor

[0, 1, 0, 1, 0, 1, 0, 1, 0, 1, …, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0,…, 0, 1]

Accept/Reject Z_i based on whether P_i is too high

Data frame with 10 data in it

425th index

Defender's time between each take move is 425 minutes, because 425th index is the first time when the past 10 data contains at least 6 rejected bad data.

Result vector is indexed from 1 to 1440 which represent one bad data per minute for duration of one day

Unclassified Unlimited Release