Lab2 – Queueing Theory and Its Applications

Objective: Through this lab exercise, students should learn: (1) how to apply queueing theory to model a simple web application system; (2) how to use fundamental laws to investigate the relations of system quantities and gain insights about the system performance; (3) how to design and implement a simulation study based on the queueing model; and (4) how to analyze simulation results statistically and comparing them to analytical results derived from queueing theory.

Part A: Consider a system modeled as shown in the figure below. A web request send to web server must wait in queue if the server is busy, otherwise it will begin the processing at the server. When the server finishes processing the request, it sends a response to the user. A user will think some times before generating next request.

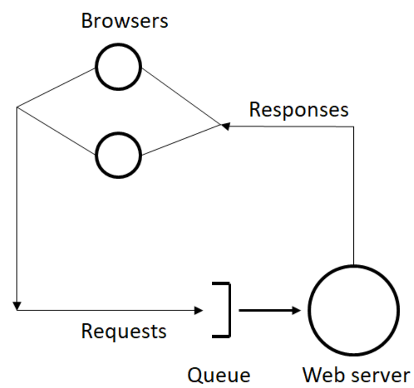

Fig. L2 - a model of a simple web application system.

(a) If there are 100 active users each with 20 seconds average think time (IID). The average system response time (IID) is 10 seconds which is the sum of queueing time and processing time. What is the throughput of the web server (responses/second)?
(b) If average queueing time is 9.75 second, what is the average processing time? What is the web server utilization (busy time/idle time)?
(c) Based on your study, without increasing system response time, what is the maximum number of active users the system can support?

Provide analytical solutions to questions (a) to (c) using queueing theory. Show the calculation steps and which fundament laws you applied.

Part B: Consider the model shown in Fig. L2, construct a simulation program use the knowledge you learned about simulation models so far (refer to textbook figure 1.3) and answer the above question in (a) to (c) using your simulation results. To simplify the simulation, please use the following assumptions and data set for now (they may not realistic; we will reconsider it after we cover random variables and their distributions):

1. Assume 100 active users with the "empty and idle" initial condition which means there are no requests in queue and web server is idle at t=0.
2. Assume the following: (i) user think time is randomly distributed in [18, 22] seconds uniformly; (ii) system response time is randomly distributed in [9, 11] seconds uniformly of which 97.5% is queueing time and 2.5% is processing time.
3. Assume, at t=0, all users are thinking. Then they generate next requests (arrival list) randomly described by assumption 2(i). The first request goes to server without waiting; others may wait in queue if server is busy.
4. Let the simulation runs for 5 hours then stop.

Lab2 Deliverables (the items you need to submit):

- A lab report details the following
  1. Your answer about Part-A questions (a) to (c) using queueing theory. Show the calculation steps and which fundament laws you applied.
  2. Your answer about Part-B questions (a) to (c) using simulation. Show the simulation data and explain how the data support your answers.
  3. You may use tables and/or graphical charts to represent the simulation data and show how close or how far to your analytical results in Part-A to question (a) to (c).
- Conclude the report with your explanation why your analytical results agree or disagree with your simulation results.
- Lab report be in WS Word or PDF format and does not contain any handwritten contents.
- Submit your source code.

Node:

1. You must submit all the deliverables via Blackboard. Email submission will not be accepted.
2. You can submit all the deliverables in two separated files or in one zip file. The lab report much be in PDF or MS Word file format. The source code file must be in the default format that the language uses, such as filename.py for Python code.
3. For students who prefer MATLab, C++, they can write their simulation in preferred language.