Midterm Report:
Hybrid Malware Detection based on "API call sequence" and "network traffic"

Notice

- Must do all the experiments in the Virtual Machine.
- Never use our school’s internet (or you’ll receive a virus incident report from the school).
- Be Careful.

Step 1. Malware Detection based on "API call sequence"

- You could read the below paper:
  - Mingdong Tang, and Quan Qian, "Dynamic API call sequence visualization for malware classification," IET Information Security 13.4 (2019): 367-377.

  - Collect samples from VirusShare: https://virusshare.com/ , or the other malware datasets
  - Collect the information of API call sequences from Cuckoo: https://github.com/cuckoosandbox/cuckoo
  - You can read this paper; however, for this midterm, you are not restricted to using its malware image generation method.

Step 2. Malware Detection based on "network traffic"

- You could read the below paper:
  - W. Wang, et al. "Malware traffic classification using convolutional neural network for representation learning." 2017 International conference on information networking (ICOIN). IEEE, 2017.

  - Use the same samples of Step 1.
  - For these malware samples, collect their network packets (.pcap files) from Cuckoo, during the same time as Step 1.
    - This means you collect the API call sequence and network packets from the same malware while it is running at the same time.
  - You can read this paper; however, for this midterm, you are not restricted to using its malware image generation method.

Step 3. Hybrid analysis based on the above two features

- Combine the features of "API call sequence" and "network traffic", then make the hybrid malware classification.

- The combination method is not restricted.

In fact, this should be where you demonstrate your novelty.

For example, you could merge two feature images into one, or create a 3-dimensional vector containing "API call sequence," "network traffic," and "time." Then, perform a hybrid analysis on this feature space.

(Optional) Few-shot Learning

- Apply a Few-shot Learning scheme on your above system.

- If you do this, please mention it in your report.

You could earn an additional **10 points**.

Upload the result to Moodle

- Upload your code and the ***detailed*** reports to Moodle.
  - **Code**: You must add some comments in your code for easy understanding.
  - **Report Files**: Document (Word) and Presentation (PPT).
    - Including a table in your report outlining the responsibilities of each team member for this midterm.
  - Zipped to a file. The file name should be “Team Name_Midterm.zip”