Quantitative Comparisons of MITRE ATT&CK Campaigns

An implementation for the paper Nicoletti et al. (to appear), detailing a method to perform quantitative comparisons of MITRE ATT&CK Campaigns.

How to Use

Requirements

System packages required (tested on Python versions 3.10, 3.11, and 3.12):

- `jupyter`: the python package to execute jupyter notebooks
- `pygraphviz` >= 1.13

> **NOTE**:
> In Debian-based Linux distros, the above corresponds to the following `apt` packages: `jupyter-notebook`, `python3-dev`, `python3-pygraphviz`.
> In Arch linux, the above corresponds to the following `pacman` packages: `jupyter-notebook`, `python-pygraphviz`

To create (template) ATs and compute metrics, additional packages are required. You can install them via the requirements.txt file in the bin/ folder, either with:

```
$ cd bin
$ python3 -m pip install -r requirements.txt
```

or

```
$ cd bin
$ mkdir .venv
$ pipenv install
```

These commands will install the following packages:

- `ply` >= 3.10
- `pandas` >= 2.2.0
- `astutils` >= 0.0.5
- `networkx` >= 3.2.1
- `matplotlib` >= 3.8.2
- `openpyxl` >= 3.1.5
- `odfpy` >= 1.4.1

Additionally, `ipykernel` might be needed to run Jupyter Notebooks inside VS Code: we used `ipykernel` 8.26.0.

MITRE Data and Counting

MITRE data needed to perform computations is stored in the bin/data/ folder.

The implementation manipulating MITRE data can be inspected in the two provided Python Jupyter Notebooks: bin/techniques_frequency_per_tactic.ipynb, and also bin/techniques_frequency_overall.ipynb. To replicate our results one can execute these notebooks—an example of how to do this is shown below. Note that pre-computed outputs are already stored in the cells of the notebooks, so it is not necessary to execute them in order to see these results. New outputs of this intermediate process are stored in the bin/data/ folder, which also contains the output `.json` and `.csv` files generated during our experimentation.

> **Example execution of jupyter notebook**
>
> 1. Run jupyter via the command `jupyter notebook`, or `pipenv run jupyter notebook`, in your shell
> 2. This opens a web browser tab with jupyter: double-click the notebook `techniques_frequency_per_tactic.ipynb`
> 3. In the top menu of the notebook, select `Kernel` → `Restart Kernel and Run All Cells...`
> 4. When prompted, click the `Restart` button
> 5. Then all cells of the notebook will be (re-) executed, and you will be able to observe the progress live as it executes.
> 6. Results external to the cells will be stored in the results/ folder—see the *Results* section at the bottom of this README.

Generating MITRE ATT&CK Templates and Computing Metrics

There are three entry points of this program. To reproduce custom-made AT templates for C0014 and C0022 — and compute metrics on them — you can run the code (as shown above) in the respective Python Jupyter Notebooks stored in the bin/ folder: C0014_Wocao_custom_full_AT.ipynb and C0022_DreamJob_custom_full_AT.ipynb.

To automatically generate MITRE ATT&CK Templates (MATTs) and compute metrics on them, you can run the bin/create_MITRE_AT_templates.py:

```
$ cd bin
$ python3 create_MITRE_AT_templates.py
```

This script generates and computes the *hard*, *default*, and *easy* templates for C0014 (campaign 14, codename *Wocao*) and C0022 (*Dream Job*) by default. This behaviour can be altered by inserting MITRE campaign codes for desired campaigns in the list contained in the CAMPAIGNS variable in bin/create_MITRE_AT_templates.py, e.g.:

```
CAMPAIGNS = [
    "C0014",  # generate the MATT for the Wocao     campaign
```

```
    "C0022",  # generate the MATT for the Dream Job campaign
    "C0018",
#    "",  # generate the complete MATT (all techniques and tactics)
]
```

Results

Results are stored in the results/ folder.

results/results_custom/ contains pre-computed results for the custom-made ATs of C0014 and C0022: these are `.out` files containing probability values for the respective top events, and `.pdf` files representing these custom ATs graphically.

results/results_MATT/ contains pre-computed results for the automatically-generated template ATs of C0014 and C0022, in the *hard*, *default* and *easy* variants: these are `.out` (and `.err`) files containing probability values for the respective template top events and `.pdf` files representing these templated ATs.

Results contained in the two aforementioned folders are used to generate Table 4 in Nicoletti et al. (to appear).

When re-running scripts as described in the previous section of this README, new folders will be created in results/.