10–28–14 CSE 4243/6243 Homework 4 - Extending the Filter

Introduction

In this lab, you will extend the filter you developed in the previous assignment to work properly with fully qualified domain names (FQDNs). Your filter will now add and remove hosts in IP address or FQDN form. The filtering code itself will attempt to prevent evasion by providing an IP address of a host instead of a blocked FQDN (or vice versa).

Requirements

- You are required to use your code from Homework 3, and it should still function correctly with regards to Homework 3's specifications. You may make any changes necessary to meet those specifications and any new specifications in this homework assignment.
- `add_host` and `remove_host` should now accept a string that either contains an IP address or a FQDN
  - Note: It may or may not be necessary for you to make any changes to these functions, depending on how you implemented your module).
- `get_blacklist` should still retain its Homework 3 functionality, though the list may now be mixed with some strings representing IP addresses, and others representing FQDNs
  - Note: Like the add and remove functions, it may not be necessary for you to make any changes.
- `get_web_index` should now behave as follows:
  - Accepts a string consisting of an IP address or a FQDN.
  - This function will check the blacklist to see if the host provided is blocked. If the host is blocked, then it should raise a `StandardError` exception describing the issue.
    - The requested host should also be considered blocked if:
      - It is a FQDN and it resolves to an IP address that is on the blacklist
      - It is an IP address that reverse lookup shows as being associated with a blacklisted FQDN
    - HINT: The Python `socket` library has functions that can help you with resolving names to IP addresses and vice versa.
    - Note: Reverse lookups of names can be tricky. A given IP addresses doesn't always resolve back to a domain name that resolves in the forward direction to that IP. There can be many-to-one relationships between multiple domain names resolving to a single IP address. While I do expect you to reliably block names that resolve to blocked IP addresses, I will not be testing you on unusual reverse lookup behavior. Test it against IP addresses and names that consistently resolve in both directions.
  - If the host provided is not blocked, then this function should issue a request for the web page located at `http://<host>/`
  - Returns HTML text of the page at the web root of that host.

    - If there is no web page at that location, or some other error occurs while issuing the request, the string `"No Page"` should be returned. No exception should be raised back to the calling function in this case.
    - If the request is successful, and a page is retrieved, a string representing the text of the HTML of that page should be returned to the calling function.
- Your deliverables include:
  - The code for your module, meeting the above requirements
  - Code for the Python program you used to test for correct operation of the module
  - A professional-looking report that includes a discussion of your design and implementation of the module, and documentation of your testing.
    - Your testing should verify all operations, not just what has changed since the last assignment, since you may have introduced new problems in old functionality.
- Your submission is due by Friday, November 7th, 5:00PM

Academic Integrity

Homework assignments in this class are strictly individual work. Do not discuss any aspect of this assignment with fellow students. Assignments will be checked for evidence of plagiarism. Use and cite resources appropriately.

You are only allowed to refer to the official Python documentation available at **https://www.python.org/doc/** for assistance in completing this assignment. The Tutorial, Library Reference, and Language Reference available there provide sufficient documentation of everything needed to complete the work. If it is determined that your code is based on or written while consulting another resource (such as, but not limited to, StackExchange), you will receive a zero for the assignment, even if it is cited correctly.

Grading *(points for not following instructions taken off of the grade determined by this scheme)*

- 10% - Spelling, grammar, appropriate technical writing style. Avoid conversational/informal tone.
  - All students are encouraged to seek (and acknowledge) writing and proofreading assistance from one of the university's writing centers.
- 10% - Professional formatting, organization of your document
- 60% - Meeting the requirements for the required Python module
- 20% - Documentation of testing of the module's functionality