PMR Insight Collective Knowledge (PICK)

Installation and Usage

Version 1.0

5/8/2020

**Installation**

To install the PICK Tool, you will need to:

* **Clone the Team11 PICK Tool GitHub repository, available on** [**GitHub**](https://docs.google.com/document/d/1p2y2RnjkFGfOTF_jg7ppoWAL4M9VuZKAGXvmQ44ZhKA/edit?usp=sharing)**.**
  + <https://github.com/CS4311-spring-2020/pick-tool-team11-binarybeasts>
* **Download dependencies the tool requires to run. This includes:**
* Python 3 (version 3.7.0 or newer)
  + Python is used to create all the backend of the PICK tool.
  + You can download the latest version of Python at the [Python website](https://www.python.org/downloads/).

* + <https://www.python.org/downloads/>
* PyQt5 (version 5.14.1)
  + This tool is used to create the GUI interface.
  + You can install this by running pip install pyqt5 or by running pip3 install pyqt5, depending on your environment. (Mac or Windows)
  + In Kali Linux, run: sudo apt-get install python-pyqt5
* MongoDB (version 4.2.6)
  + This tool is used to store important data generated by the PICK tool.
  + You can install MongoDB from the [MongoDB website](https://docs.mongodb.com/manual/installation/).

* + <https://medium.com/cyber4people/setup-mongodb-in-kali-linux-3ab86731e3ec>
* pyMongo (version 3.10.1)
  + This library is used to interact with the MongoDB database through Python.
  + You can install this by running pip install pymongo or by running pip3 install pymongo, depending on your environment. (Mac or Windows)
  + In Kali Linux, run sudo python -m easy\_install pymongo-3.6-py2.7-linux-x86\_64.egg
* pandas (version 1.0.3)
  + This tool is used as an aid for cleansing the log entries.
  + You can install this by running pip install pandas or pip3 install pandas, depending on your environment. (Mac or Windows)
  + In Kali Linux, run: sudo apt-get install python-pandas
* Splunk-SDK (version 1.6.12)
  + This tool is used to import all the log files to the PICK tool and obtain the log entries.
  + You can download Splunk Enterprise from the [Splunk website](https://www.splunk.com).

* + <https://www.splunk.com>
* QGraphViz (version 0.13.2)
  + This tool is used for creating the graph.
  + You can install this by running pip install QGraphViz or by running pip3 install QGraphViz, depending on your environment. (Mac or Windows)
  + In Kali Linux, run: sudo apt-get install graphviz
* Pytesseract (version 0.3.4)
  + You can install this by running pip install pytesseract or pip3 install pytesseract.(Mac or Windows)
  + In Kali Linux, run: sudo apt-get install tesseract-ocr

* + <https://tesseract-ocr.github.io/tessdoc/Downloads>

**Usage**

**Before you run the tool:**

* Be sure to install the dependencies and libraries described in the previous section.
* Run your local Splunk instance, and create an index called “splunk\_test”. Keep the tool open when running the PICK Tool.
* You will need to edit the hardcoded Splunk login info to match your credentials in the file pick-tool-team11-binarybeasts/src/ingestion/splunk.py
* Run MongoDB on your machine.

**Running the tool:**

* In a command line window, navigate to the cloned repository pick-tool-team11-binarybeasts/src.
* In the same command line window, you’ll need to run python -m ui.windows.main or python3 -m ui.windows.main, depending on your environment.
* After running this command, the tool should run and display the main window.
  + If the main window doesn’t pop up, make sure you have installed all dependencies and followed the steps in “before you run the tool.”

**Set up your configurations:**

* To open up the configurations window, click the **Event Configuration** button from the main window.
* From this window’s tabs, you can:
  + Edit your user settings in **Team Configuration** to specify the **lead’s IP address**, and whether or not you are a **lead user**. To save the info, click Connect.
  + Edit the current event’s **name**, **description**, **start time**, and **end time**. To save the info, click Save Event.
  + Edit the **directories** of the log file locations by either typing in the fully qualified path of each directory, or by using the directory picker button to the right of each text field. To upload the files in the given directories to Splunk, click on Start Data Ingestion.
  + Add any number of **vectors** with names and descriptions. Add a vector by typing in the info and clicking Save Vector. To delete a Vector, click the vector you want to delete, then click Delete Vector.
  + Add any number of **icons** with names and source file locations. The file location can be entered either by typing in the fully qualified path of each directory, or by using the directory picker button. Add an icon by typing in the info and clicking Add Icon. To delete an Icon, click the Icon you want to delete, then click Delete Icon.
* The information you enter into this window is data persistent (i.e., it will still be present if you close and open the tool again. Just be sure to confirm your entries and selection on each window before closing).

**The Enforcement Action Report:**

* To open up the search and filter window, click on the **Enforcement Action Report** button from the main window.
* Our Enforcement Action Report is able to pull log entries from Splunk, but does not have any of the cleansing or validation stuff implemented into it.
  + On our **cleansing.py**, we have the code that is able to cleanse log files by removing blank rows.
    - First you would need to go into the code and input the location path you would like to have your saved file as shown below in ***Line 32 of cleansing.py***:
      * data.to\_csv(r'C:\**locationPath**\Cleansed.csv',index=False)
    - After specifying your location path, run this code separately by using the following command: cleansing.py “filename.csv”
  + On our **ocr\_feeder.py**, we have the code that is able to extract text from images and save it as a text file. From that text file, the code will then convert it into a CSV file which can then be ingested into Splunk.
    - First, in ***Line 10 of ocr\_feeder.py***, specify the location path for your tesseract executable as shown below:
      * pytesseract.pytesseract.tesseract\_cmd = r'C:\**locationPath**\tesseract.exe'
    - Secondly, you would need to go into the code and input the location path you would like to have your saved file as shown below in ***Line 23 of ocr\_feeder.py***:
      * data.to\_csv(r'C:\**locationPath**\Cleansed.csv',index=False)
    - Run this code separately by using the following command: ocr\_feeder.py “imageName.png”
* Our **cleansing.py** and **ocr\_feeder.py** are not yet connected into the PICK tool, nor is there validation going on, but they do function separately.

**Searching, Filtering, and Associating Splunk Log entries to Vectors:**

* To open up the search and filter window, click on the **Search/Filter** button from the main window.
  + When you click the Search/Filter button, a popup window will display explaining that log entries might take a little while to import from Splunk. This window is a standard warning window. Please be patient, especially if you have imported several log files or large files into Splunk. Log entries will display once the log entries are imported from Splunk.
* A full, comprehensive list of **log entries** uploaded to Splunk from the tool will display on the top right hand side of the window.
* To mark a log entry as significant by associating it to a vector, you can select the **vector** you want to associate to from the dropdown menu at the bottom half of the window, then click Select Vector. Then, click the log entry you want to associate to it, then click Associate to Vector.
  + The newly associated vector will be immediately displayed in the vector’s list of associated log entries.
* While associating log entries to vectors, you can quickly switch between vectors by selecting the vector from the dropdown menu from the bottom half of the window and clicking Select Vector. When switching between vectors, the list of that vector’s log entries will immediately display after clicking Select Vector.
* Searching and Filtering Splunk log entries is not functional at this time.

**Creating a graph:**

* To open up the search and filter window, click on the **Manage Graph** button from the main window.
* From this window you can:
  + Select the vector you wish to work with, based on the vectors created on the vector configuration tab from the **Event Configuration** window
  + Add a new node by clicking on the “Add Node” button and fill out the dialog box.
    - You can select a circle/square or an icon to represent the node in the graph
  + Once you have added at least two nodes on the graph, the user can create a new relationship between two nodes:
    - This can be done in the relationship table, by typing the parent node name and typing the child node name
    - Or by clicking in the “Add relationship” button on the graph view
      * and dragging a line between the parent and child node on the graph.
      * Every time the user creates a new relationship the user has to click on the “Add relationship” button on the graph view.
  + Delete an existing relationship by clicking on the “Rem Relationship” button and select the relationship you wish to delete.
  + Delete an existing node by clicking on the “Rem Node” button and select the node you wish you to remove from the graph
  + Export the graph by clicking on the “Export” button and clicking on the “Save” button once you have chosen the place in your machine to save it. This will automatically download and save the graph as a CSV file in the location chosen on your machine.
  + Open an existing graph by clicking on the “Open” button and selecting the desired CSV file.
  + Create a new graph and clear all the existing data on the graph window by clicking on the “New” button.
  + Manipulate the graph by clicking on the “Manipulate” button and selecting the nodes and moving them around the graph to a desired space.