CS 547 Lecture 24: Approximate Mean-Value Analysis

Daniel Myers

Previously, we stated that the mean-value analysis (MVA) algorithm would produce exact results for product-form queueing networks (PFQNs). This note covers two approximate forms of the MVA algorithm. The first

result, Schweitzer's approximation, allows us to compute the solution to a network without iterating over the full set of N customers. The second adapts the basic MVA approach to work with queueing centers with non-exponential service times.

Schweitzer's Approximation

The basic MVA algorithm relied on the *Arrival Theorem*. In a network with N total customers, the expected number observed at an arrival to queueing center k, denoted $\overline{A}_k(N)$, is simply the average number in the system with one customer removed.

$$\overline{A}_k(N) = \overline{Q}_k(N-1)$$

Schweitzer's approximation is based on a variation of the arrival theorem that removes the need to solve the network with $N-1$ customers. The approximation is

$$\overline{A}_k(N) \approx \frac{N-1}{N}\overline{Q}_k(N)$$

There are, on average, $\overline{Q}_k(N)$ customers at center k when there are N total customers in the system. If we know that one customer is arriving to center k, there are only $N-1$ other customers that could *potentially* be present at the arrival instant. Therefore, we expect the number of customers in the queue at an arrival to be a little less than $\overline{Q}_k(N)$.

Schweitzer's approximation assumes that all N customers make an equal contribution to the value of $\overline{Q}_k(N)$. Each customer's partial contribution is simply $\frac{1}{N}\overline{Q}_k(N)$. Therefore, if we know that one customer is arriving to the queue, there are $N-1$ other customers that *could* be in the queue at an arrival instant, and the sum of all their fractional contributions is $\frac{N-1}{N}\overline{Q}_k(N)$.

This approximation makes it possible to solve an MVA model without iterating from 1 to N. This isn't a major concern for single-class models, but it's very helpful for multi-class models.

Here's a sample implementation of MVA using Schweitzer's approximation.

```
# Initizalize queue lengths for all queueing centers
Q(k) = N/K for all k = 1 to K

converged = False
R_old = 0

# Loop until change in R is small
while not converged
```

```
  # Calculate residence time at each center
  R(k) = s(k) * (1 + (N-1)/N * Q(k)) for all k = 1 to K

  # Total residence time
  R_old = R
  R = sum V(k) * R(k) over k = 1 to K

  # Throughput
  X = N / (R + Z)

  # New approximate queue length
  Q(k) = X * V(k) * R(k) for k = 1 to K

  # Test for convergence
  if abs(R_old - R) < tolerance
    converged = True
  end

end
```

Schweitzer's approximation will converge, usually in a small number of iterations. Further, it will be *pessimistic* – it may yield estimates of residence time that are larger than the true residence times, but it will never predict a residence time lower than the true value.

Another Interpretation of Schweitzer's Approximation

The basic residence time equation in Schweitzer's approximation is

$$\overline{R}_k = \overline{s}_k(1 + \frac{N-1}{N}\overline{Q}_k)$$

If we use Little's law to combine our calculations of $\overline{Q}_k$ and the throughput Λ into one equation, we obtain

$$\begin{aligned}\overline{R}_k &= \overline{s}_k(1 + \frac{N-1}{N}\frac{N}{\overline{R}+\overline{Z}}\overline{V}_k\overline{R}_k) \\ &= \overline{s}_k(1 + \frac{\overline{V}_k\overline{R}_k}{\overline{R}+\overline{Z}}(N-1))\end{aligned}$$

This way of repeatedly calculating $\overline{R}_k$ leads to an interesting interpretation of Schweitzer's approximation. The term

$$\frac{\overline{V}_k\overline{R}_k}{\overline{R}+\overline{Z}}$$

is the average fraction of time customers spend at center k out of the total time required to make one trip through the network. If we consider one customer arriving to center k, there are $N-1$ other customers distributed throughout the network. The expected number of those other customer that will be at center k is

$$\frac{\overline{V}_k\overline{R}_k}{\overline{R}+\overline{Z}}(N-1)$$

Therefore, solving a closed model with Schweitzer's approximation is equivalent to approximating the fraction of time that each customer is at center k, and using that value to predice the number of customers in the queue at an arrival instant. You may recall that we used this approach to solve a simple OS semaphore model on the first day of class.

General Service Time Distributions

The original rules for product form networks stipulated that any FCFS queue had to have exponential service times. In networks with FCFS queues and non-exponential service times, the MVA algorithm will not produce exact results. In practice, however, MVA has proven to be reasonably accurate in these situations, so approximate MVA models are widely used for systems analysis.

Recall that solving the M/G/1 model required adapting our tagged customer analysis to account for the waiting time due to a customer currently in service.

We'll assume the Arrival Theorem still holds, so that $\overline{A}_k(N) = \overline{Q}_k(N-1)$. We'll also assume that the probability that a queue is busy at an arrival instant is simply $U_k(N-1)$, the utilization of center k when there are $N-1$ customers in the system. Trivially, $U_k(1) = 0$.

Combining these results,

$$\overline{R}_k(N) = U_k(N-1)\frac{\overline{s}_k(1+c_k^2)}{2} + (\overline{Q}_k(N-1) - U_k(N-1))\overline{s}_k + \overline{s}_k$$

Here's a sample implementation of this algorithm. Note the extra step to compute the new utilizations at the end of each iteration.

```
# Initialize queue lengths and utilizations
Q(k) = 0 for k = 1 to K
U(k) = 0 for k = 1 to K

for n = 1 to N

  # Residence time calculation
  R(k) = U(k) * (s(k) / 2) * (1 + c(k)^2) + (Q(k) - U(k)) * s(k) + s(k)

  # Total residence time
  R = sum V(k) * R(k) over all k = 1 to K

  # Throughput
  X = n / (R + Z)

  # Calculate new queue lengths and utilizations
  Q(k) = X * V(k) * R(k) for all k = 1 to K
  U(k) = X * V(k) * s(k) for all k = 1 to K

end
```