Service Engineering

Empirical Analysis of a Call Center via SEEStat

October 8th 2001 was "Columbus Day" in the USA. 'US bank' advertised that, immediately following the holiday, it would promote the opening of accounts to invest in felt (fabric, or in "Hebrew" - HASHKAOT BE-LEVED).

The promotion took place over the three days of 9-11/10/01, and turned out a huge success. Indeed, many customers called the **Telesales** category during those days, in an attempt to take advantage of the opportunity.

Remark on SEEStat analysis: Unless otherwise stated, carry out your performance analysis in 30-minutes resolution, and cover the hours of 7:00-23:00 per day.

For your convenience, you may use the glossary provided to this homework, available on ServEng website.

Part 1

Question 1.1:

Compare the three days 9-11/10/01 with other "similar" days (namely, with both equal week-days in other weeks of the same month, as well as days after long weekends in other months). The comparison here covers only "parameters" of the day (as opposed to performance measures), specifically arrival volume/rates, service time and number of agents (for the latter, use "Average agents in system").

Comment shortly on your findings.

Question 1.2:

Recall the formula for the "Offered Load per Server": $\rho = \dfrac{\lambda \cdot E(S)}{\# Servers}$

Extract from SEEStat the appropriate values and use a spreadsheet to calculate and plot "Offered Load per Server" for the days: 10/10/01, 11/10/01, and 16/10/01.

Compare the three graphs and describe any differences.
What do you expect the performance measures P(Abandon), E(W) and P(Wait>0) to be, for each of these three days? Answer without numerical calculations.

Question 1.3:

For 9.10.2001 only, **using SEEStat**, plot P(Abandon), the Offered Load and the Offered Load **per Server**. Fit all three on a single graph, using the "percent to mean" option.

Note:

1. Use SEEStat's variables "Offered load" and "Average Agents in System" to compute the Offered Load per Server.

2. SEEStat's feature "percent to mean" can not be used here as it does not fit our needs, Specifically, SEEStat uses the daily mean, while we must restrict the mean to our interval only (i.e. 7am-11pm). As a result, you are required to build your own "percent to mean" graph, using a spreadsheet (extract the proper values from SEEStat, calculate the proper averages and compute each interval's "percent to mean".)

Do any of the curves plotted exhibit similar trends? Discuss which measure, the Offered Load or the Offered Load per Server, can predict P(Abandon) better.

Note: As taught in class, the Offered Load has a time-varying version, which is more appropriate than simply using $R = \lambda \cdot E(S)$ at each time. (The latter formula is appropriate in steady state, with a constant arrival rate.) Since actual call center face time-varying arrival rates, the time-varying version of the Offered Load is the one to use.

To compute the time-varying version of the Offered Load, SEEStat actually computes the *workload* of a given realization (in our case, a single-day realization – that of 9.10.2001).

Part 2

For this part, consider the **2 days of 9-10/10/01 between 9:00-23:00**. (Note that the interval that ends at 23:00 is denoted in SEEStat as 22:30):

The high commotion at the call center, during the promotion days, resulted in errors of data accumulation. Specifically, it turns out that the call center did not do a very good job at recording the number of servers/agents. (This fact was ignored in Part 1- do not change anything in your analysis there.)

Question 2.1:

We learned in class that a service system could operate in any one of three regimes, denoted QD, ED and QED. Restricting attention to the above two days, which operational regime was the system functioning in? Support your claim with empirical results via SEEStat.
Next, without any calculations or further data analysis, provide your best assessment for the value of the MOPs P(Abandon), P(Wait>0) and E(Wait), had the system been functioning under each of the other two regimes.

Question 2.2:

The number of servers, denote it n, is in fact, the number of servers *available* for service (or FTEs, a term that was introduced in class). Explain why, during our 2 days (9-10/10/01 between 9:00-23:00, in 30 minutes resolution), n can be interpreted *also* as the number of *working* servers.

One way of estimating the number of *working* servers n, is by using the formula $n \cdot \mu = \lambda_{Ser}$; here λ_{Ser} is the arrival-rate to service (namely the rate of service-starts).

Compute the number of *both available and working* servers n using the above formula.

Note: In the sequel, use this method of estimating n wherever appropriate.

Question 2.3:

Plot P(Abandon) vs. the Offered Load per Server (as before, during 9-10/10/01 between 9:00-23:00, and for all 30 minute intervals). In calculating the latter, use your estimates of n from Question 2.2.
Now recall the theoretical relation between Offered Load per Server and P(Abandon), in the ED regime. Then plot, on the same graph, this theoretical relation, and compare it against practice - comment on your findings.

Part 3

Consider the day 10.10, during 7:00-12:00 in 30 minute intervals. (Note that the interval that ends at 12:00 is referred to in SEEStat as 11:30):

Question 3.1:

Assume constant average patience of customers over the period described – estimate this average patience. (Hint: You need to first compute properly E(Wait) and P(Abandon) for the period in question, resulting in a single value for each of these two MOPs, and then compute the average patience.)

In the sequel use this estimated value whenever patience estimation is needed.

Note: to compute the average wait time, use SEEStat's "Average wait time (without filter)" (this is useful for computing E(W) when there are very long waiting times)

Question 3.2:

We now wish to staff the call center so that it operates in the QED regime. Using the main characterization of the QED regime, reasonable target for a Measure of Performance (MOP) is $P(Wait > 0) \leq 0.5$. Using the appropriate Garnett function and the square root rule (here, use SEEStat's Offered load), determine, for each interval, the minimal staffing level that guarantees such performance (for each of the intervals separately).

Question 3.3:

In this question, we are going to use 4CC to support staffing of the call center. Specifically, 4CC will help identify staffing levels that guarantee predetermined performance goals, now given two service-level constraints (as opposed to the single constraint in the previous question):

1. $P(Wait > 0) \leq 0.5$
2. $P(Abandon) \leq 0.025$

4CC will be applied to each ½ hour interval separately. Its input parameters will be based on data imported from SEEStat. However, a careful application of 4CC turns out to require input parameters that must be calculated based on events outside of this interval. In the present question, we first demonstrate the magnitude of errors that could arise from an inappropriate data-translation (from SEEStat to 4CC), and then lead you to the right way of doing things.

a. Let us take for example the half-hour between 11:30-12:00 on 10.10. Extract from SEEStat both the number of calls and the average service time, during this interval. Input into 4CC the extracted values and calculate the least number of servers needed to adhere only to the above first constraint. Compare your answer to the staffing level obtained in Question 3.2 – is the difference reasonable?

One of the parameters of 4CC is "Calls per Interval". In theoretical models, this parameter equals the arrival-rate in a steady-state system. As already mentioned, estimating the value of this parameter from actual data requires care. (This is especially true in circumstances such as ours, namely the ED regime, where there is a significant overlap between adjacent intervals.)

The value of "Calls per Interval" in 4CC is the total number of **calls that require service** in that interval. SEEStat, when computing "Arrivals to Queue" in a specific interval, takes into consideration only calls that **arrived in that interval**.

b. Explain why, during the days discussed, calls that **require service** in some interval and calls that **arrive** in that interval may differ.

Another parameter that 4CC requires is "average time to handle one call". This parameter reflects the **average service time** of **calls that require service** during an interval. SEEStat, however, when computing "Average service time" in an interval, takes into account only calls that **started service during that interval.** Under the ED regime, a call that arrives in one interval often starts service in a later one. Thus, SEEStat's "Average service time" clearly cannot be used as input for 4CC in its computations.

The above difficulties do not arise in SEEStat's calculation of "Offered load" over an interval. Indeed, its calculation takes into consideration the **actual workload** during that interval (i.e. it is based on **calls that require service** during the interval and their corresponding **average service time**). Hence, SEEStat's calculation of "average workload" will be safely used in the sequel.

c. In the following Question d., you will be required to staff the call center so that it operates in the QED operational regime. In this regime, it is reasonable to assume that, on average, almost all calls that join the queue during a certain ½ hour interval also start service during that interval. Therefore, for QED staffing, one can actually rely on SEEStat's "Arrivals to queue" for the purpose of imputing the parameter "calls per interval".

   Here, you are required to find the proper values of 4CC's parameters for the period discussed.

   Given a specific interval, how does one determine the parameters of 4CC, specifically **calls that require service, denoted λ, and the average service time, denoted $E(S)$?** The following has been proved successful: start with the steady-state relation that relates the offered load, the calls that require service and the average service time, namely $R = \lambda \cdot E(S)$. Then use SEEStat's "Arrivals to queue" and "Offered load", to compute the average service time, separately for each ½ hour interval.

d. Propose staffing levels such that the call center maintains both $P(Wait > 0) \leq 0.5$ and $P(Abandon) \leq 0.025$, over each ½ hour interval between 7:00 and 12:00 (thus assuming that staffing level can be changed every 30 minutes). Describe and explain any discrepancies between your answer here and that in Question 3.2.

e. Qualitatively, compare your recommended staffing against actual staffing in the call center: was the call center well prepared for the load during the day of 10.10.2001?

f. Suppose that the hourly cost of an agent is 40 shekels per hour (and 20 per half hour). Experience has shown that the average profit from a served call during sale day is 7.5 shekels per call. Under these circumstances, would you

recommend changing into the QED staffing from Question d. above? Give an answer that is restricted to the time interval 11:00-11:30 only.
What is the conclusion from this answer?