Modeling Cyber-Threats

Adopting Bayes' principles in the Attack Graphs Theory

Author

Mattia Zago

VR367531

Supervisor

Prof.

Isabella Mastroeni

Examiner

Prof.

Alessandro Farinelli

Co-Supervisor

Full Prof.

Gregorio Martinez

Our Agenda for Today

A Motivation Example

Understand the idea behind this work

A Real-World Example

Understand the actual issues

The Simulator

Our solution for the comparison problem

Conclusions

A short summary

A Motivation Example

Understand the idea behind this work

> “ Hardware:
> The parts of a computer system
> that can be kicked. ”

How to access to my sister’s laptop?

design an attack graph

How to access to my sister's laptop?

design an attack graph

Physical Access

Remote Access

Get the laptop

Steal it

Ask her

Get the password

Use a Tool

Catch it

Connect to it

Laptop Connect

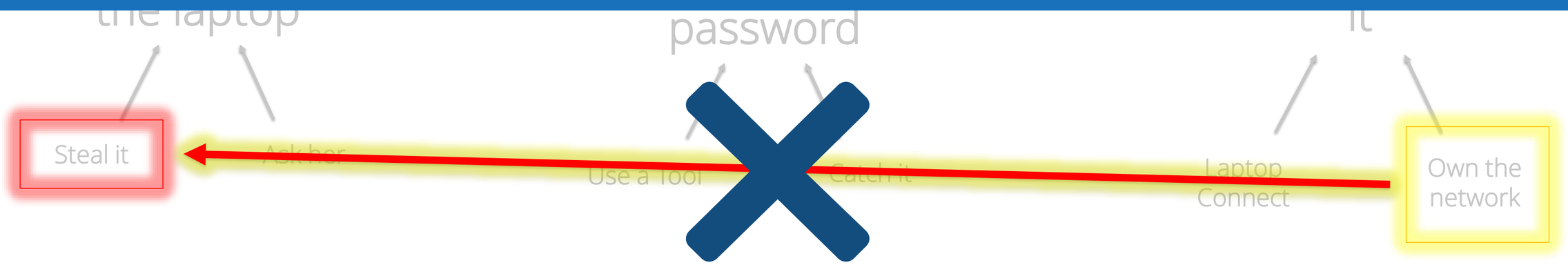
Own the network

How to access to my sister’s laptop?

design an attack graph

“ If I already own the network,
which is the probability of the event “steal the laptop” ? ”

ATTACK GRAPHS CANNOT ANSWER TO THIS QUESTION



How to access to my sister's laptop?

design an attack graph

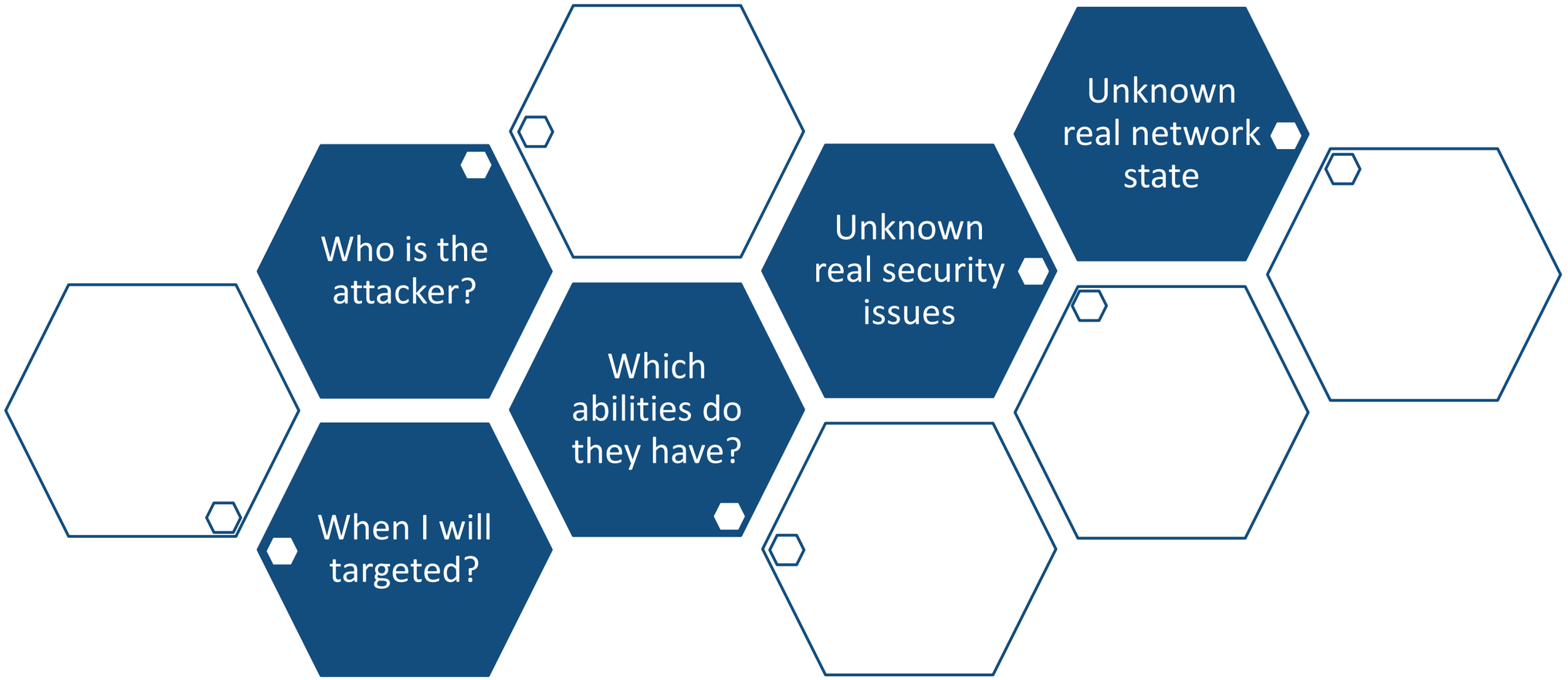
"If I already own the network,
which is the probability of the event "steal the laptop" ?"

BUT
BAYESIAN NETWORKS CAN!

What is a Bayesian Network?

and how it can be useful?

BNs are **statistical** models built on the attack graph structure aiming to deal with the **uncertainty**



Essentially they are ***Directed Acyclic Graphs*** decorated with **probabilities**

What is a Bayesian Network?

and how it can be useful?

In BNs there are 4 probability function

PRIOR

Express the uncertainty of an event before any else information is available

$$\mathcal{P}(A)$$

CONDITIONAL

Express the probability of an event A given the occurrence of another event B

$$P(A|B) = \frac{P(A \cap B)}{P(A)}$$

JOINT

Express the probability of an event A given the occurrence of all its ancestor

$$\text{JPT}(A) = P(A|\,\text{parent}(A))$$

POSTERIOR

Express the probability of an ancestor event A given the knowledge about a child event C

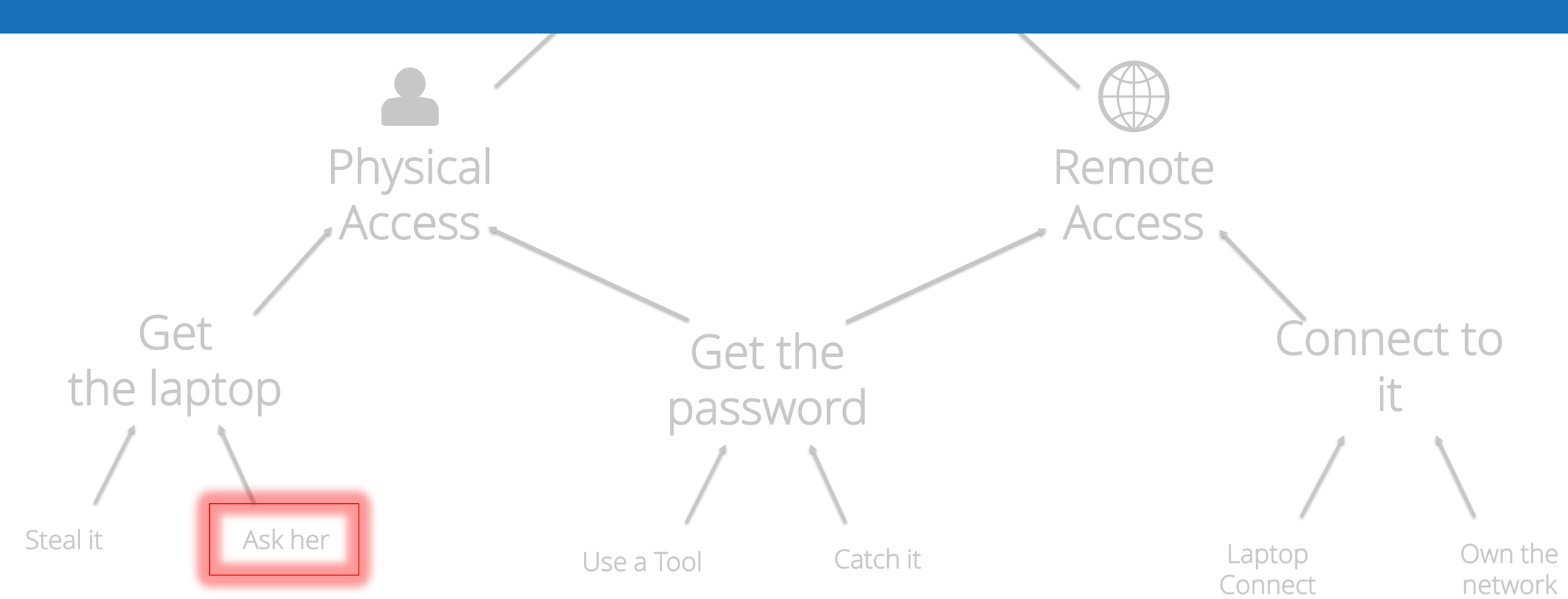
$$P(A|C) = \frac{\text{JPT}(A, C)}{M_A(C)}$$

What is a Bayesian Network?

and how it can be useful?

“Which is the probability of
“she gives me her laptop” ?”

PRIOR PROBABILITY

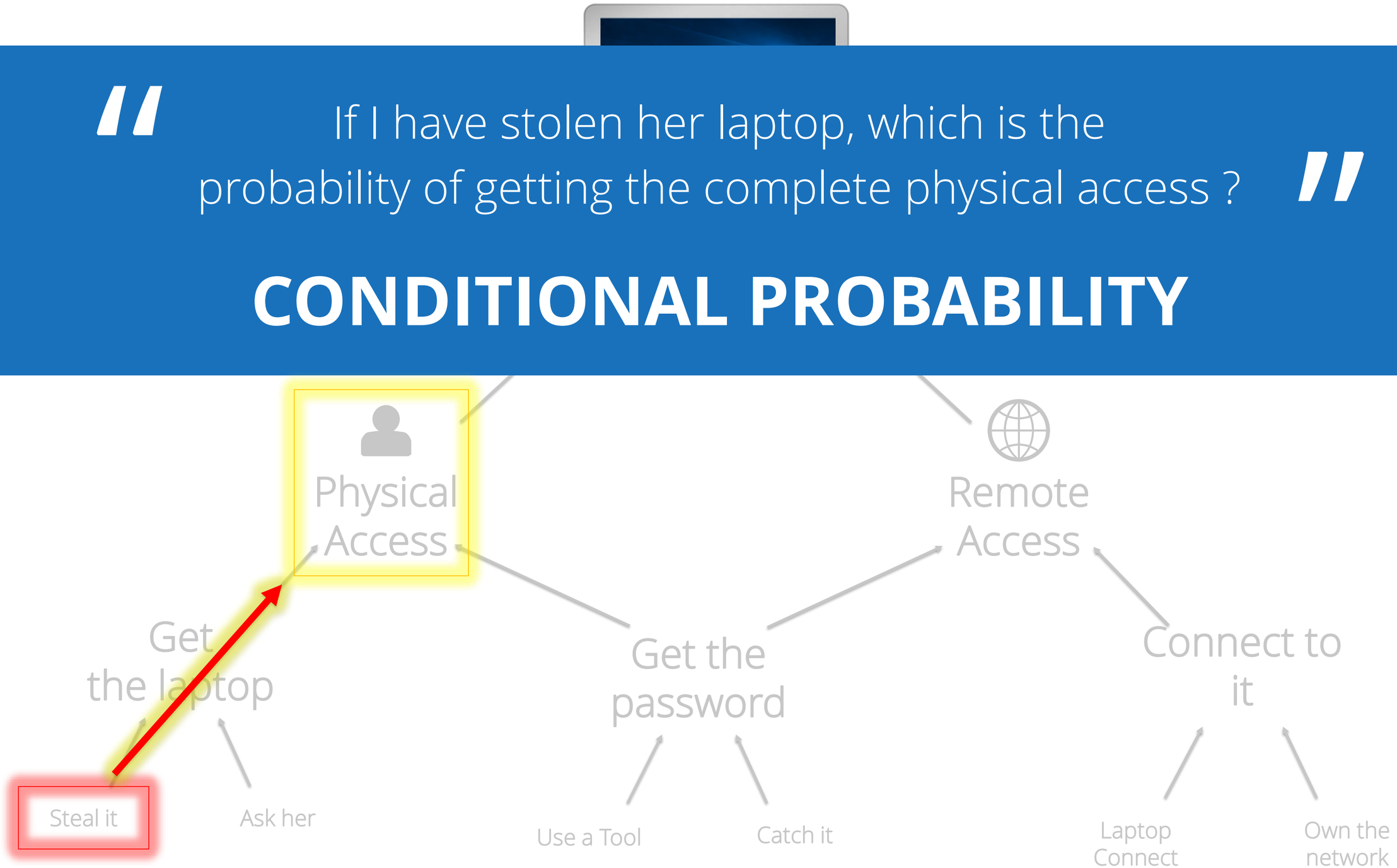


What is a Bayesian Network?

and how it can be useful?

"If I have stolen her laptop, which is the probability of getting the complete physical access ?"

CONDITIONAL PROBABILITY

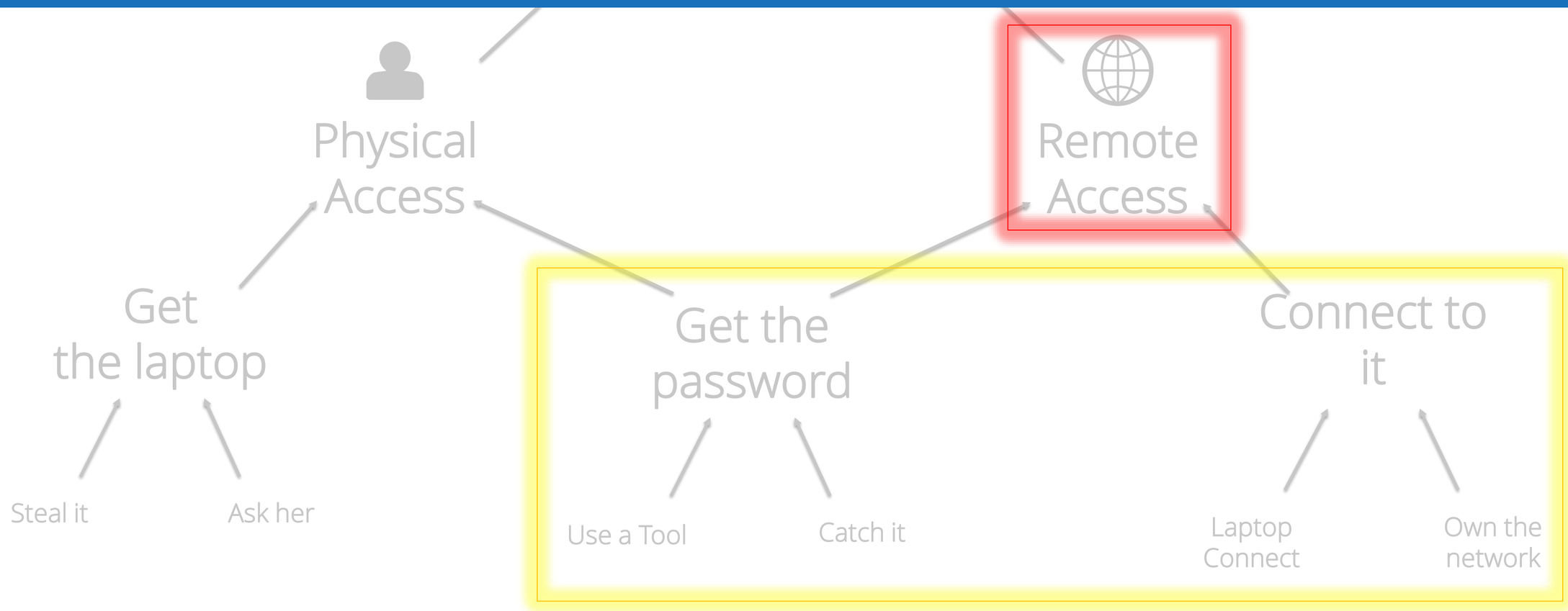


What is a Bayesian Network?

and how it can be useful?

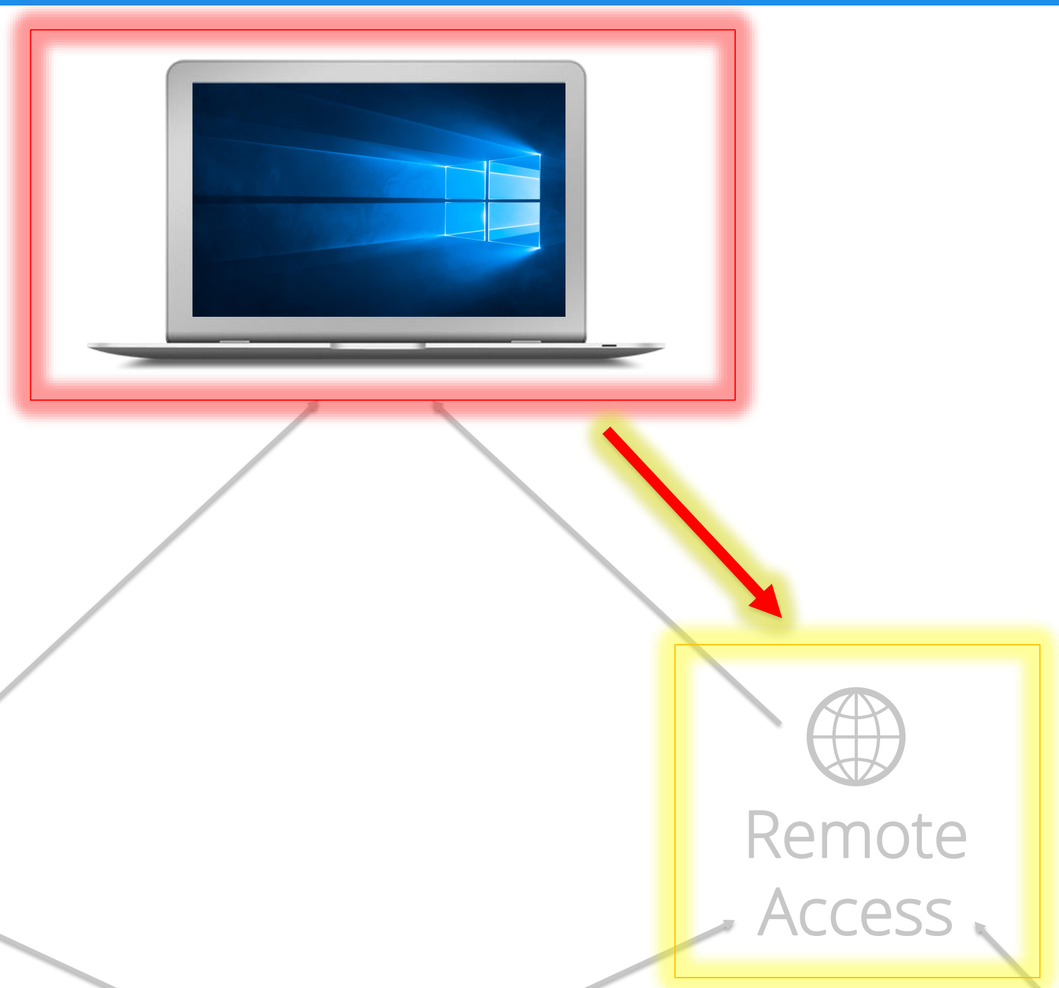
“Which is the probability of “complete remote access” ?”

JOINT PROBABILITY



What is a Bayesian Network?

and how it can be useful?



"She finds out that I had access to the laptop.
Which is the probability of the "Remote Access" event?"

POSTERIOR PROBABILITY

Bayesian Network – Comparison Table

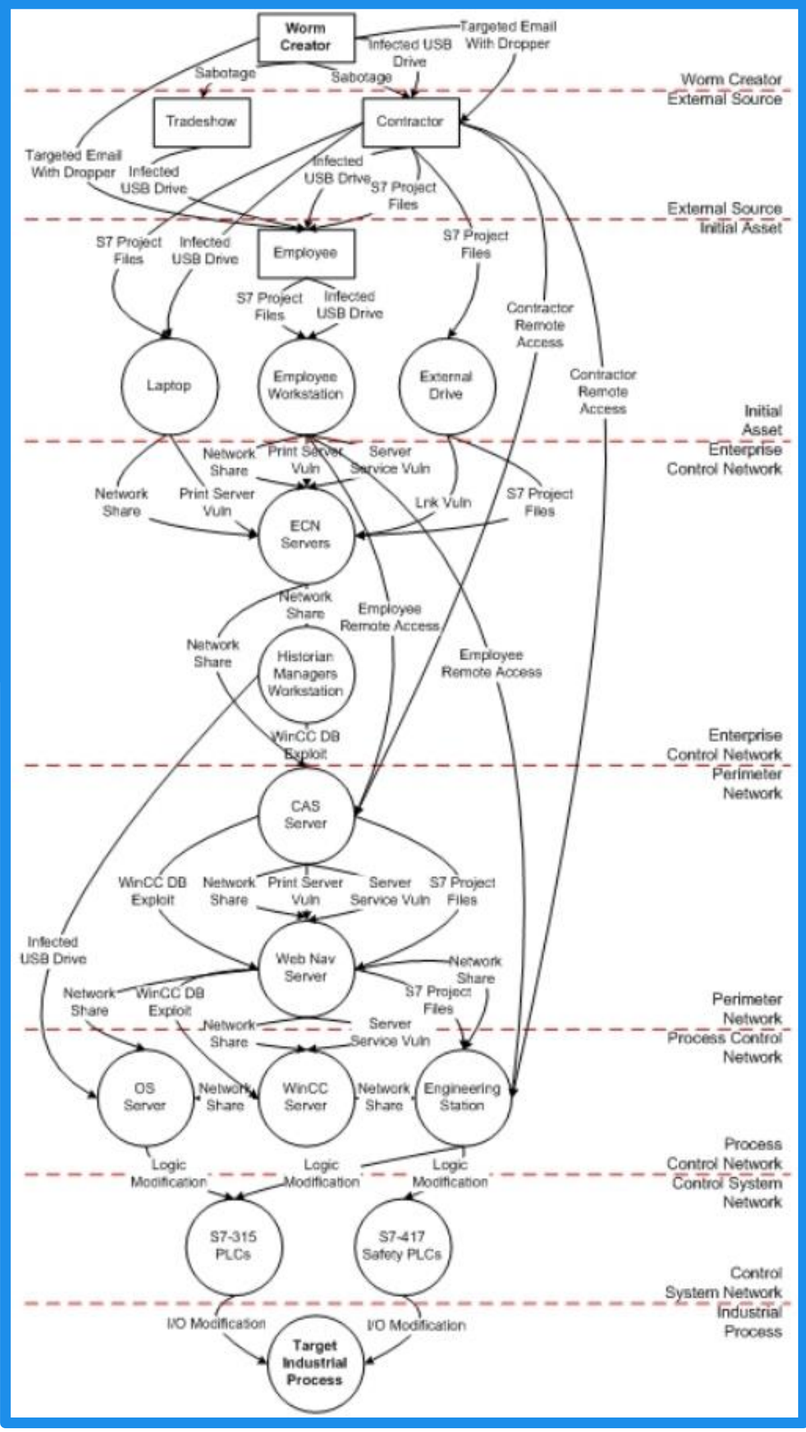
BN are statistical models build on the attack graph structure

Attack Graphs	Attack Trees	Bayesian Networks
Generic Graphs	Directed Acyclic Graphs	Directed Acyclic Graphs
Exponential Path Analysis	Polynomial Path Analysis	Polynomial Path Analysis
Can represents any dependencies	Can represents almost any dependencies	Can represents almost any dependencies
Cannot quantify the relationships	Cannot quantify the relationships	Can quantify the relationships
Generic domains	Generic domains	Finite, exhaustive, mutually exclusive domains
X	Monotonicity	Monotonicity
X	X	Probabilistic and Statistic functions

A Real-World Example

Understand the actual issues

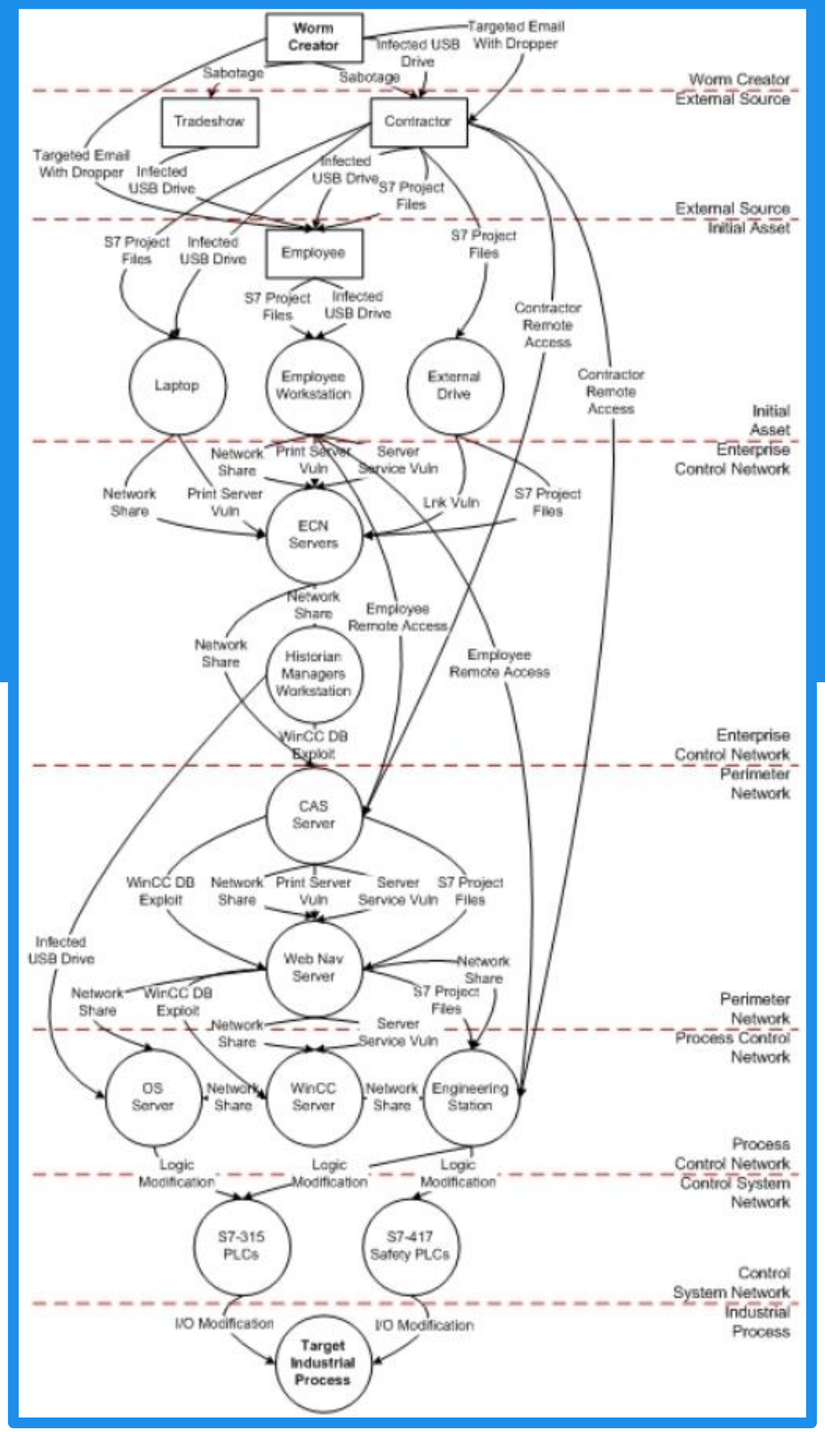
> “If brute force
> doesn't solve your problems,
> then you aren't using enough.”



A Real-World Example

Understand the actual issues

- ✓ Not always clear
- ✓ Reflect the complexity of the network
- ✓ Thousands of nodes



A Real-World Example

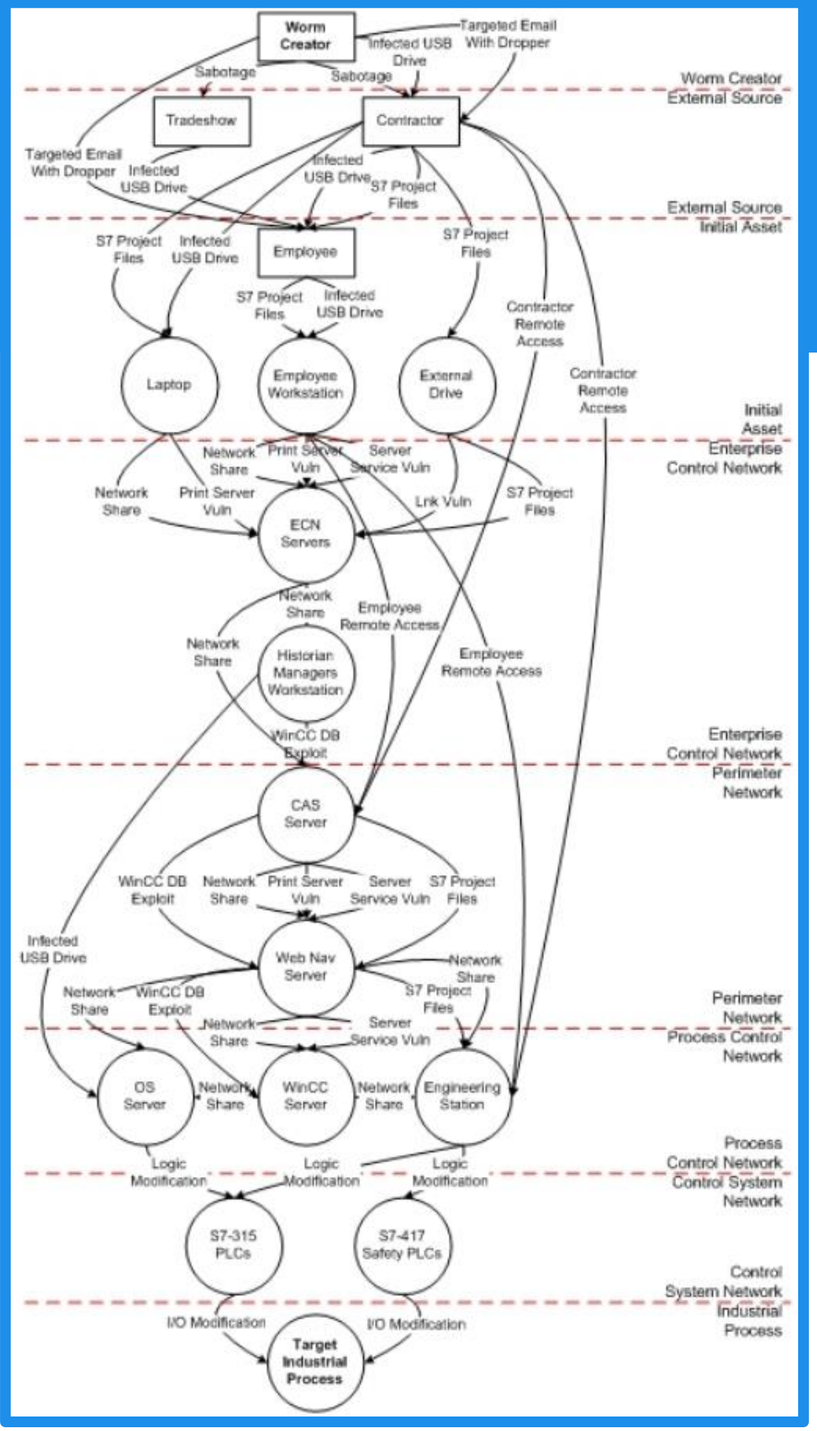
Understand the actual issues

- ✓ Not always clear
- ✓ Reflect the complexity of the network
- ✓ Thousands of nodes

In the past 15 years
were designed **several models**
to address these problems.

Today **does not exists** any model
that solved them completely
within a *reasonable* time.

Which is the best model?



A Real-World Example

Technical Difficulties

We cannot easily compare them

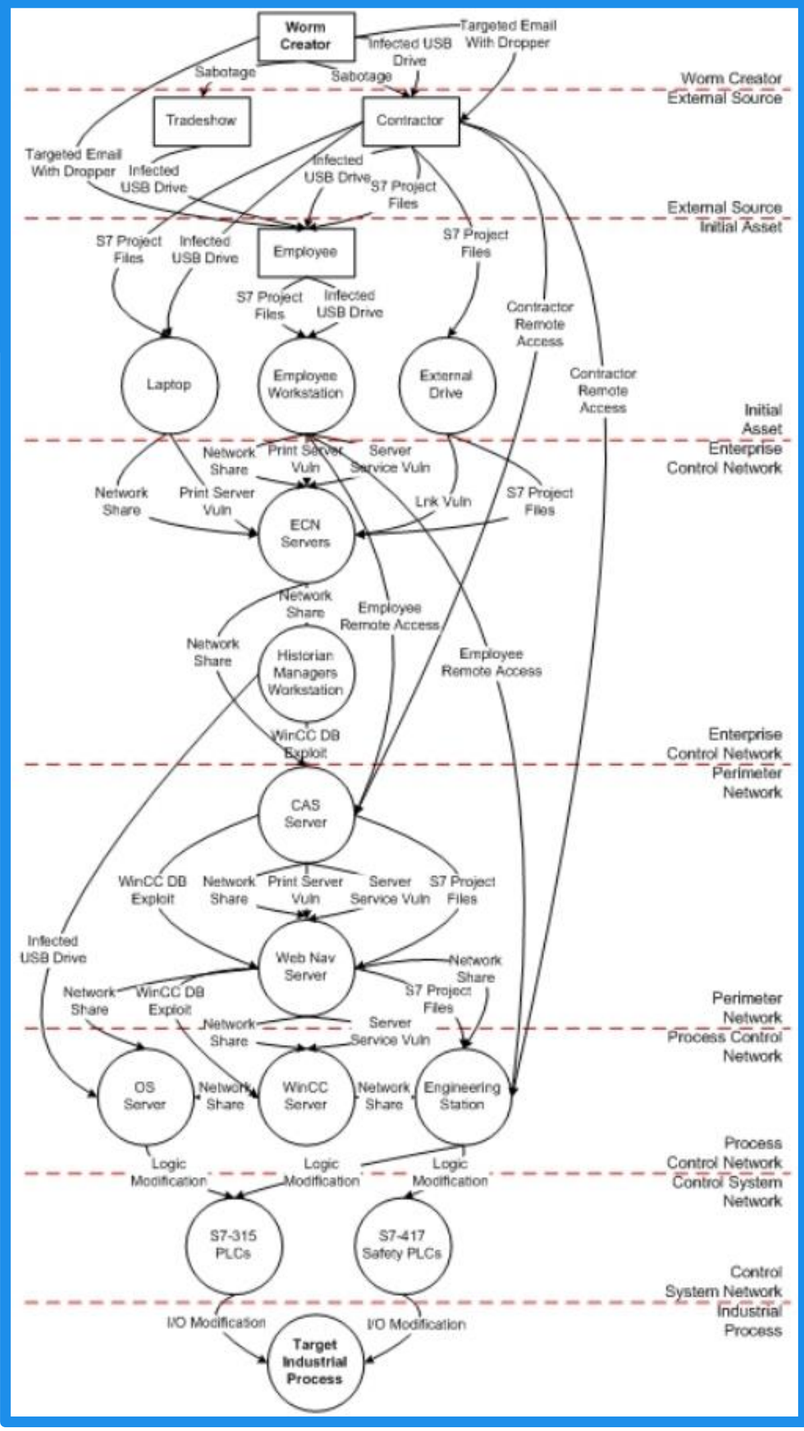
Literature Model A

- ✓ Pro 1
- ✓ Pro 2
- ✓ Pro 3
- × Cons 1
- × Cons 2
- × Cons 3

Literature Model B

- ✓ Pro 1
- ✓ Pro 2
- ✓ Pro 3
- × Cons 1
- × Cons 2
- × Cons 3

Given two generic models they are not immediately comparable.
The main issue is related to “how” compare the results because most of them are not complete or they analyze different aspect of the same problem.



A Real-World Example

Technical Difficulties

Literature Model A

Literature Model B

- ▶ Different approaches
- ▶ Different structures

- Uncompleted models ◀
- Impracticable algorithms ◀

BUT

They are solving the same problem, with a similar basic structure

The Simulator

Our solution for the comparison problem

> “ My software never has bugs.
> It just develops random features. ”

Multigraph Simulator

Design, Objectives and Functionalities

Joint project between the
University of Verona
and the
University of Murcia

JNIC2015

M. Zago, M. Gil Perez, G. Martinez Pérez e J. J. Andreu Blazquez

Multigraph Project: First steps towards the definition of a multiple attack graph model simulator

JNIC - Jornadas Nacionales de investigación en Ciberseguridad
Leon (ES), 2015

Multigraph Simulator

Design, Objectives and Functionalities

Objectives

Simulate

Implement different literature's models and run them with the same network examples

Observe

Analyze and measure the behavior of each implemented model

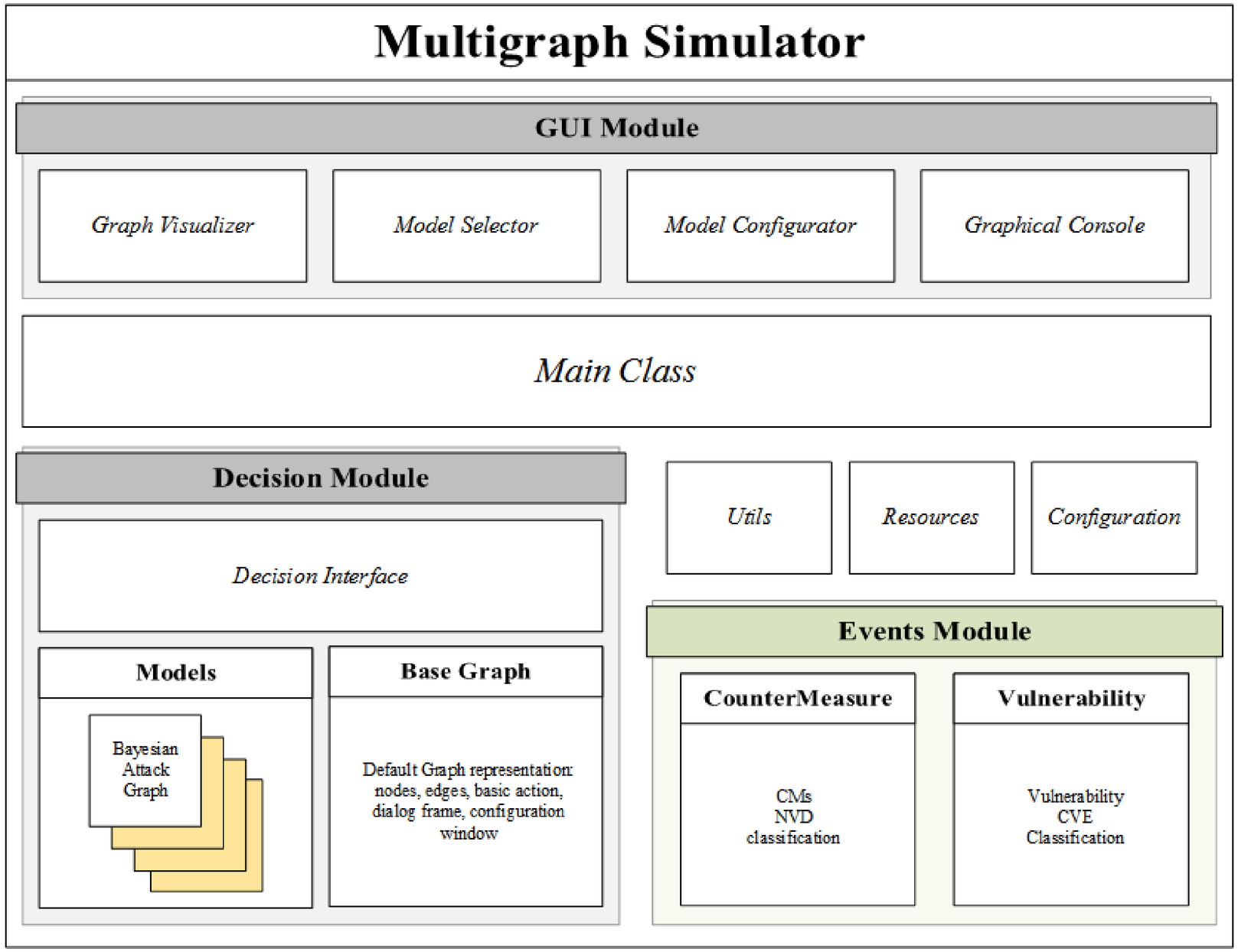
Compare

Compare the results and the performances when the models are comparable

Calibrate

Find incomplete aspects of implemented models and adjust its performances

Architecture Design

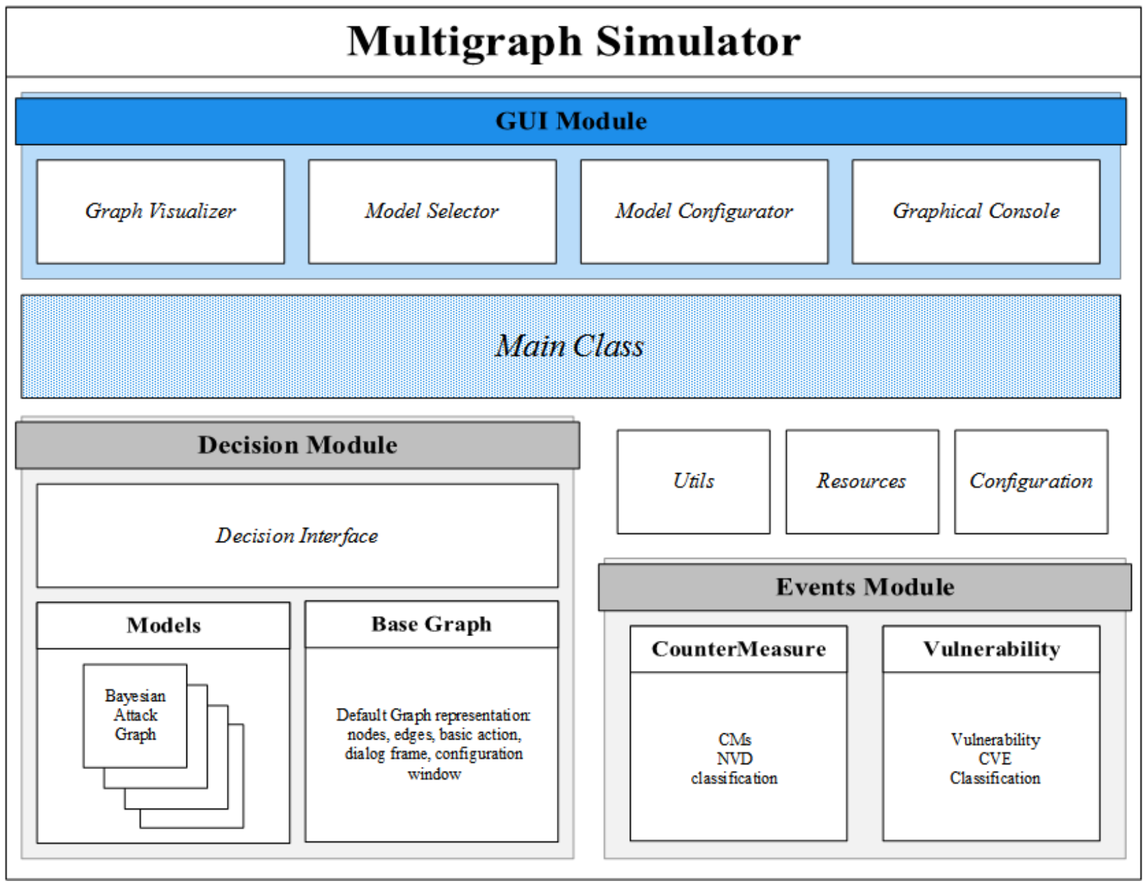


Juan José Andreu Blázquez
Bachelor Degree

Manuel Gil Perez
PhD

Multigraph Simulator

Design, Objectives and Functionalities

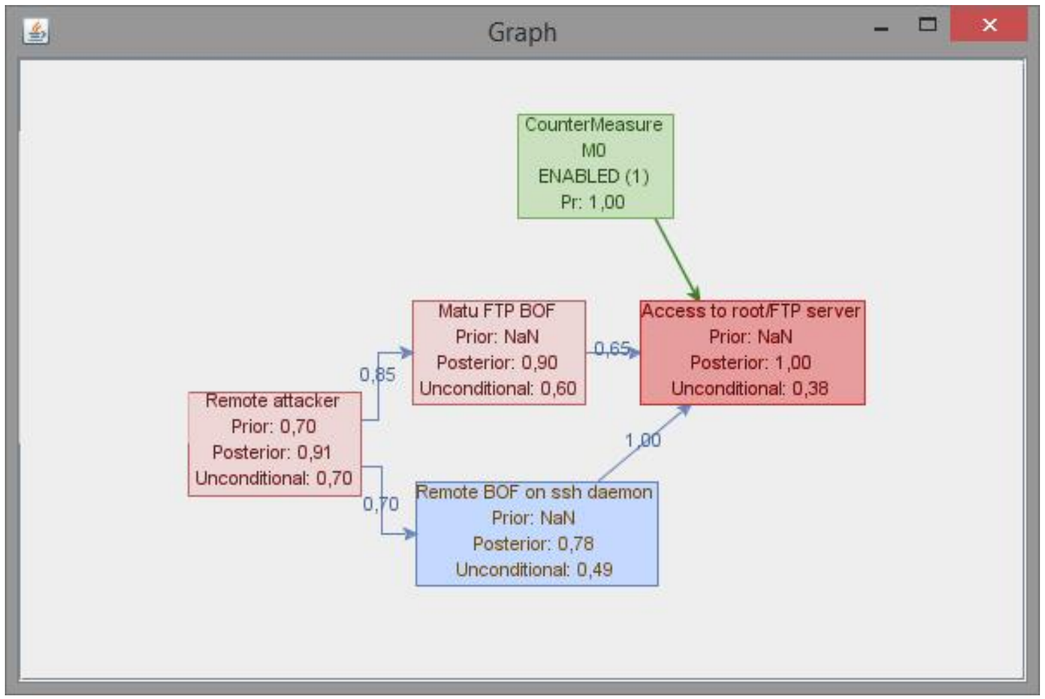


GUI MODULE

How users can interact with the simulator

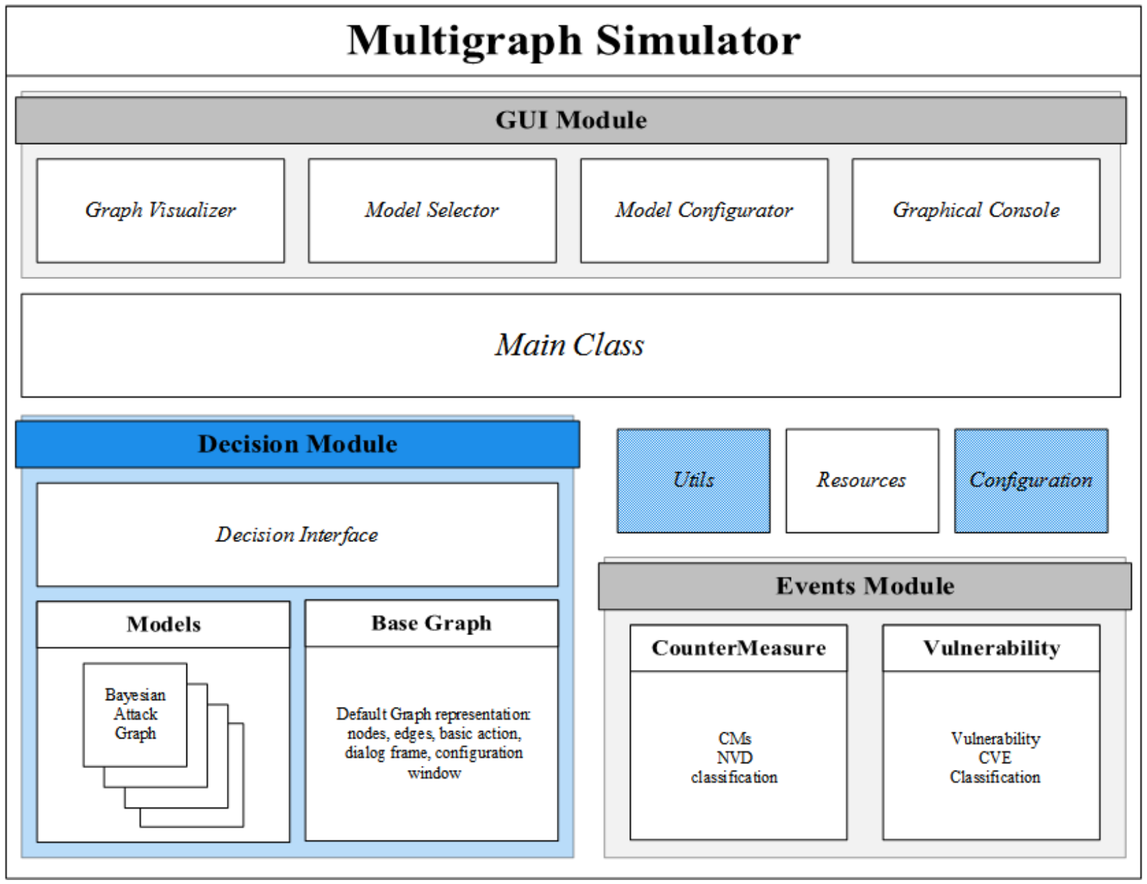
- Generic configurator (using Java reflection) that can be overwritten by the author of each specific model.
- Include the standard output directly to the main window.
- Show the general performances and the basic comparison charts (memory, time, size of the graph, etc.)
- Allows to control the simulator (start/stop threads etc.)

Graph Visualizer



Multigraph Simulator

Design, Objectives and Functionalities



DECISION MODULE

Extend the simulator with custom models

- Public API interface.
- Basic graph structure that provide canonical functionality (such as path algorithms, parent and children's pointer etc.)
- Can be extended with custom comparison metrics (risk scale, performance monitor, etc.)

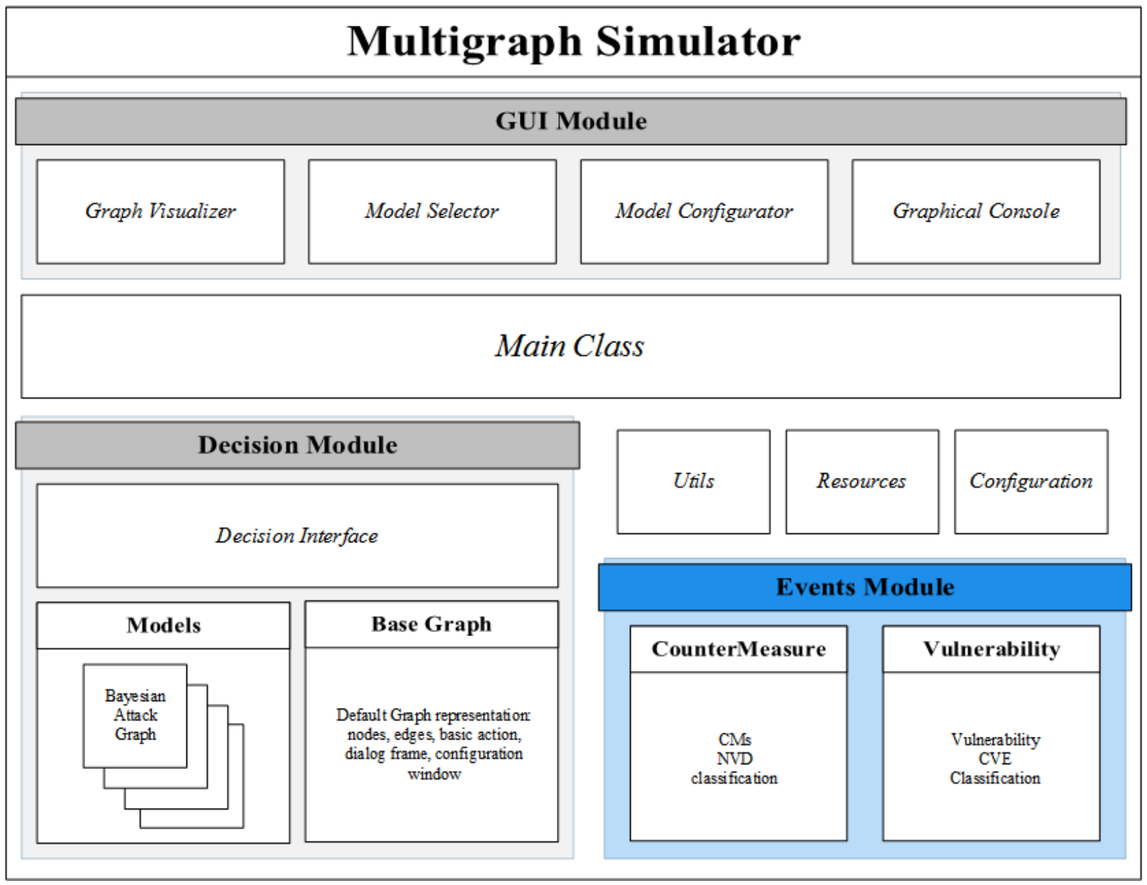
Implemented Models as Proof of Concept

✓ N. Poolsappasit, R. Dewri, I. Ray
Dynamic security risk management using Bayesian attack graphs
IEEE Transactions on Dependable and Secure Computing, 2012 – Thesis ref.: 23

✓ A. Roy, D.S. Kim, K.S. Trivedi.
Scalable optimal countermeasure selection using implicit enumeration on attack countermeasure trees
IEEE/IFIP International Conference on Dependable Systems and Networks, 2012.

Multigraph Simulator

Design, Objectives and Functionalities

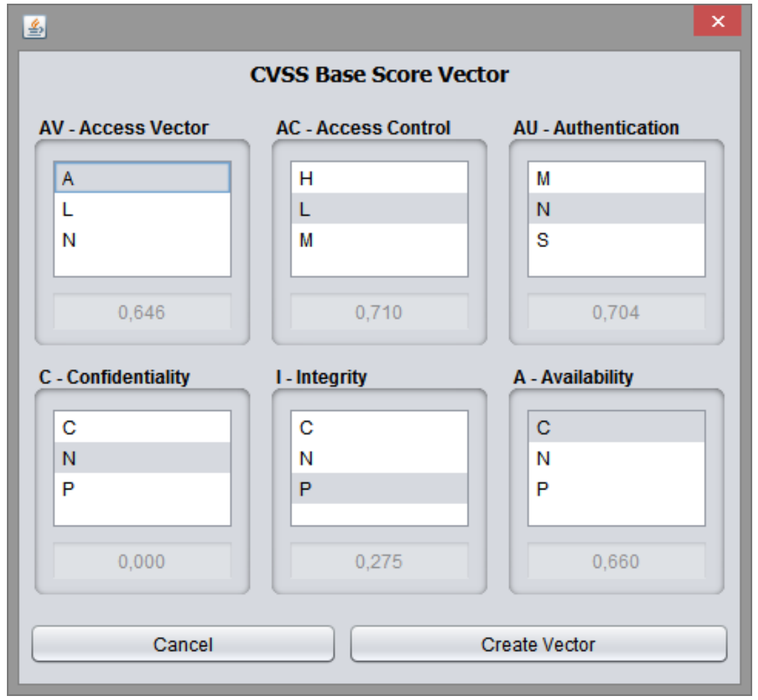


EVENT MODULE

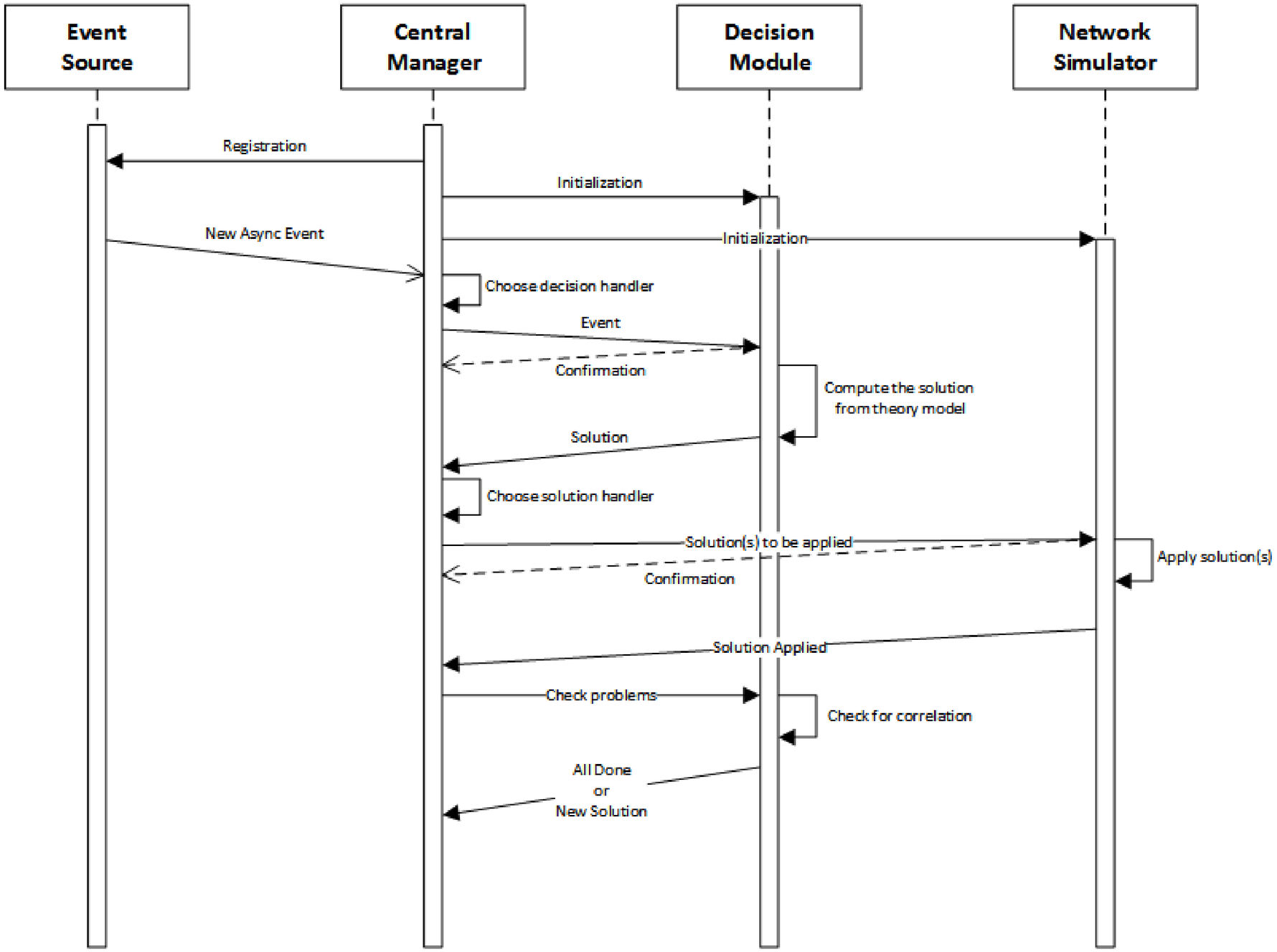
External interaction with the simulator

- Provides the standard metrics for the vulnerabilities and the countermeasures (CVSS, NVD classification, etc.).
- Will standardize the external input for the simulator (requires further study)
- Will allows the automatic analysis of the aggregated log received from the IDS/IPS.

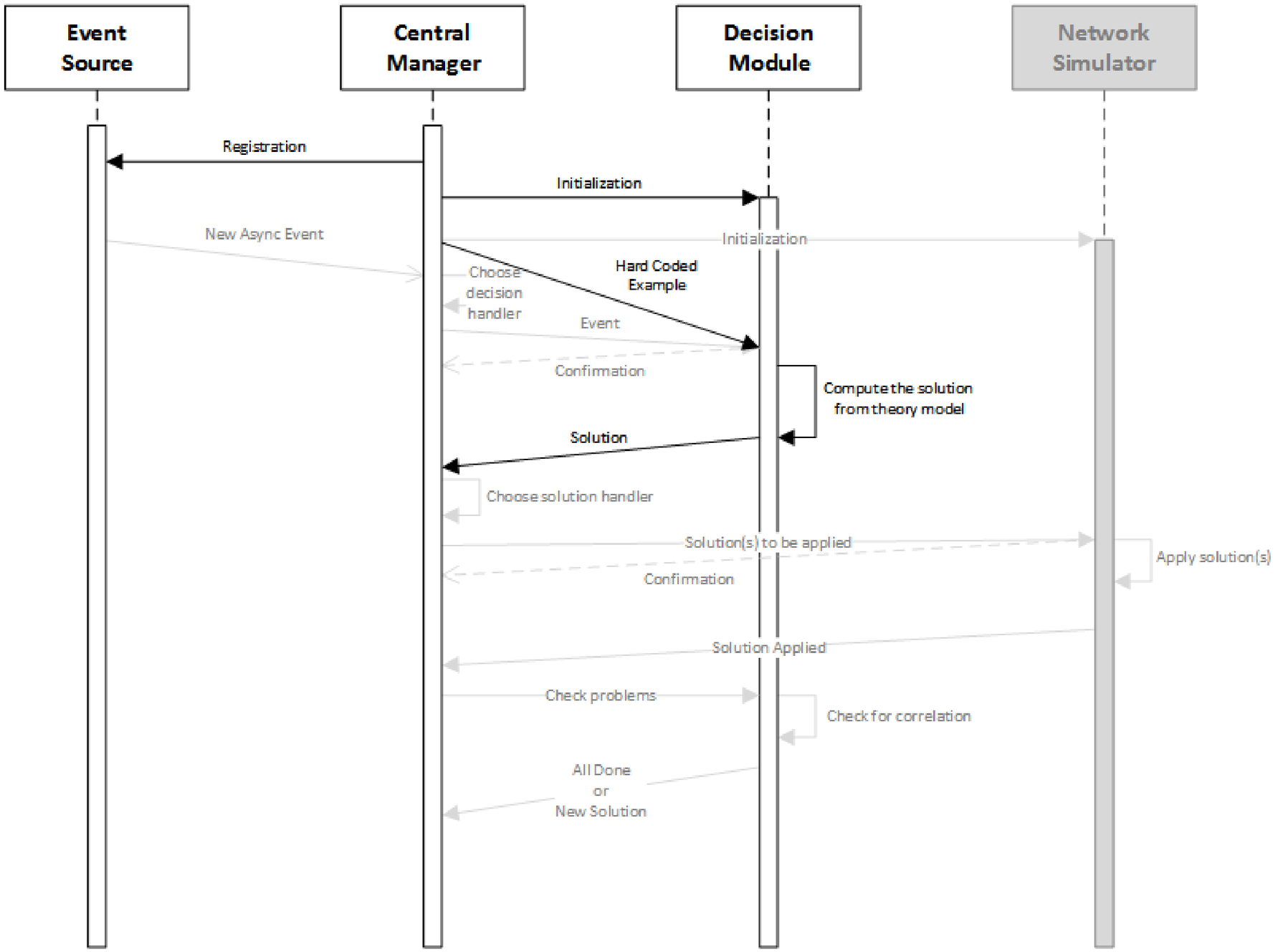
CVSS v2.0 Library



Control Flow – Goal



Control Flow – Actual Implementation



Conclusions

A short summary

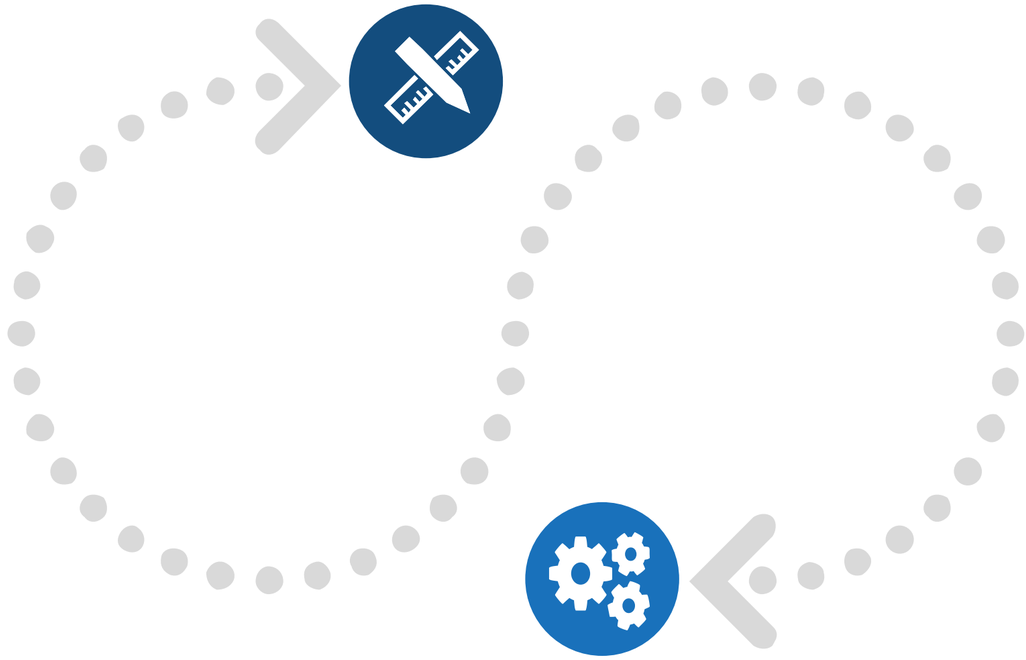
> “ Programmers are tools
> for converting caffeine into code. ”

What's next?

Future work and research interests

Complete the simulator…

- Define a (few) comparison metric(s)
- Develop the events' manager
- Analyze how the user can interact with the visualized graph
- Apply abstraction theory to the visualized graph
- Extend the generic model with temporal capabilities
- … and much more!



… and integrate it with commercial tools

In order to use it a real-world context Multigraph must be able to talk with the majority of commercial tools.
Integrating it with several IDPS components will allow the simulator to predict the next attack steps, while the coordination with a reaction module will grant to Multigraph the ability to fight off the intruders.

Conclusions

A short summary

Literature

This thesis provides a survey on the Bayesian-related models, showing that a standard structure or a well-defined methodology is required in order to compare different models.

That is mandatory when aiming to **combine different approaches for the best and most efficient solution to the generic problem.**

Multigraph

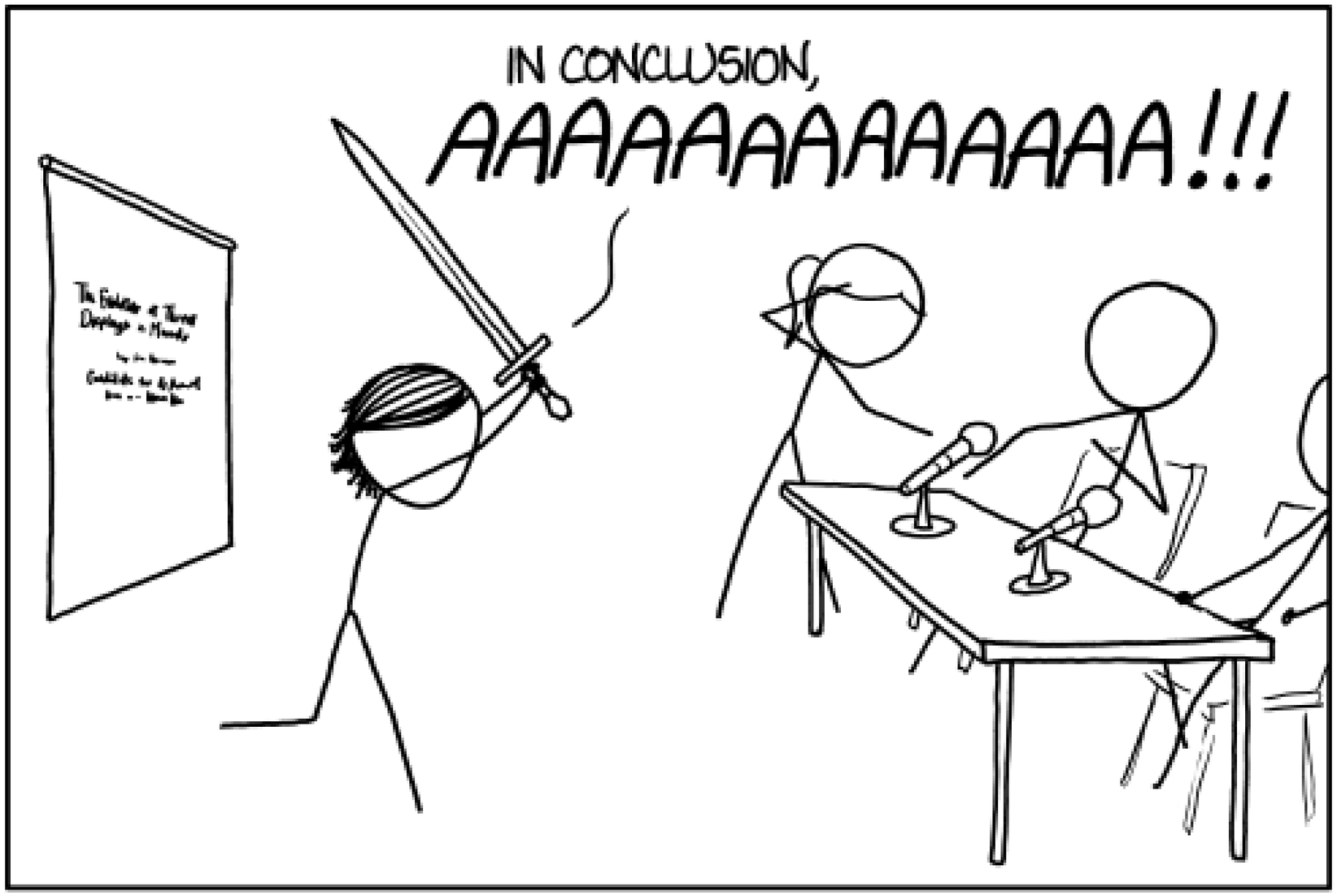
The simulator highlighted that most of the actual models are incomplete or not feasible in practice.

Further study are required in order to **define a common structure and a set of comparison metrics** in order to study different approaches' results, regardless that the simulator can help to develop new models in this field.

Achievement

This master thesis starts the cooperation between two universities boosting the research on the security fields, we are thinking that this joint work will lead to new exciting outcomes.

The simulator was also published during the first JNIC conference in Sep. 2015 showing a general interest on the approach.



THE BEST THESIS DEFENSE IS A GOOD THESIS OFFENSE.

Thanks for your attention!

MATTIA ZAGO

MSc Student in Cyber Security Engineering