Assignment 4

ITE 423 – Information Security

DUE: June 7. *Submit in Class*

Programming Assignment (version 1)

You will create a program (or script) that read a sequence of numbers and draw several analysis graphs. The given number represents the abnormal security events (daily) over the last 10 years (2002-2012) in the Internet link. Such events represent the Internet Worms, DDoS attacks, and etc. Your job is to create a set of analysis graphs to observe any irregularity within its occurrence.

Your graphs should be the following:

1. (40%) Time-series dotted graph (x-axis: days, y-axis: daily counts of abnormal event)

2. (60%) Cumulative Distribution Function (CDF) graph of daily counts (x-axis: daily counts, y-axis: ratios)

3. (200%) Big Bonus: Discrete Fourier Transforms (FFT) graph on frequency analysis. It measures frequency (pattern) in the given data.

You have a few choices to do this. However, the following is my recommendation:

> A. First, you will need to install NumPy and SciPy (http://www.numpy.org/) to read the data file.
>
> B. To analyze and draw figures, you may use PYTHON with matplotlib (http://matplotlib.org/), octave (http://www.gnu.org/software/octave/), gnuplot (http://www.gnuplot.info/), FFT (http://docs.scipy.org/doc/scipy/reference/fftpack.html), or any other choice that suits you the best.

The data file is given and the sequence of numbers written in the numpy format: "abnormal.npy". It contains an array of numbers in [0 1 2 3 4 …]. You may use "numpy.load(filename)" to load the data in your python program.

Create your program called: "abnormal-analysis.py". Your program should draw the above figures accordingly (PDF is recommended).

SUBMISSION:

1. Print out your source code – "abnormal-analysis.py"
2. Print out all figures that you create.
3. Print one paragraph explaining your program and any difficulties you had. (No less than 250 words)

NOTE:

1. Python is recommended. If you are willing to use other programming languages, that's fine too. But use it at your own risk.
2. Command Line base is recommended.
3. Any platform of your choice is fine, but I'd like to see many Linux as possible (e.g., Ubuntu).

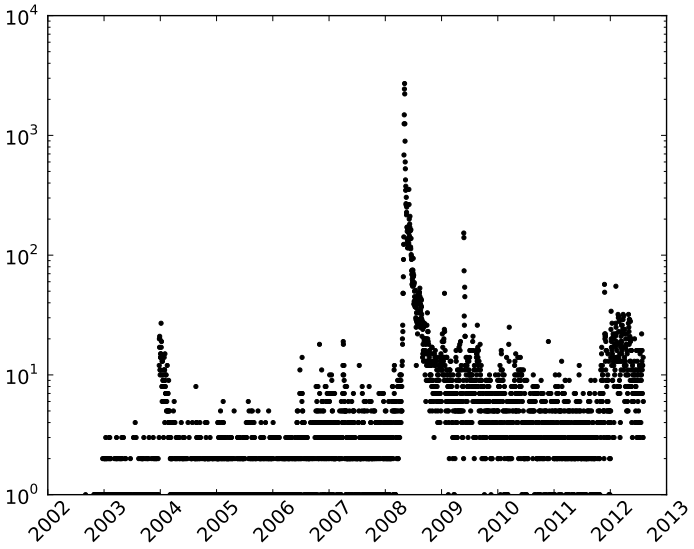
EXAMPLE:

>>python abnormal-analysis.py abnormal.npy

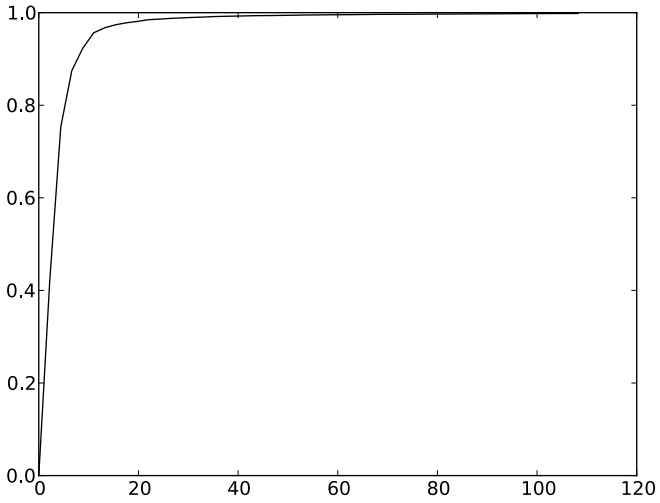
timeseries.pdf is created.
cdf.pdf is created.
fft.pdf is created.

EXAMPLE (graphs):

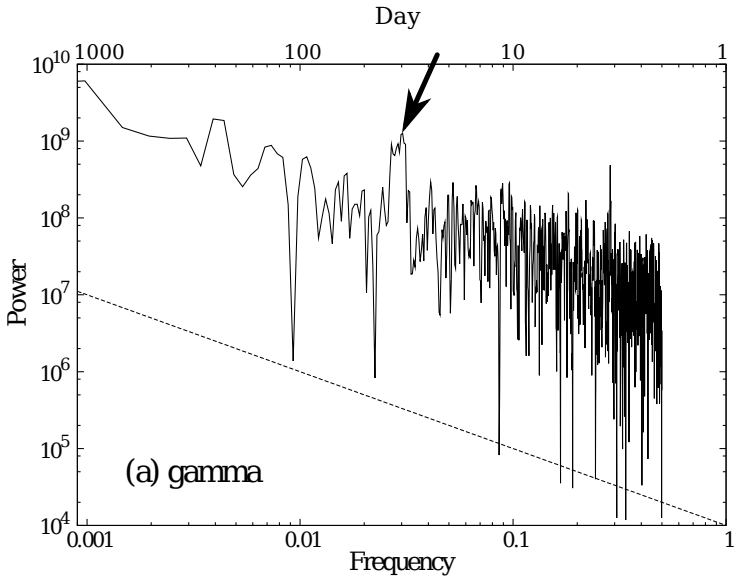
1. timeseries.pdf example:



2. cdf.pdf example:



3. fft.pdf example:



IMPORTANT:

Do your best. Even if you can't do the whole assignment, submit as much as you can (with explanation why you can't do this). If you get caught cheating, I will deduct your other assignment marks. And to make sure that you do your own assignment, I will randomly select a few students in class and ask them to explain their code. The BIG Bonus is really BIG.

REFERENCE: python

For download and installation,

http://www.python.org/download/

For tutorial,

http://www.learnpython.org/ http://docs.python.org/tutorial/

TRIVIA:

Google is your friend and teacher. Search!
Discuss with your classmates! (DO NOT send me an email first)
I will go over the assignment in class, so don't worry too much.