Project 1 – Csci 244 Operating Systems

Tracking "abnormal" behavior in processes

Goal

The goal of this project is to monitor the evolution of the characteristics of processes over time.

Rationale

The main program will create processes. In these processes, some functions will be performed, and the choice of these functions will be unknown from the main program that did create the processes. The goal of the main program is to know when the type of processes that are being run is changing, without knowing what is running.

The story

The parent will create a child. The child will run, finish. Then, the parent is going to recreate a child process, and so on…

Each child process will run a function of your choice in which the main parameter will be taken from a normal distribution defined by its mean and standard deviation. This parameter can correspond to the amount of the **memory** used by the process.

During the 500 first child processes, the parameters will be taken from a distribution D1 while after for the next 500 child processes, it will be taken from the distribution D2.

The parent process will retrieve the amount of memory allocated for the different child processes using appropriate functions (it depends on the OS of your choice).

After the data obtained from the first 250 child processes, create a one-class classifier using a method of your choice (e.g. Gaussian model, Nearest neighbor,…). These algorithms will be covered during the class.

For the next child processes, determine if a child process is coming from the distribution D1 or D2.

At the end, report the confusion matrix corresponding to the number of child processes belonging to D1 and detected as such, those detected belonging to D2; the number of child processes belonging to D2 and detected as such, those detected belonging to D1.

You can use any Operating System you want (Windows, Linux,…)

Deliverables

- The code
  - It has comments.
  - It is elegant.
  - It compiles.
  - It runs.
  - It displays appropriate messages in the console to determine what is happening.
- Source
  - The main process function
  - The child process function
  - The function to generate the values of the amount of memory allocated by the child process in relation to a normal distribution defined by its mean and standard deviation.
  - The function to retrieve the amount of memory used by a child process
  - The function to create the one-class classifier model
  - The function to use (test) the one-class model
- Short report
  - Short introduction about the goal of the project
  - Brief description of the method used to track the child processes.
  - A table corresponding to the results from experiments using different distribution.
    - Try different distributions, i.e. different differences between the mean and standard deviation of these distributions.
    - Plot the distributions obtained from 1 to 250, 251 to 500, and 501 to 1000, using histograms.
  - Discussion about the results
  - Personal reflection about novelty detection for tracking process behavior