R · I · T

Rochester Institute of Technology
Golisano College of Computing and Information Sciences
Dept. of Networking, Security, & Systems Administration

4055-760 Computer Viruses and Malicious Software

First Generation Virus Scanner Project

General requirements

The goal of this project is to write a first generation virus scanner with your favorite programming language such as C, C++, Java, VB, etc. Read Chapter 11 of the book “The Art of Computer Virus Research and Defense” by Peter Szor for a detailed description of the first generation virus scanner.

Here are some the basic requirements and specifications.

1. There are two inputs to the program. One is a directory; the other is a file containing a list of virus signatures, i.e., binary strings.
2. The program should examine recursively all files and subdirectories in the directory specified.
3. For each files, the program will test if it contains a binary substring than matches a signature listed in the input file.
4. The output of the program is a report that indicates which files in the directory contain which signatures in the list.
5. You have the freedom to design the format of the signature file for fast performance. The simplest format can be just one line containing one binary string. If you design your own file format, a tool should be provided to convert the simplest format to your faster format.
6. Any string matching algorithm is acceptable. Build-in library functions are allowed. However, faster string matching algorithms such as Bayer-Moore algorithm will increase the performance of the scanner. Bonus points will be awarded for faster implementations. The original article by Bayer and Moore is provided.

Deliverables

You need to submit all your deliverables in a single zip file to the drop box specified on mycourses. Your deliverables includes at least these items.

1. Source code of your program
2. Compiled executable on either Window XP or Linux platform
3. A write-up containing how to run your program, your program design, flow charts, file formats, etc.

Final Note

This is an individual project. Although discussions about the project among students are allowed, you should complete your program independently. Coping parts of source code of any sizes is not allowed.