DIRECTION

Project Goal: The project is a **group activity** that aims to create software that transforms an execution graph path (see Figure 1) in the MITRE ATT&CK into a Bayesian representation, incorporating the capacity to elicit and measure uncertainty in the new model representation.

Figure 1: Project Solution

Task Description:

- It is a group activity:
  - Two groups of **3 students**.
  - Two groups of **4 students**.
- The project must be saved in GitHub with a GPL or BSD license.
- The student must develop the solution and write a one or two-page position paper explaining your solution.
- The demo will be presented on the last day of class (Seminar).

Deliverables:

a. **17 JUN (1 point)**:
    a. The leader of Groups must submit the group members' names on the Blackboard.

b. **16/JUL (19 points)**:
    a. The student must submit the link to the GitHub project
    b. The project site must have:
        i. the license files.
        ii. source code
        iii. tutorial → how to install and use the library
    c. A late submission causes -10/100 points of the Project evaluation (MAXIMUM LATE IS 17/JUL 11:59 PM -> The syllabus policy is not applied for this task).

c. 17/JUL (60 points):
    a. Demo Session: each group will have 30 minutes during the session.
    b. In these 30 minutes, groups must:
        i. 5 minutes (PPT or demo): shows how to install, configure, and the software architecture of the tool
        ii. 5 minutes (PPT): Explains two scenarios where they are used to show the tool's functionality.
        iii. 20 minutes (demo): Run the tool, explain its use, and analyze the result.
    c. It is not accepted late submission.

d. 18/JUL (20 points):
    a. A position paper that describes why the solution is required, its architecture, the scenarios, and the scenario results using the tool.
    b. Maximum Size: 2 pages.

PROJECT DESCRIPTION

Security analysis is a far more imprecise process than deterministic reasoning. We do not know the attacker's choices; thus, there is uncertainty about unknown attacker behaviors[1]. Cyber-attacks are not always guaranteed to succeed; thus, there is uncertainty from the imperfect nature of exploits. The defender's observations on potential attack activities are limited, and as a result, we have uncertainty about the false positives and negatives of intrusion detection system (IDS) sensors. Nevertheless, the logical causality encoded in a deterministic attack graph is invaluable to understanding security events. It will help build practical network defense tools if we can appropriately account for the uncertainty inherent in the reasoning process.

Risk assessment, a crucial part of security analysis, is a systematic process for identifying, analyzing, and controlling potential hazards and risks. It's a powerful decision-making tool that guides us in determining which measures to implement to eliminate or control risks and which to prioritize.

However, any other security reasoning process could have these sources of uncertainties:

- Uncertainty in attack structure (kill chain)
- Uncertainty about the required preconditions to the attack (vulnerability and threat)
- Uncertainty about attacker actions (exploit)
- Uncertainty about the observations (alerts)
- Uncertainty about the consequences (impact)

The MITRE ATT&CK is a vital framework for risk assessment. MITRE ATT&CK® is a globally accessible knowledge base of adversary tactics and techniques based on real-world observations. The ATT&CK knowledge base is a foundation for developing specific threat models and methodologies in the private sector, government, and cybersecurity product and service community. It provides

[1] Xie, P., Li, J.H., Ou, X., Liu, P. and Levy, R., 2010, June. Using Bayesian networks for cyber security analysis. In *2010 IEEE/IFIP International Conference on Dependable Systems & Networks (DSN)* (pp. 211-220). IEEE.

information on the motivation, capabilities, interests, tactics, techniques, and procedures (TTPs) used by the threat actors.

Tactics represent the "why": the reason for acting. For example, an attacker may want to achieve credential access. Techniques represent the "how," e.g., the attacker may dump credentials to achieve credential access. Procedures are the specific steps an adversary uses to implement a technique or sub-technique to achieve their objective. Procedures differ from sub-techniques, are categorized by behavior, and can vary even when the underlying technique is the same. For example, an adversary might use various tools, skills, or resources.

Finally, MITRE ATT&CK combines the above three main components with the mitigations, representing the technologies that can prevent tactics and techniques.

When combined with the above information, it is easy to understand that MITRE ATT&CK perfectly matches all the phases of the risk assessment process[2], as demonstrated in Figure 1.

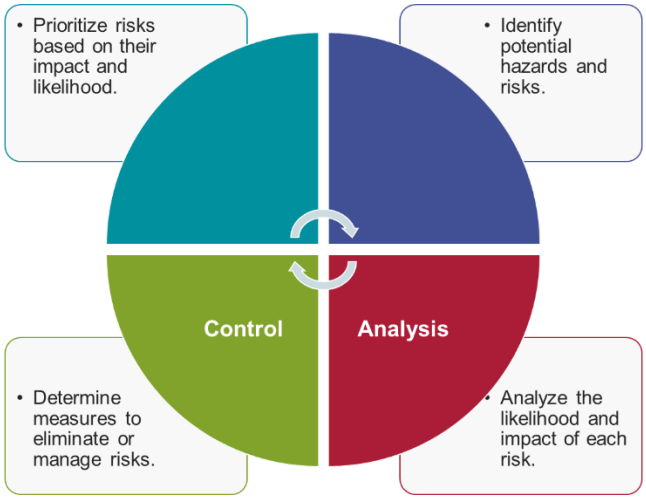

Figure 1 – Risk Assessment Phases

[2] Ahmed, M., Panda, S., Xenakis, C. and Panaousis, E., 2022, August. MITRE ATT&CK-driven cyber risk assessment. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (pp. 1-10).

The main problem with using MITRE ATT&CK is that it creates these maps using a deterministic approach, requiring it to be used in a risk assessment process to insert the uncertainty modeling in the tool.

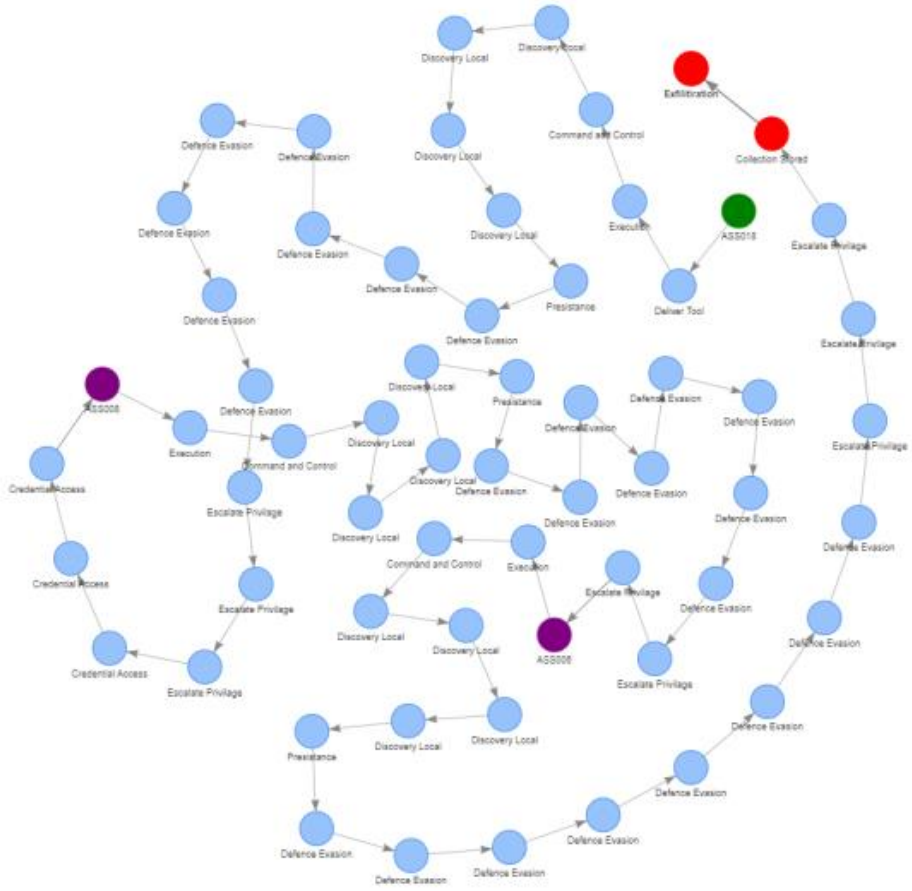

Figure 2: Kill Chain using MITRE ATT&CK

Project Goal: The project is a **group activity (3 students)**, which aims to create a library that transforms an execution graph (see Figure 2) in the MITRE ATT&CK into a Bayesian representation, incorporating the capacity to elicit and measure uncertainty into the model.

To perform the task, a data representation of the MITRE ATT&CK matrix must be connected to a Bayesian Network (BN) library. It is important to note that it is not a simple graph generation, given that a Bayesian Network is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed **acyclic graph** (DAG).

The solution must create (import) a kill chain based on the MITRE ATT&CK and produce a Bayesian representation of that kill chain. It is not required that the created BN analyze the type of variable to define the distribution type; you can simplify it using a normal distribution as a default. However, the BN must consider different types of relationships, like OR - AND gateways between each variable.

The end solution does not require a graphical interface; it can receive a file as input and generate a BN, similar to Figure 3.

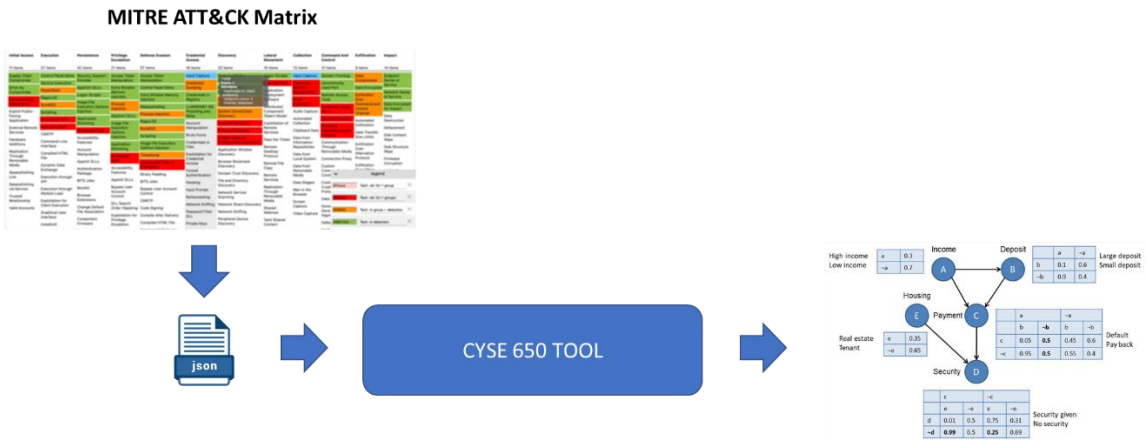

Figure 3: Project Solution

The student can select the libraries, languages, tools, and approaches to implement the solution. However, I suggest customizing the MITRE ATT&CK Navigator[3] to export a layer to the desired format to import into your toolbox.

You have two options for developing the BN. The first one, you can use the UNBBayes[4]. It is a Java tool and framework that you can use to develop simplistic and advanced BN. GMU uses it for many DARPA projects. Another approach is to use GeNie / SMILE[5]. It is another library that allows you to code in Java, Python, or C++.

[3] https://github.com/mitre-attack/attack-navigator/tree/master
[4] https://unbbayes.sourceforge.net/
[5] https://www.bayesfusion.com/genie/

The UNBBayes has a large number of samples, but the primary reference is the (Shou et al., 2011)[6] and its Wiki page[7] and how to use its API[8]. Also, the UNBBayes has a large number of publications explaining the different functionalities of the library[9].

GeNie is a graphical BN tool with an API (SMILE). It contains significant documentation, which you can find on its website: GeNie[10], C++ and MATLAB[11], and Java, Python, R and NET[12], as a page where you can find other information and several examples and source codes[13].

[6] Matsumoto, S., Carvalho, R.N., Ladeira, M., Costa, P.C., Santos, L.L., Silva, D., Onishi, M., Machado, E. and Cai, K., 2011. UnBBayes: a java framework for probabilistic models in AI. *Java in academia and research*, p.34.
[7] https://sourceforge.net/p/unbbayes/wiki/browse_pages/
[8] https://sourceforge.net/p/unbbayes/wiki/Using%20UnBBayes%20as%20API/
[9] https://unbbayes.sourceforge.net/publications.html
[10] https://support.bayesfusion.com/docs/GeNIe/
[11] https://support.bayesfusion.com/docs/SMILE/
[12] https://support.bayesfusion.com/docs/Wrappers/
[13] https://support.bayesfusion.com/docs/

CYSE-650 - CYBER RISK MODELING AND ANALYSIS TOOLS

RUBRIC

TASK 1 – Project Files (17 JUN)		
Grade Component	**Poor -10 points**	**Satisfactory 1 point**
Submission of the group list	No submission or late submission.	Submit the group list on the BB until 17 JUN 11:59 pm.

TASK 2 – Repository (16 JUL)			
Grade Component	**Needs improvement. - 10 points**	**Satisfactory 2 points**	**Good 9.5 points**
GitHub with License file	No submission or Late submission MAXIMUM LATE IS 17/JUL 11:59 PM -> The syllabus policy is not applied for this task.	The repository has the license file.	
Source Code		The source code is in the folder.	The source code is complete and contains all dependencies or scripts to compile (ex., maven, NPM, etc.)
Tutorial: how to install and use the library		The tutorial contains the installation procedure and how to use the library.	The installation procedure and how to use the library works without any errors.

CYSE-650 - CYBER RISK MODELING AND ANALYSIS TOOLS

TASK 3 – Demo Session (17/JUL) IT IS NOT ACCEPTED FOR LATE SUBMISSION.				
Grade Component	**Poor 0 points**	**Needs improvement. 5 points**	**Satisfactory 10 points**	**Excellent 25 points**
Shows how to install and configure the software architecture of the tool. **Time: 5 minutes**	Not present or participate in the demo session.	The presentation is unclear, does not show all the steps involved in the installation and configuration, and does not discuss required software dependencies.	The presentation clearly shows all the steps involved in the installation and configuration, including discussing any required software dependencies.	[redacted]
Explains two scenarios where they are used to show the tool's functionality. **Time: 5 minutes**		The student presents only one scenario.	The students present two scenarios; however, the presentation is unclear and does not explain the importance of measuring uncertainty and why it is required.	The students presented two scenarios and did a good presentation.
Run the tool, explain its use, and analyze the result. **Time: 20 minutes**		The student did not make a live demo.	The demo ran with some bugs, or the students did not explain and demonstrate how to use the result.	The demo works without bugs, and the students explain and demonstrate how to use the results.

CYSE-650 - CYBER RISK MODELING AND ANALYSIS TOOLS

TASK 4 – Position Paper (18/JUL)				
Grade Component	**Poor 0 points**	**Needs improvement. 1 point**	**Satisfactory 3 points**	**Excellent 4 points**
Compliance with the template provided.	No submission	The paper does not comply with the template provided.	The abstract contains more than 200 words or less than 80 words. OR The abstract fails in one of these features: a) The abstract lets readers get the gist or essence of your paper or article quickly to decide whether to read the full paper; b) the abstract prepares readers to follow the paper's detailed information, analyses, and arguments; or c) the abstract helps readers remember key points from your paper.	The paper complies with the template provided. The paper includes a bibliography section. The paper has one full page and no more than two pages. The abstract achieves all the requirements presented in the previous column.
Introduction Section		The paper does not introduce the problem.	The paper motivates the reader to the problem but does not show the problem solved by the tools.	The paper motivates the reader to the problem and shows the problem solved by the tools.
Software Architecture		The paper does not explain the architecture that was developed.	The paper presents, but does not explain, the architecture that was developed,	The paper explains the architecture that was developed.
Scenario		The paper does not present at least one scenario.	The paper contains at least one scenario; however, it does not explain, or the explanation is weak.	The paper comprises at least one scenario and explains how it relates to the problem it handles.
Results		The paper does not show how the tools could be used to generate or explain the results.	The paper shows how the tools generate the results; however, it does not explain the results.	The paper shows how the tools are used to generate the results and explains its results.