



Making Sound Design Decisions Using Quantitative Security Metrics

Bill Sanders

The Problem: Assessing Security and Resilience

- **Systems operate in adversarial environments**
 - Adversaries seek to degrade system operation by affecting the confidentiality, integrity, and/or availability of the system information and services
 - “Resilient” systems aim to meet their ongoing operational objectives despite attack attempts by adversaries
- **System security is not absolute**
 - No real system is perfectly secure
 - Some systems are more secure than others
 - *But which ones are more secure?*
 - *And how much more secure are they?*

Practical Applications of Security Metrics

Organizational-level Metrics

Questions the CIO cannot answer:

- How much risk am I carrying?
- Am I better off now than I was this time last year?
- Am I spending the right amount of money on the right things?
- How do I compare to my peers?
- What risk transfer options do I have?

(From CRA, Four Grand Challenges in Trustworthy Computing, 2003)

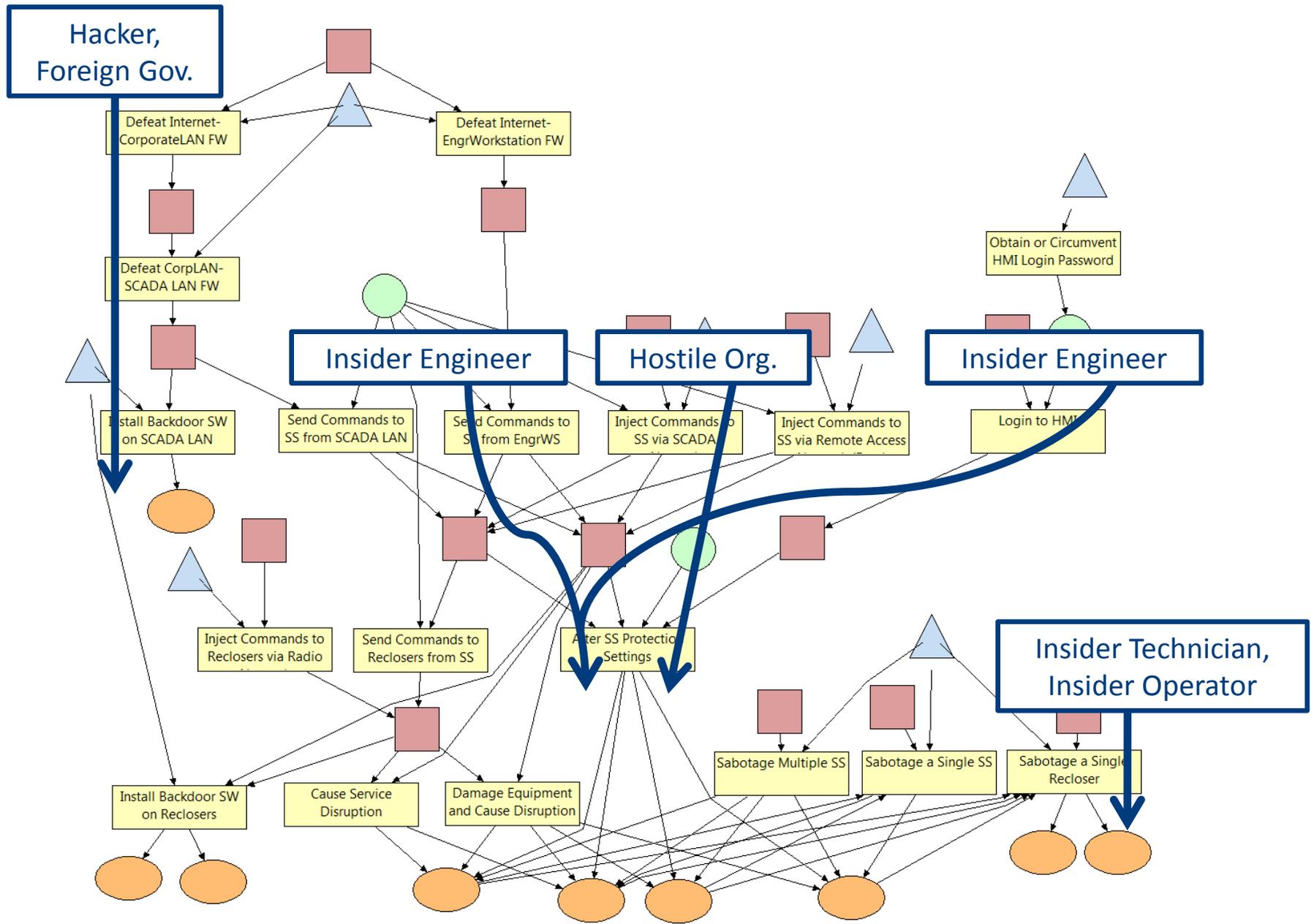
A Question neither can answer:

- How do the technical metrics impact the organizational-level security metrics?

Technical Metrics

Questions the design engineer cannot answer:

- Is design A or B more secure (confidentially, integrity, availability, privacy)?
- Have I made the appropriate design trade off between timeliness, security, and cost?
- How will the system, as implemented, respond to a specific attack scenario?
- What is the most critical part of the system to test, from a security point of view?



Related Work Motivating ADVISE

- Model-based security analysis
 - Attack Trees
 - Attack Graphs and Privilege Graphs
- Adversary-based security analysis
 - MORDA (Mission-Oriented Risk and Design Analysis)
 - NRAT (Network Risk Assessment Tool)

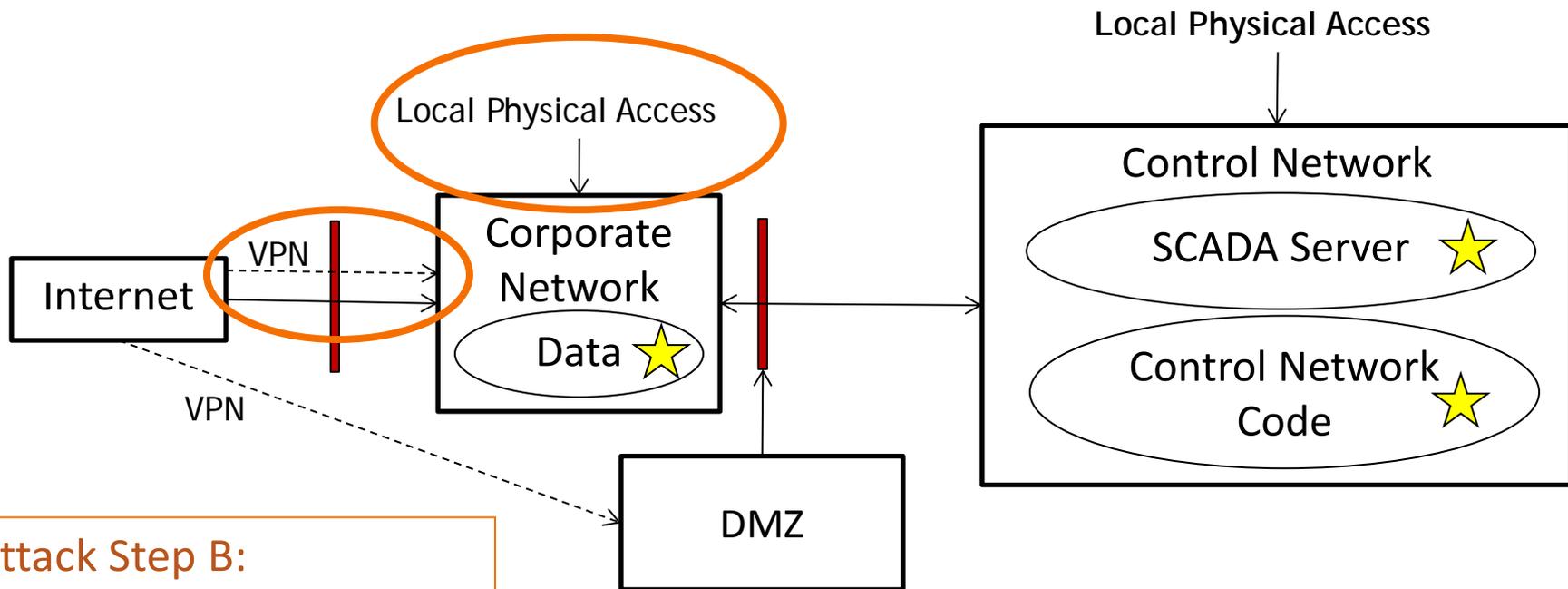
ADVISE integrates the benefits of both model-based and adversary-based security analysis

ADversary Vlew Security Evaluation (ADVISE) approach

- **Adversary-driven analysis**
 - Considers characteristics and capabilities of adversaries
- **State-based analysis**
 - Considers multi-step attacks
- **Quantitative metrics**
 - Enables trade-off comparisons among alternatives
- **Mission-relevant metrics**
 - Measures the aspects of security important to owners/operators of the system

Example: SCADA System Attack

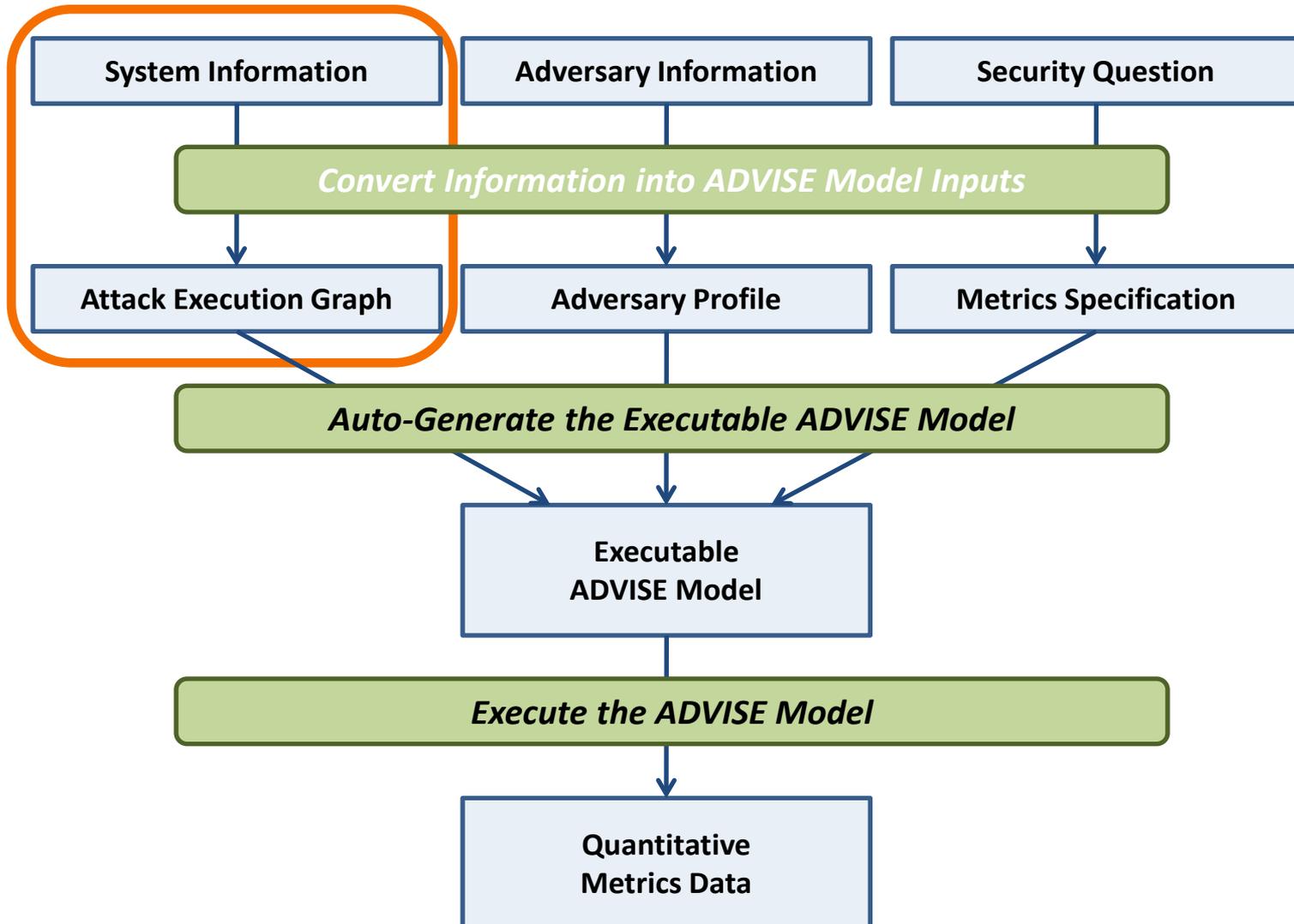
Attack Step A:
Gain Corporate Network Access
Through Local Physical Access



Attack Step B:
Gain Corporate Network
Access Through VPN

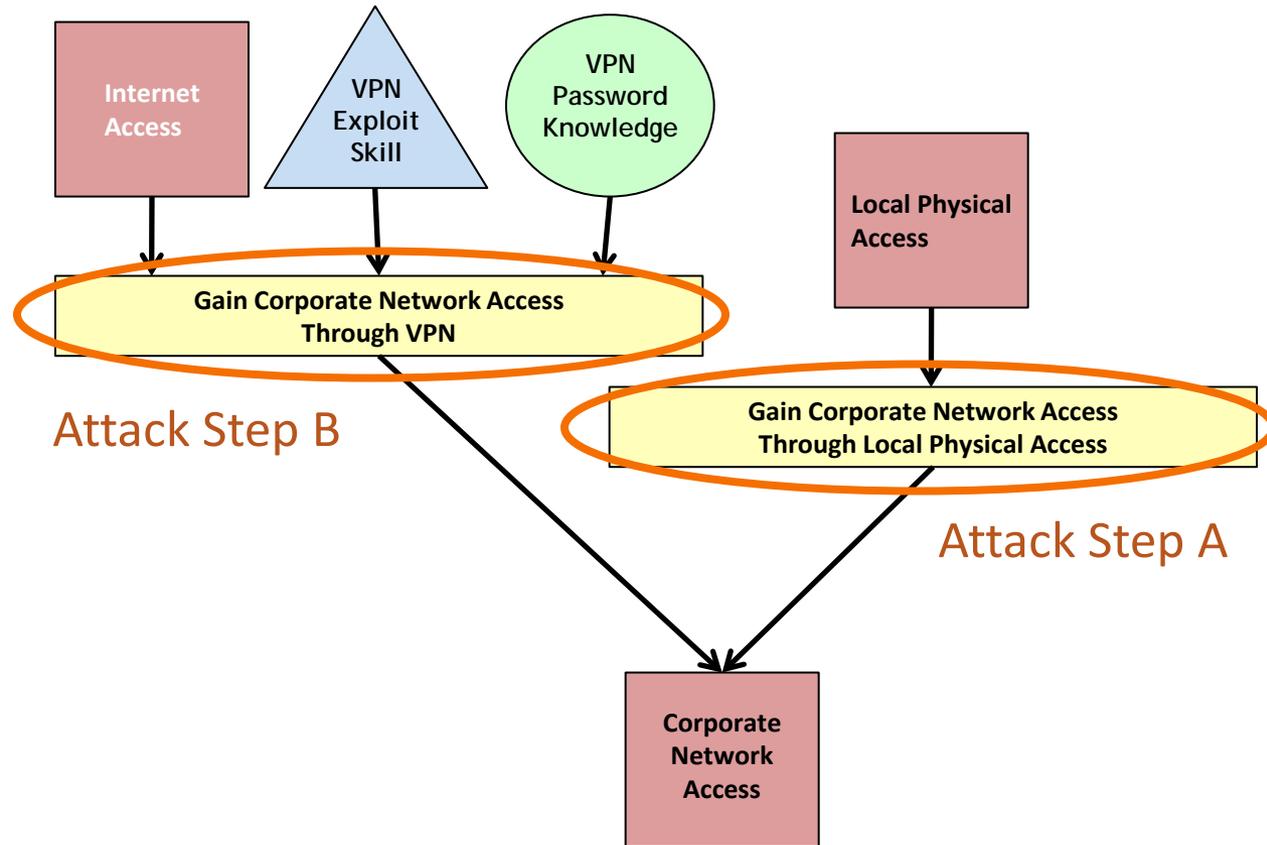
★ = Attack Target

ADVISE Method Overview



Representing Attacks Against the System

An “attack execution graph” describes potential attack vectors against the system from an attacker point of view. Attempting an attack step requires certain skills, access, and knowledge about the system. The outcome of an attack can affect the adversary’s access and knowledge about the system.



ADVISE System Information: Attack Execution Graph

An attack execution graph is defined by

$\langle A, R, K, S, G \rangle$,

where

A is the set of **attack steps**,
e.g., “Access the network using the VPN,”

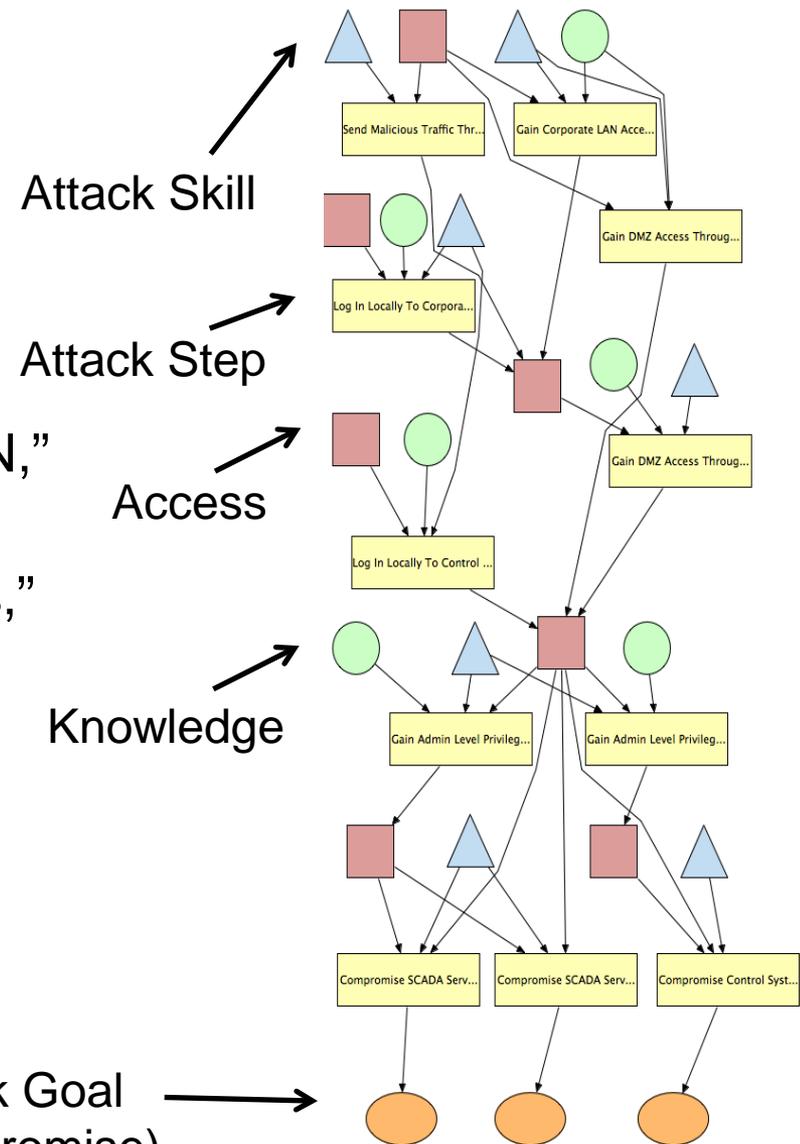
R is the set of **access domains**,
e.g., “Internet access,” “Network access,”

K is the set of **knowledge items**,
e.g., “VPN username and password”

S is the set of **adversary attack skills**,
e.g., “VPN exploit skill,” and

G is the set of **adversary attack goals**,
e.g., “View contents of network.”

Attack Goal
(System Compromise)



Attack Step Definition

An attack step a_i is a tuple:
 $a_i = \langle B_i, T_i, C_i, O_i, Pr_i, D_i, E_i \rangle$

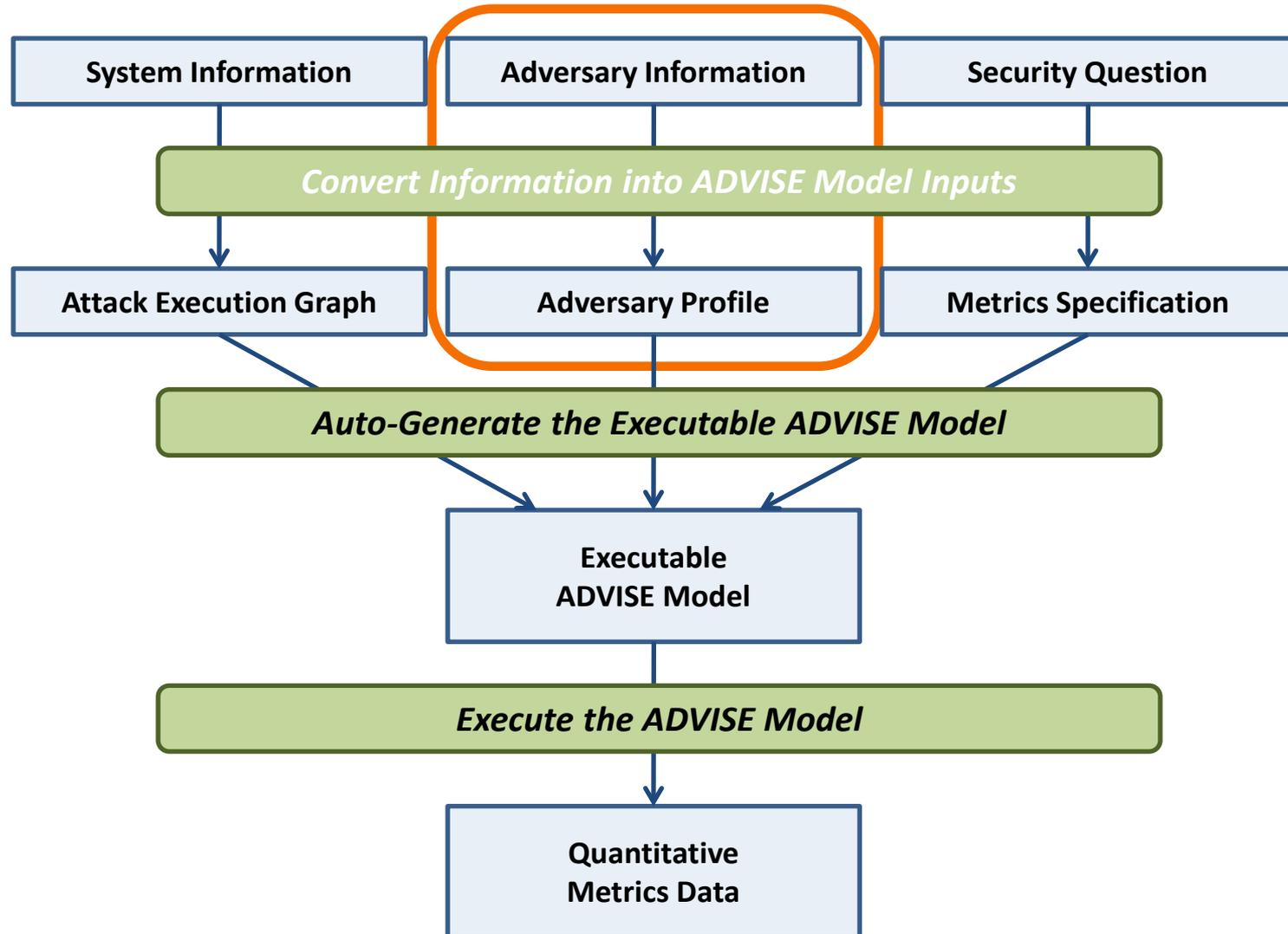
Note: X is the set of all states in the model.

- $B_i: X \rightarrow \{True, False\}$ is a **Boolean precondition**,
e.g., (Internet Access) AND ((VPN account info) OR (VPN exploit skill)).
- $T_i: X \times R^+ \rightarrow [0, 1]$ is the **distribution of the time to attempt the attack step**,
e.g., normally distributed with mean 5 hours and variance 1 hour.
- $C_i: X \rightarrow R^{\geq 0}$ is the **cost of attempting the attack step**, e.g., \$1000.
- O_i is a finite set of **outcomes**, e.g., {Success, Failure}.
- $Pr_i: X \times O_i \rightarrow [0, 1]$ is the **probability of outcome $o \in O_i$ occurring**,
e.g., if (VPN exploit skill > 0.8) {0.9, 0.1} else {0.5, 0.5}.
- $D_i: X \times O_i \rightarrow [0, 1]$ is the **probability of the attack being detected when outcome $o \in O_i$ occurs**, e.g., {0.01, 0.2}.
- $E_i: X \times O_i \rightarrow X$ is the **next-state that results when outcome $o \in O_i$ occurs**,
e.g., {gain Network Access, no effect}.

The “Do-Nothing” Attack Step

- Contained in every attack execution graph
- Represents the option of an adversary to refrain from attempting any active attack
 - The precondition $B_{\text{DoNothing}}$ is always true.
- For most attack execution graphs,
 - the cost $C_{\text{DoNothing}}$ is zero,
 - the detection probability $D_{\text{DoNothing}}$ is zero, and
 - the next-state is the same as the current state.
- The existence of the “do-nothing” attack step means that, regardless of the model state, there is always at least one attack step in the attack execution graph whose precondition is satisfied

ADVISE Method Overview



ADVISE Adversary Information: Adversary Profile

The adversary profile is defined by the tuple
 $\langle s_0, L, V, w_C, w_P, w_D, U_C, U_P, U_D, N \rangle$,

where

$s_0 \in X$ is the **initial model state**, e.g., has Internet Access & VPN password,

L is the **attack skill level function**, e.g. has VPN exploit skill level = 0.3,

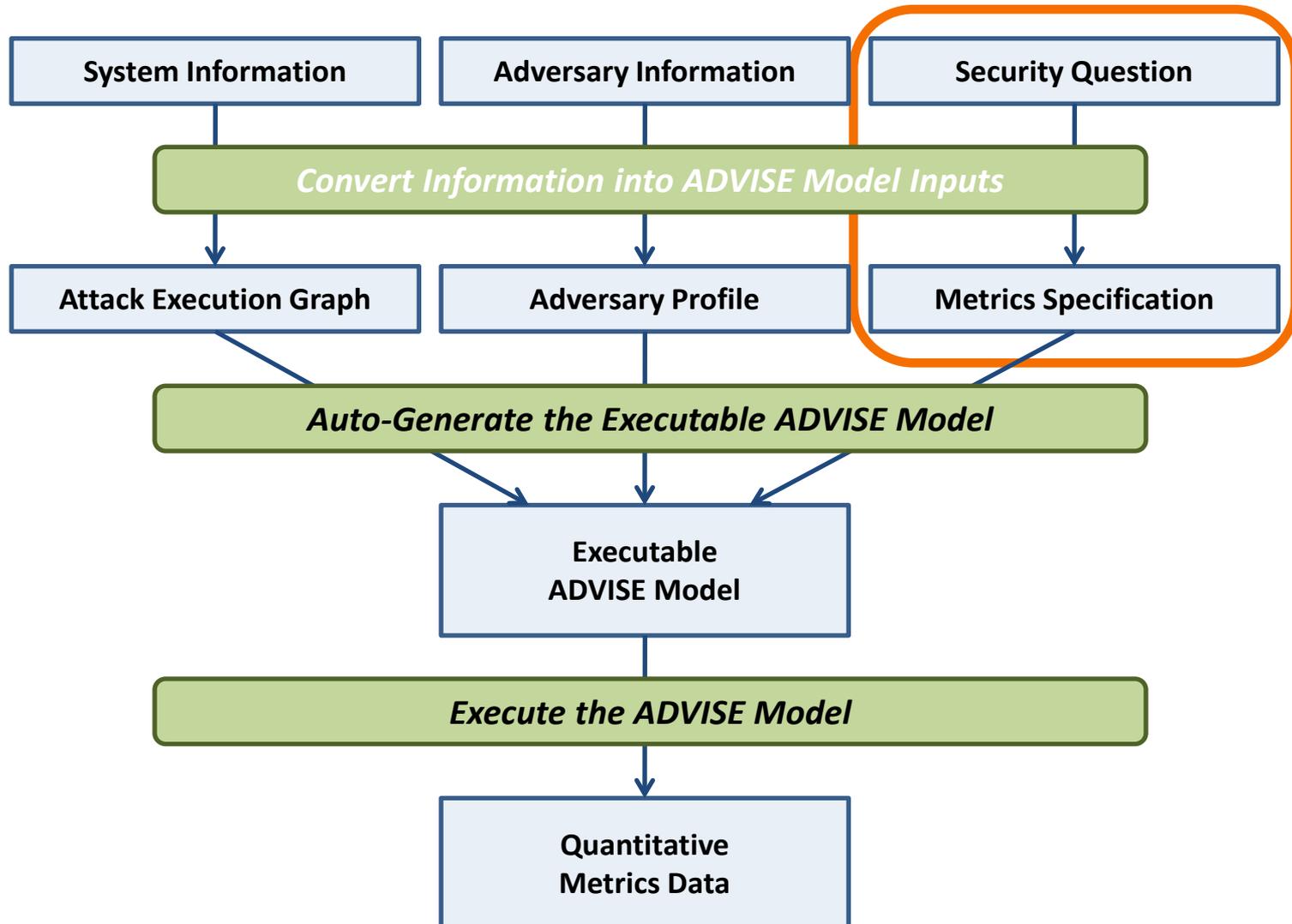
V is the **attack goal value function**, e.g., values “View contents of network” at \$5000,

w_C , w_P , and w_D are the **attack preference weights for cost, payoff, and detection probability**, e.g., $w_C = 0.7$, $w_P = 0.2$, and $w_D = 0.1$,

U_C , U_P , and U_D are the **utility functions for cost, payoff, and detection probability**, e.g., $U_C(c) = 1 - c/10000$, $U_P(p) = p/10000$, $U_D(d) = 1 - d$, and

N is the **planning horizon**, e.g., $N = 4$.

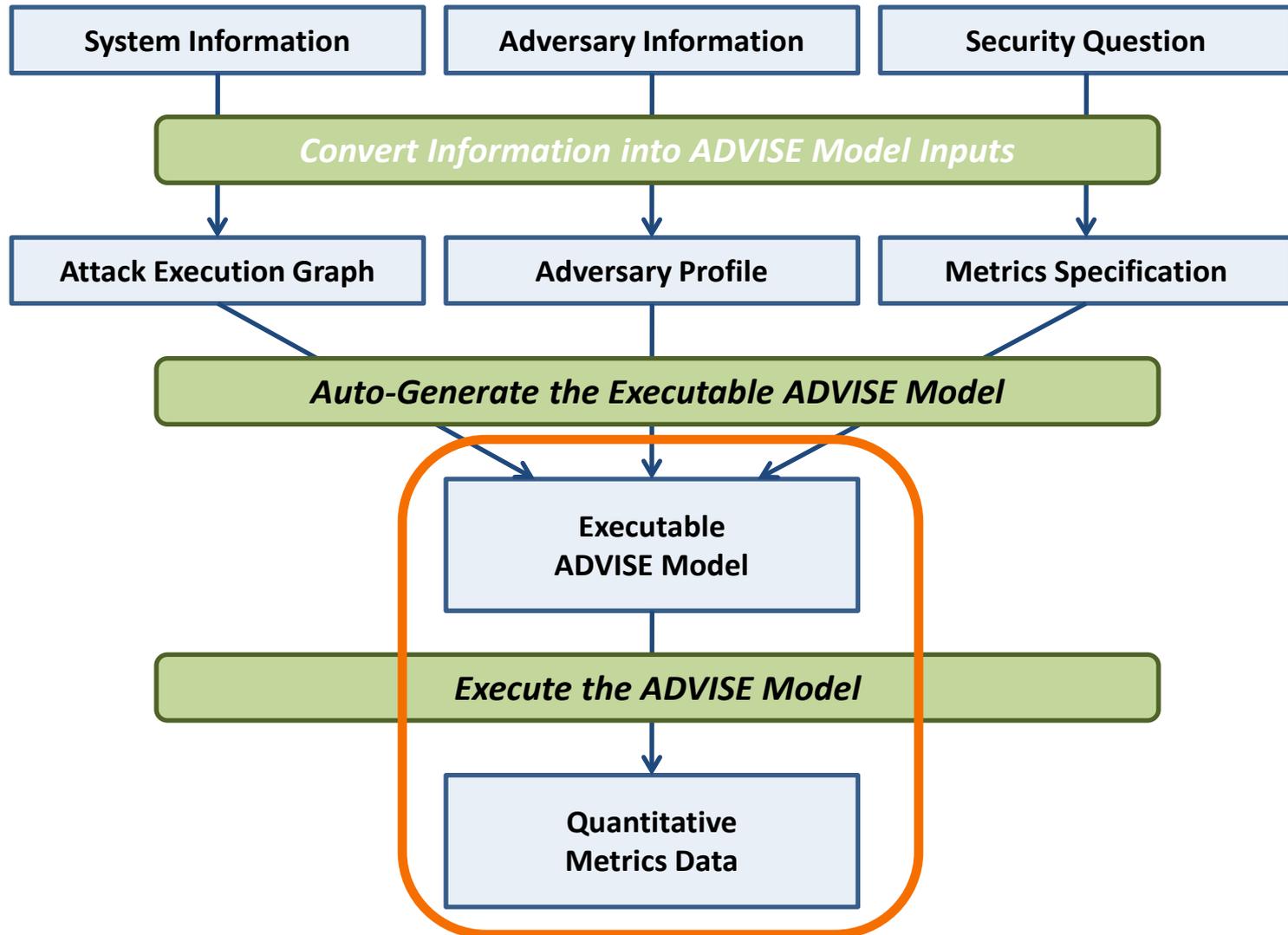
ADVISE Method Overview



ADVISE Security Question: Metrics Specification

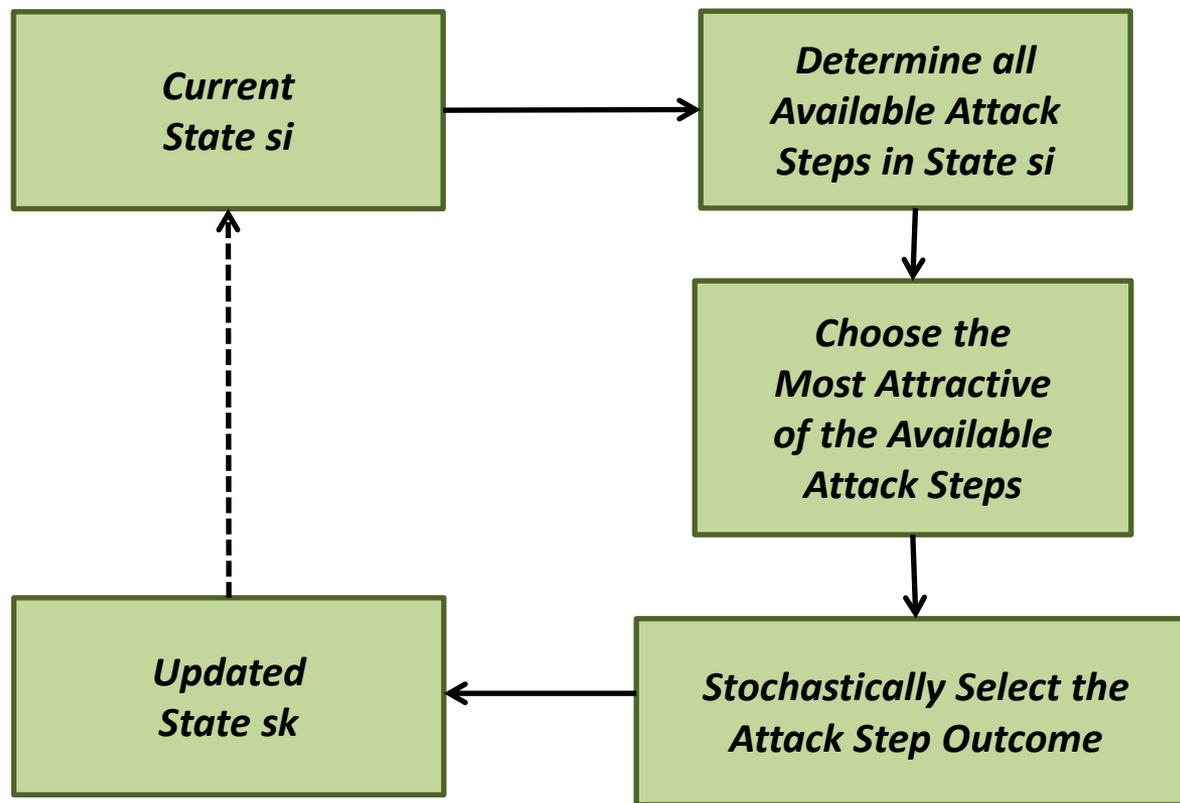
- State metrics analyze the model state
 - State occupancy probability metric (probability that the model is in a certain state at a certain time)
 - Average time metric (average amount of time during the time interval spent in a certain model state)
- Event metrics analyze events (state changes, attack step attempts, and attack step outcomes)
 - Frequency metric (average number of occurrences of an event during the time interval)
 - Probability of occurrence metric (probability that the event occurs at least once during the time interval)

ADVISE Method Overview



Model Execution: the Attack Decision Cycle

- The adversary selects the most attractive available attack step based on his attack preferences.
- State transitions are determined by the outcome of the attack step chosen by the adversary.



ADVISE Model Execution Algorithm

- 1: Time $\leftarrow 0$ Simulation time and model state initialization
- 2: State $\leftarrow s_0$
- 3: **while** Time < EndTime **do**
- 4: Attack_i $\leftarrow \beta^N(\text{State})$ Adversary attack decision
- 5: Outcome $\leftarrow o$, where $o \sim \text{Prob}_i(\text{State})$ Stochastic outcome
- 6: Time $\leftarrow \text{Time} + t$, where $t \sim T_i(\text{State})$ Time update
- 7: State $\leftarrow E_i(\text{State}, \text{Outcome})$ State update
- 8: **end while**

$\beta^N(s)$ selects the most attractive available attack step in model state s using a planning horizon of N

Goal-driven Adversary Decision Function

When the planning horizon N is greater than 1, the attractiveness of an available next step

is a function of

the payoff in the expected states

N attack steps from the current state

(the **expected horizon payoff**)

and

the expected cost and detection probability

of those N attack steps

(the **expected path cost** and **expected path detection**).

Goal-driven Adversary Decision Function

$$\text{Attractiveness of an attack step } a_i \\ \text{to an adversary with planning horizon } N = \\ UC(E[C]) * w_c + UP(E[P]) * w_p + UD(E[D]) * w_d$$

$E[C]$ = **Expected Path Cost** to get to a state N attack steps away via attack step a_i .

$E[P]$ = **Expected Horizon Payoff** in a state N attack steps away via attack step a_i .

$E[D]$ = **Expected Path Detection** to get to a state N attack steps away via attack step a_i .

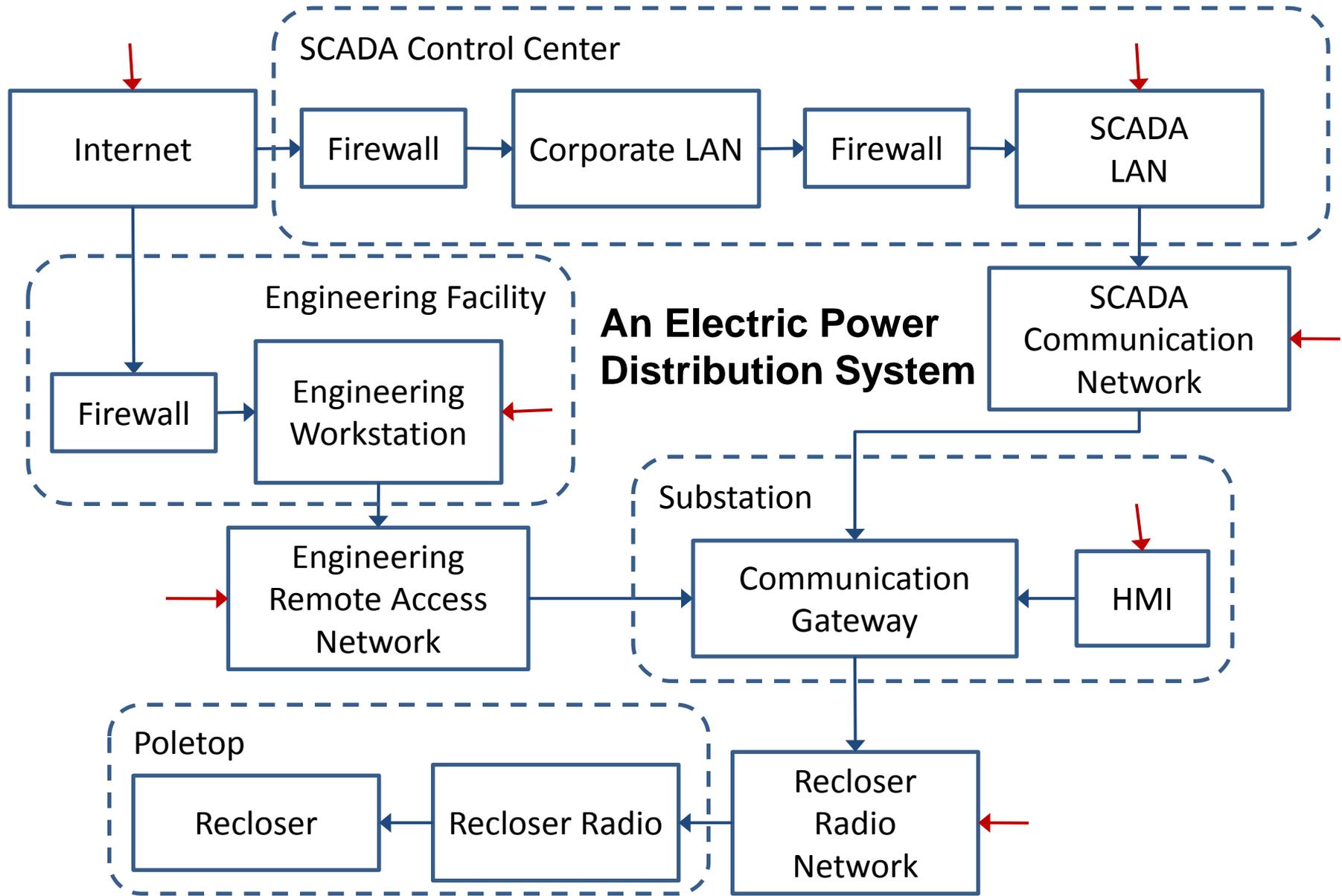
$E[C]$, $E[P]$, and $E[D]$ are computed using a **State Look-Ahead Tree**.

Practical Implications of Algorithm Optimality

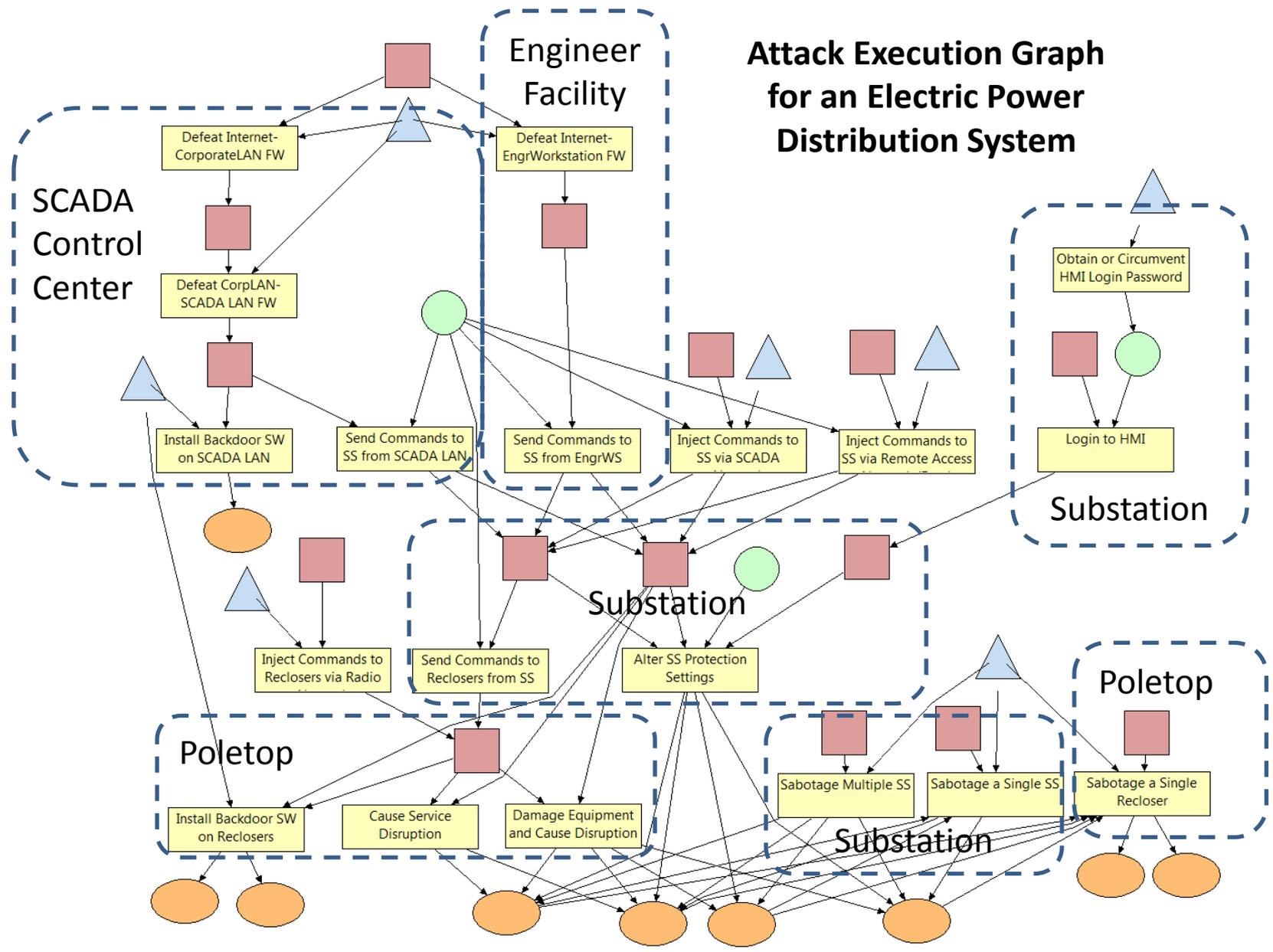
- Adversaries modeled using this algorithm exhibit “worst case” behavior, that is, they always select a next attack step that is best for them considering
 - Adversary attack preferences
 - Adversary planning horizon
 - Available attack steps
 - Attractiveness function definition

Case Study

- Investigates the effects of architectural changes on the security of an electric power distribution system
- In particular, analyze the security impact of adding radio communication between substations and poletop reclosers



Attack Execution Graph for an Electric Power Distribution System



Adversary Profiles: Decision Parameters

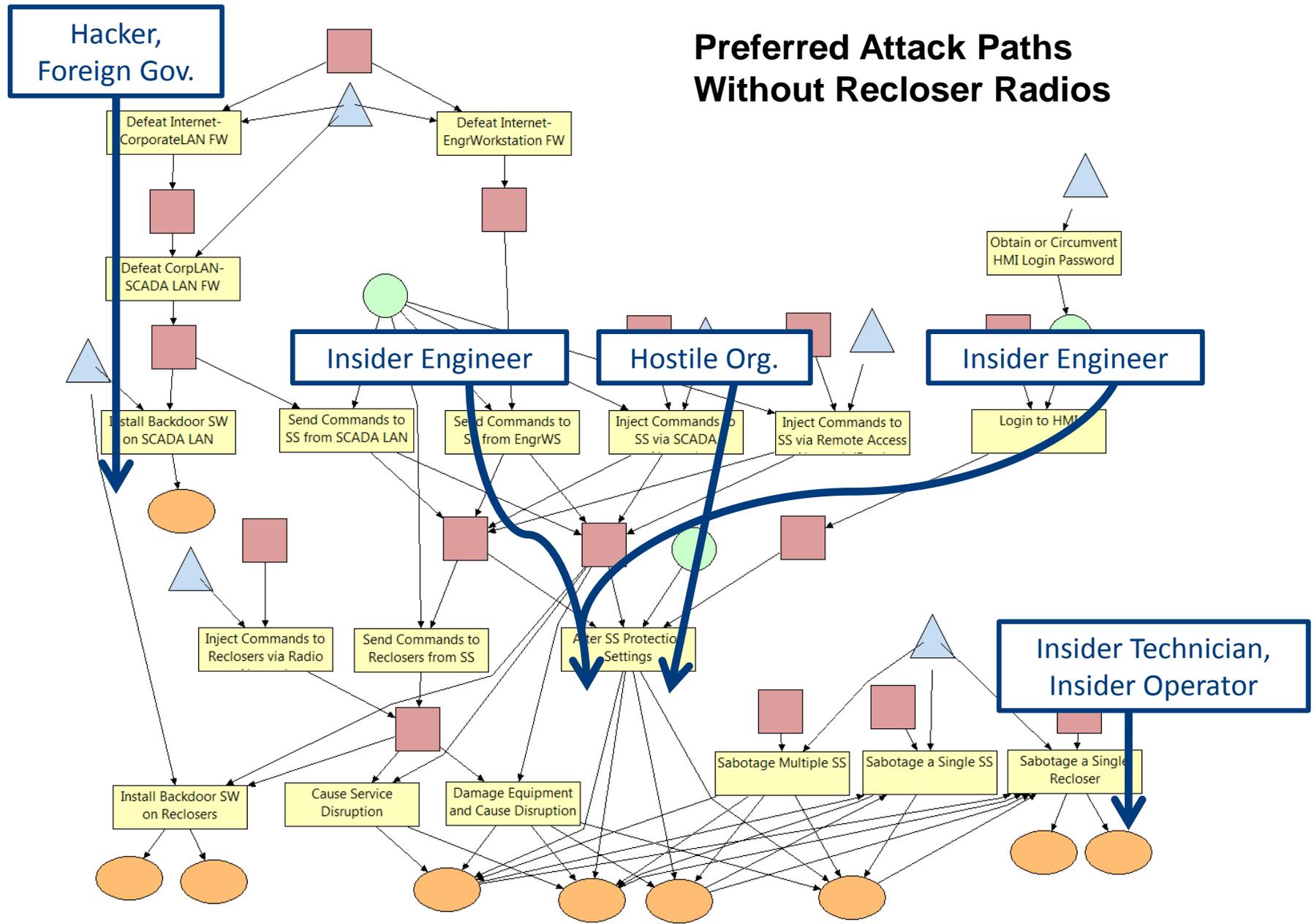
	Foreign Government	Hacker	Hostile Organization	Insider Engineer	Insider SCADA Operator	Insider Remote Technician
Cost Preference Weight	0	0.2	0.05	0.2	0.2	0.2
Detection Preference Weight	0.5	0.4	0.2	0.1	0.1	0.1
Payoff Preference Weight	0.5	0.4	0.75	0.7	0.7	0.7

- The Foreign Government adversary is very well-funded but risk-averse.
- The Hacker is resourced-constrained.
- The Hostile Organization is moderately well-funded and more driven by payoff than the others.
- The Insider Engineer, Insider Technician, and Insider Operator are resource-constrained but willing to take risks.

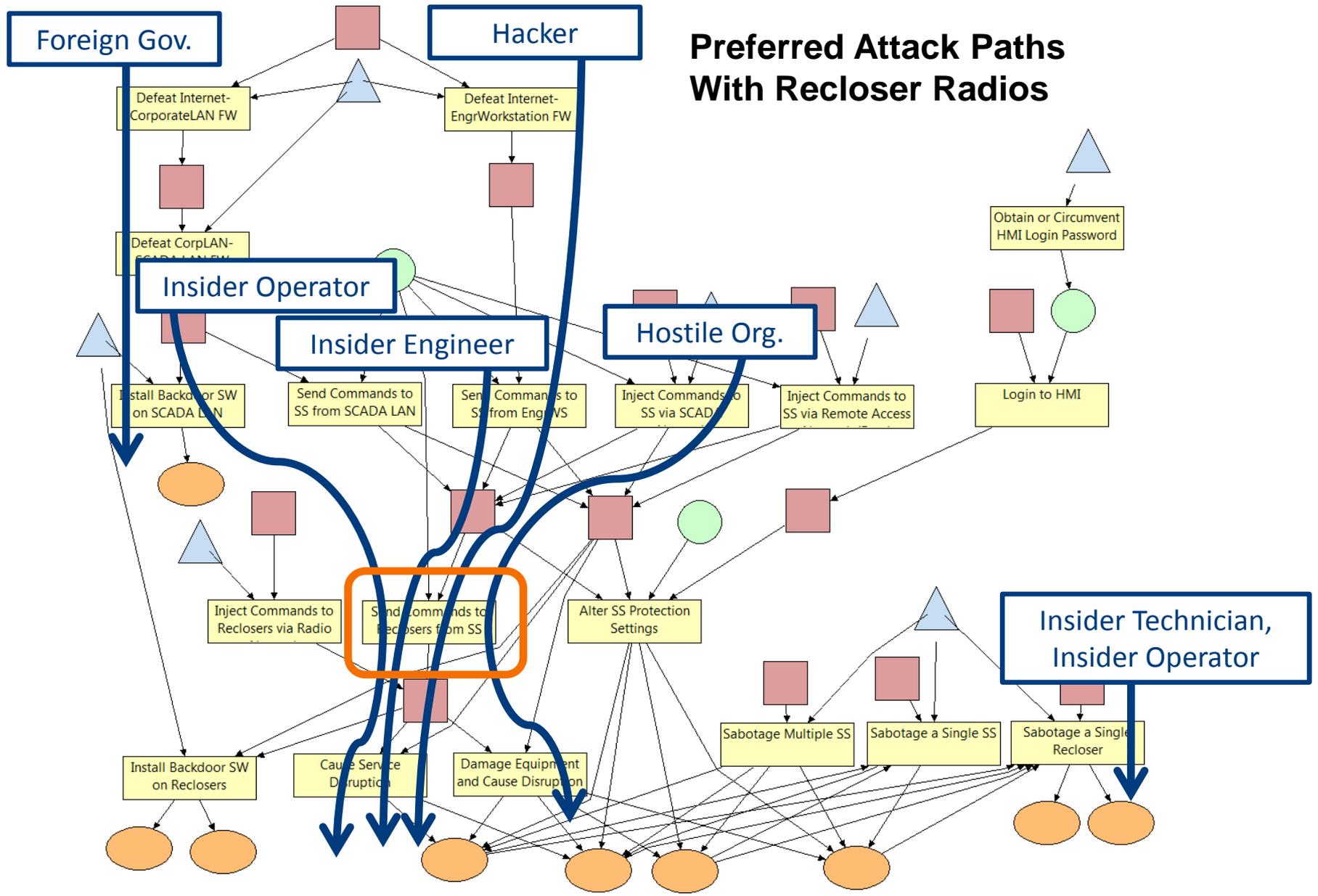
Security Metrics

- **Average Number of Attempts**
 - Report for each attack step
 - Gives insight on preferred attack path of adversary
- **Probability of Attack Goal Achieved at End Time**
 - Report for each attack goal
 - Gives insight on what goals the adversary is actively pursuing and reaching
- **Average Time-To-Achieve-Goal**
 - For attack goals where the above probability metric is 1 (or close to 1)
 - Gives insight on the speed of the adversary's attack

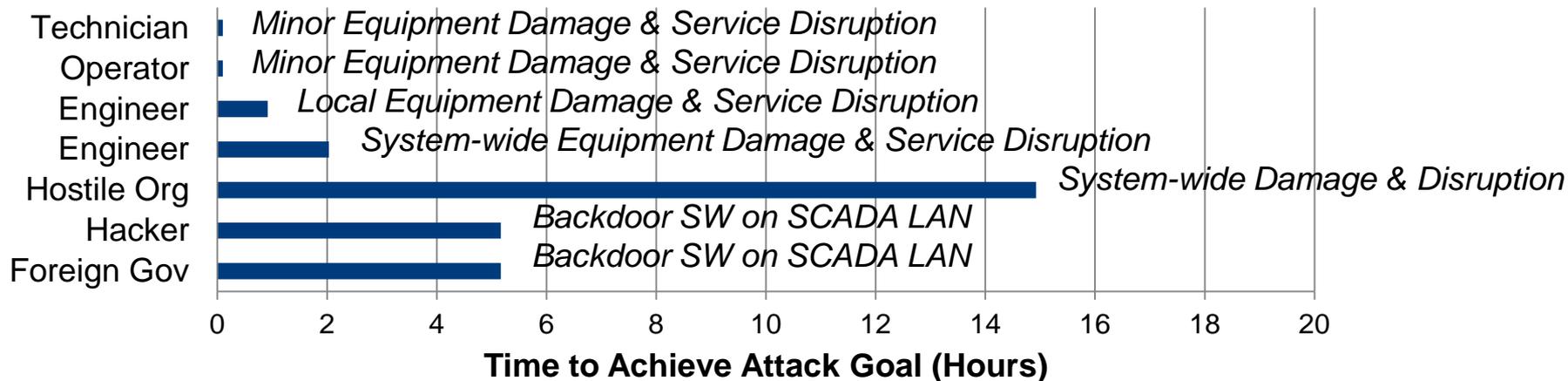
Preferred Attack Paths Without Recloser Radios



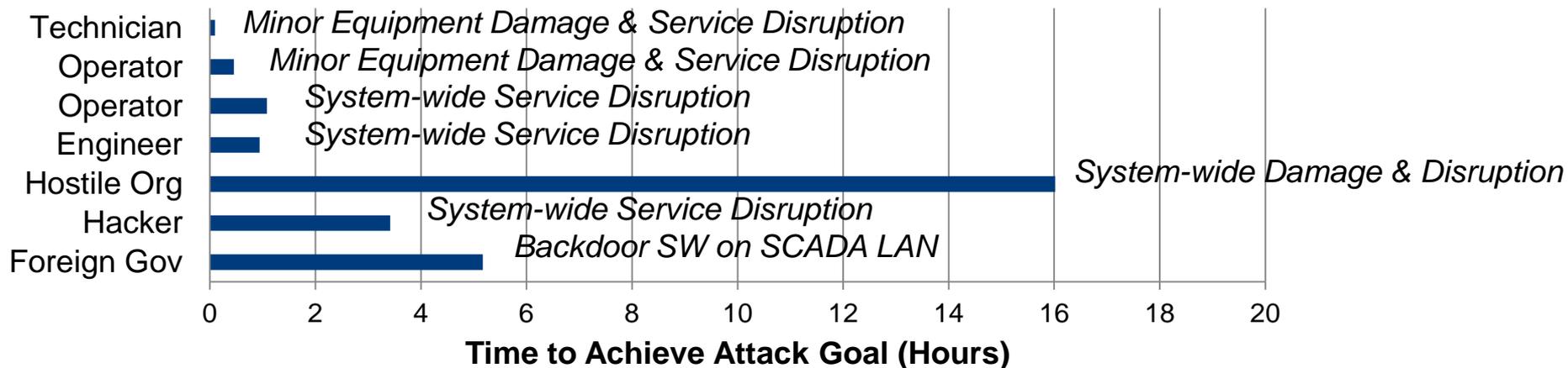
Preferred Attack Paths With Recloser Radios



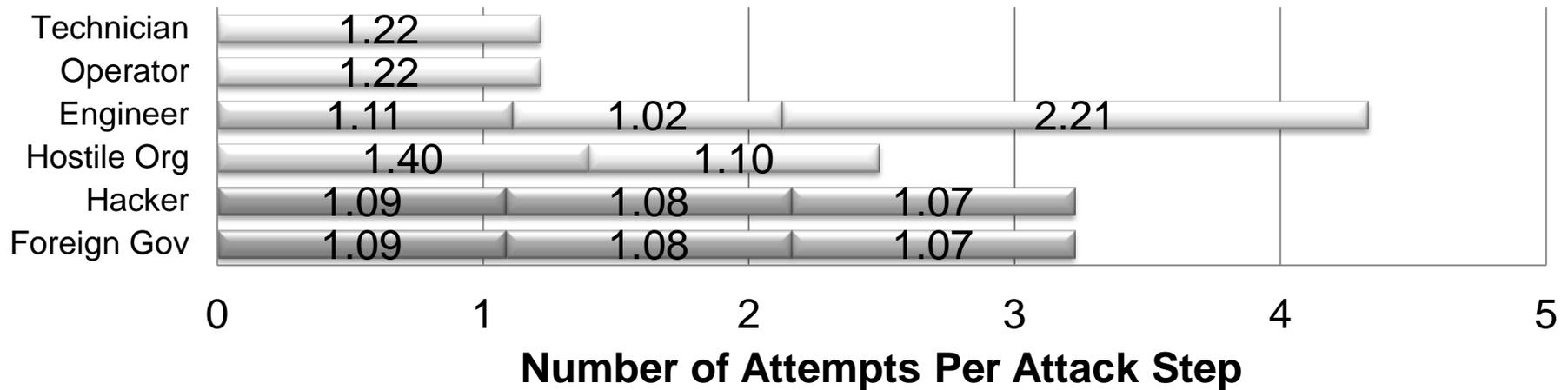
Attack Speed Without Recloser Radios



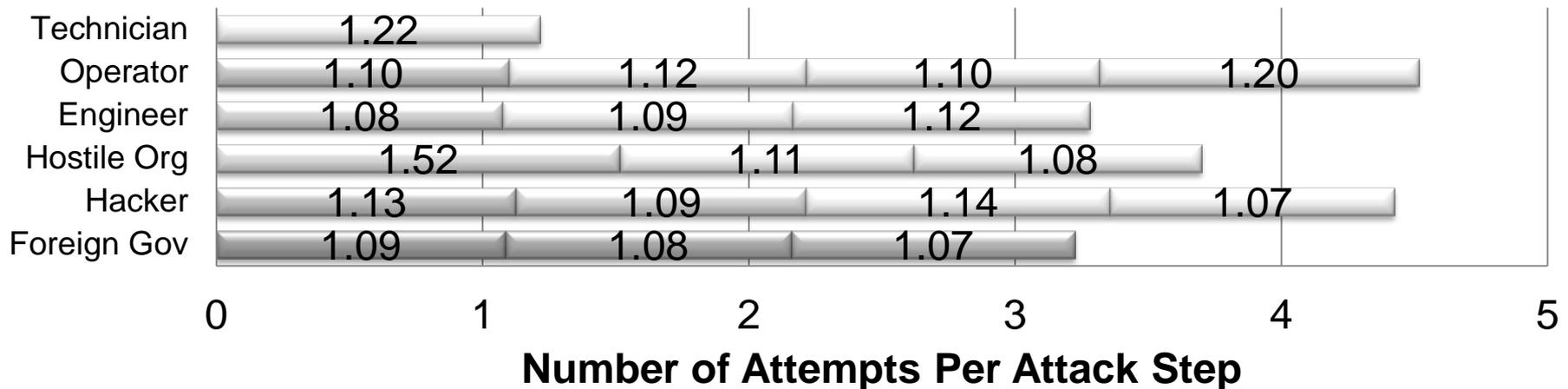
Attack Speed With Recloser Radios



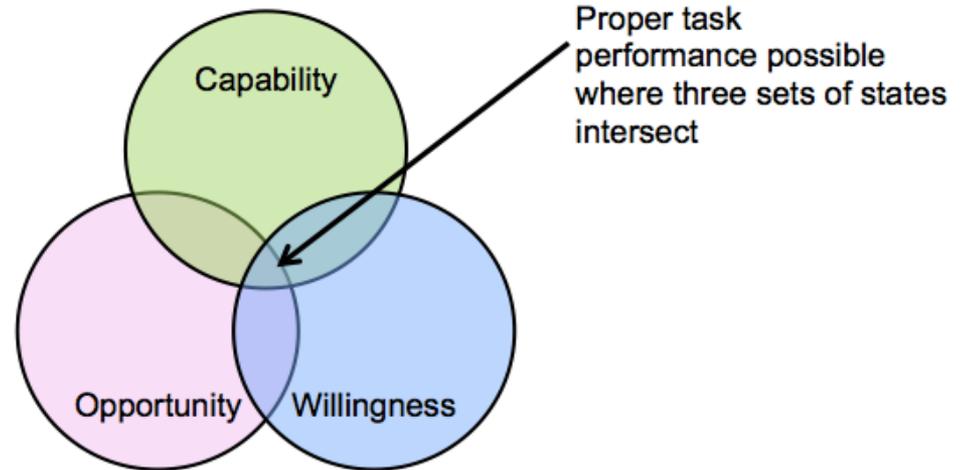
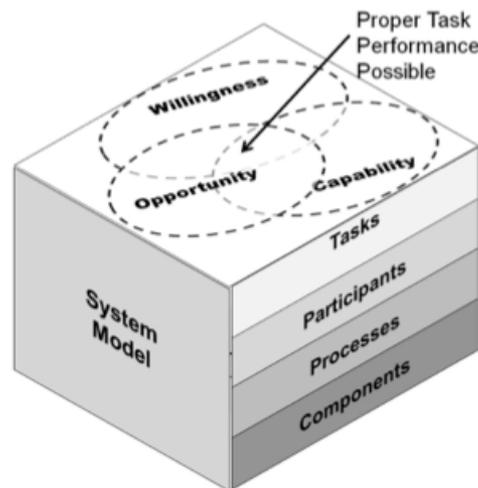
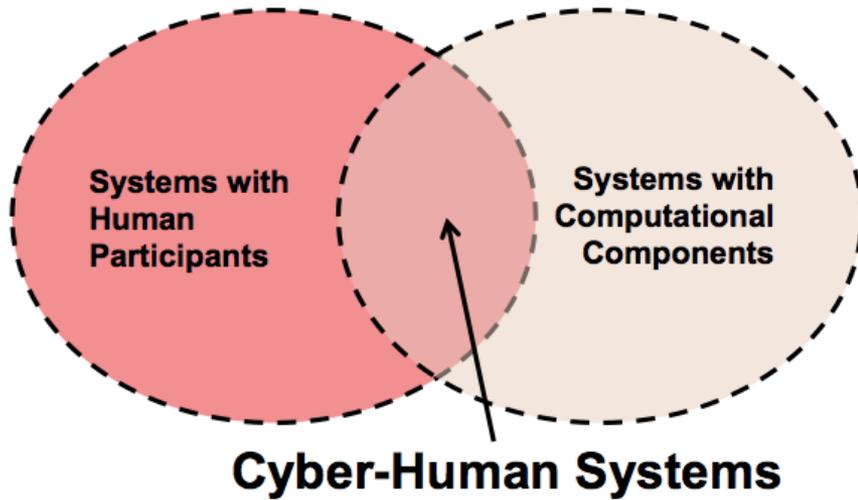
Number of Attack Attempts Without Recloser Radios



Number of Attack Attempts With Recloser Radios



Ongoing Research: Modeling Cyber-Human-Physical Systems



ADVISE Team

University of Illinois Urbana-Champaign

Mike Ford

Ken Keefe

Elizabeth LeMay

Bill Sanders

Cyber Defense Agency, Inc.

Carol Muehrcke

Case study collaborators

Bruce Barnett and Michael Dell' Anno,
GE Research

Research sponsored by Science and Technology Directorate,
Department of Homeland Security, GE Research, NSA Science of
Security Center

Conclusions

- Since system security cannot be absolute, quantifiable security metrics are needed
- Metrics are useful even if not perfect; e.g., relative metrics can aid in critical design decisions
- The ADVISE formalism, and its implementation in Mobius-SE
 - Is rich enough to adversary, user, and system behavior
 - Natural for security analysts
 - Semantically precise
- Mobius-SE is in alpha-test, and has been distributed to 10 organizations (industry, govt., & academics) who are using it in real case studies.
- Public release in Early September: mobius.illinois.edu
- Work is on going on modeling human user behavior

Thank you!

Bill Sanders

mobius.illinois.edu

perform.csl.illinois.edu

whs@illinois.edu