22.4UC2

Numerical Computation and Statistics in Engineering

Unit 8: Simulations

Simulation

- Simulation enables the study of, and experimentation with, the interactions of a complex system (or a subsystem thereof).
- Informational, organisational, and environmental changes can be simulated and the changes to the model's behaviour can be observed.
- Knowledge gained by studying the model can be of great value, and may suggest improvement to the system under investigation.
- Changing simulation inputs can help understanding how certain components of the system interact.
- Simulation is, among others, used to experiment with new designs or policies prior to implementation, and may save huge amounts of money.
- Simulations are also used to verify analytical solutions.

Advantages and Disadvantages

Following Pegden, Shannon, and Sadowski (1995) we list some of the advantages and disadvantages of simulations:

- New policies, operating procedures, decision rules, information flows, organisational procedures, etc. can be tested without disrupting ongoing operations of the real system.
- New hardware designs, physical layouts, transportations systems, etc. can be tested without committing resources.
- Hypothesis about how and why certain phenomena occur can be tested for feasibility.
- Time can be compressed or expanded for a speed up or slow down of phenomena under investigation.
- Insight can be obtained about the interaction of parameters.
- Insight can be obtained about the importance of variables.

Advantages and Disadvantages

- Bottleneck analysis can be performed.
- A simulation can help in understanding how a system works, opposed to how individuals think the system works.
- Questions like ‘what if’ can be answered.

Advantages and Disadvantages

There are also some disadvantages:

- Model building requires special training, and designing complex simulations is an art form. In practice, two models designed by different competent people will usually differ considerably, although there will be similarities.
- Simulation results may be difficult to interpret.
- Simulation modeling and analysis can be time consuming and expensive.
- Simulation is sometimes used where analytical models are available and even preferable.

Applications

The following are a couple of potential applications of simulation related to electrical engineering (in the broadest sense):

- Material handling system design for semiconductor manufacturing
- Assembly operations
- Distributed models for computer integrated manufacturing
- Inventory cost model for 'just-in-time' production
- Heterogeneous networks
- Evaluating large-scale computer system performance
- Client/server system architecture

Terminology

- *Entity* — object of interest in the system.
- *Attribute* — property.
- *Activity* — performing a task over a time period of specified length.
- *State* — collection of variables necessary to describe the system.
- *Event* — Occurrence that may change the system.

We also distinguish between *endogenous* activities, which are activities within the system, and *exogenous* activities.

Steps in Simulations

- Setting objectives/Overall design
  - Problem formulation
  - Setting objectives and overall project plan
- Model building
  - Model conceptualisations
  - Data collection (for input)
  - Model translation
  - Verification of the (computer) model
  - Valididation (accurate description of the real world)
- Running the model
  - Experimental design
  - Production runs and analysis
- Documentation and reporting

Queueing Systems

A *queueing system* consists of at least a server, a waiting line, and a calling population.

Examples are

System	Customer/Population	Server
reception desk	people	receptionist
garage	cars	mechanics
warehouse	pallets	crane
computer	jobs	CPU, disc, or tapes
telephone	calls	exchange

Calling Population

Calling populations are often assumed to be 'infinite' if the real population is large. This simplifies the model. For infinite populations the *arrival rate* is not affected by the number of people that left the population and entered the queue.

Typical examples with 'infinite' populations are

- people being served at an ATM,
- cars served (serviced) in a garage.

However, the one company helicopter that can be serviced is definitely a finite population.

System Capacity

In many queueing systems there is a limit to the length of the queue:

- A buffer holding requests to a CPU may have a fixed length.
- The buffer of a printer usually has a fixed length.
- A lane of a drive-in restaurant may only hold 10 cars.

This is *not* to be confused with customer behaviour, where customers will not queue if there are already ten customers waiting in line to be served.

Arrival Process

The arrival process is characterised by the time intervals between customer arrivals. Such arrivals may be *scheduled*, or occur at random times, in which case they are characterised by a probability distribution.

For example, if arrival of customers is modeled by a Poisson process with mean λ (arrivals per time unit), then interarrival time is modelled by an exponential distribution with mean λ^{-1} (time units).

The expected number of arrivals in t time intervals is thus $t\lambda$.

The Poisson process is often used to model arrival of customers.

Next to scheduled and random arrivals there is also the possibility that there is *always* one customer in the queue. This happens for example if the customer represents raw material, and the production process guarantees that such material is always present.

Queue Behaviour and Discipline

Queue behaviour refers to the actions customers take while arriving or waiting in the queue. Customers may

- balk (leave, if the queue is too long),
- renege (leave, if the queue proceeds too slowly), or
- jockey (move between lines according to what the customer thinks is the shortest/fastest line).

Discipline refers to the logical order in which customers are served. Models are

- FIFO (first-in-first-out), LIFO (last-in-first-out)
- SIRO (service in random order)
- SPT (shortest processing time first)
- PR (service according to priority)

Clock Time

Suppose interarrival times are determined by rolling a die.

Customer	Interarrival time	Arrival time on clock
1	-	0
2	6	6
3	1	7
4	4	11
5	3	14
6	6	20
7	5	25

Suppose also that the service time for each customer is

$$2 \quad 3, \quad 1, \quad 1, \quad 1, \quad 1, \quad 2$$

Clock Time

time units respectively, which yields the following data

Customer	Arrival time	Service begin	Service time	Service ends
1	0	0	2	2
2	6	6	3	9
3	7	9	1	10
4	11	11	1	12
5	14	14	1	15
6	20	20	1	21
7	25	25	2	27

Single Queue Model

On the following slides we will go through a single queue system in some detail. We consider a server model with interarrival times between 1 and 8 time units. The probability of each time interval of length 1 is 1/8. We will use random numbers to generate arrivals as follows:

Time betw. arrivals	probability	cum. prob.	random digits
1	0.125	0.125	001–125
2	0.125	0.250	126–250
3	0.125	0.375	251–375
4	0.125	0.500	376–500
5	0.125	0.625	501–625
6	0.125	0.750	626–750
7	0.125	0.875	751–875
8	0.125	1.000	876–000

Single Queue Model

The service time, between 1 and 6 time units, is as follows:

Service time	probability	cum. prob.	random digits
1	0.10	0.10	01–10
2	0.20	0.30	11–30
3	0.30	0.60	31–60
4	0.25	0.85	61–85
5	0.10	0.95	86–95
6	0.05	1.00	96-00

Using C++ we generate the following table of random numbers, 60 for simulating 20 customer arrivals, and another 40 to simulate their service time:

Single Queue Model

customer	random digits	time	customer	random digits	time
1	—	—	11	535	5
2	629	6	12	127	2
3	093	1	13	606	5
4	261	3	14	879	8
5	202	2	15	375	4
6	922	8	16	897	8
7	361	3	17	293	3
8	194	2	18	784	7
9	503	5	19	610	5
10	632	6	20	061	1

Single Queue Model

customer	random digits	time	customer	random digits	time
1	55	3	11	47	3
2	65	4	12	69	4
3	37	3	13	45	3
4	25	2	14	47	3
5	44	3	15	30	2
6	78	4	16	68	4
7	84	4	17	31	3
8	49	3	18	20	2
9	68	4	19	92	5
10	66	4	20	49	3

Here is the data of the system:

Single Queue Model

Customer	Time since last arrival	Arrival time	Service time	Start service	Waiting (queue)	End service	Time spent in system	Idle time
1	0	0	3	0	0	3	3	0
2	6	6	4	6	0	10	4	3
3	1	7	3	10	3	13	6	0
4	3	10	2	13	3	15	5	0
5	2	12	3	15	3	18	6	0
6	8	20	4	20	0	24	4	2
7	3	23	4	24	1	28	5	0
8	2	25	3	28	3	31	6	0
9	5	30	4	31	1	35	5	0
10	6	36	4	36	0	40	4	1

Single Queue Model

Customer	Time since last arrival	Arrival time	Service time	Start service	Waiting (queue)	End service	Time spent in system	Idle time
11	5	41	3	41	0	44	3	1
12	2	43	4	44	1	48	5	0
13	5	48	3	48	0	51	3	0
14	8	56	3	56	0	59	3	5
15	4	60	2	60	0	62	2	1
16	8	68	4	68	0	72	4	6
17	3	71	3	72	1	75	4	0
18	7	78	2	78	0	80	2	3
19	5	83	5	83	0	88	5	3
20	1	84	3	88	4	91	7	0
			66		20		86	25

Single Queue Model

Let us summarise some of the data:

1. The average waiting time per customer is 1 time unit.

$$\text{average waiting time} = \frac{\text{total waiting time}}{\text{number of customers}} = \frac{20}{20} = 1$$

2. The probability for a customer to wait in the queue is 45%.

$$\text{probability(wait)} = \frac{\text{total number of waiting customers}}{\text{number of customers}} = \frac{9}{20} = 0.45$$

Single Queue Model

3. The proportion of idle time of the server is 27%.

$$\begin{aligned}\text{Probability(server idle)} &= \frac{\text{idle time}}{\text{total run time}} \\ &= \frac{25}{91} = 0.27\end{aligned}$$

4. The average service time is 3.3 time units.

$$\begin{aligned}\text{aerage service time} &= \frac{\text{total service time}}{\text{number of customers}} \\ &= \frac{66}{20} = 3.3\end{aligned}$$

Note that the expected service time is slightly lower,

$$\begin{aligned}E(s) &= 1(0.1) + 2(0.2) + 3(0.3) + 4(0.25) + 5(0.1) + 6(0.05) \\ &= 3.2\,.\end{aligned}$$

Single Queue Model

5. The average time between arrivals is 4.42 time units.

$$\begin{aligned} \text{average(interarrival)} &= \frac{84}{19} \\ &= 4.42\,. \end{aligned}$$

The expected time between arrivals is

$$E(A) = \frac{1+8}{2} = 4.5$$

time units.

6. The average waiting time for customers that have to wait is 2.22 time units.

$$\begin{aligned} \text{average(waiting time)} &= \frac{\text{total waiting time}}{\text{number of waiting customers}} \\ &= \frac{20}{9} = 2.22 \end{aligned}$$

Single Queue Model

7. The average time spent by a customer in the system is 4.3 time units.

$$\begin{aligned} \text{avarage(time spent)} &= \frac{\text{total time spent}}{\text{number of customers}} \\ &= \frac{86}{20} \;=\; 4.3 \end{aligned}$$

Alternatively, the average time spent in the system is the average waiting time plus the average server time, which is

$$3.3 + 1 \;=\; 4.3\,.$$

Example

We consider a second simulation. Here two server serve one queue. We assume that a customer is served by the server that is available. If both servers are available then the customer is assigned randomly to one of the servers. The servers take different times to serve a customer (these could be mechanics, one being more senior and faster than the other, or network servers, where one model is older than the other). Below are the distribution functions for the two servers A and B, and the distribution for the arrival time between processes.

Interarrival time

Interarrival time	Prob.	Cum. prob.	Random digits
1	0.25	0.25	01–25
2	0.40	0.65	26–65
3	0.20	0.85	66–85
4	0.15	1.00	86-00

Example

Server A:

Service time	Prob.	Cum. prob.	Random digits
1	0.10	0.10	01–10
2	0.30	0.40	11–40
3	0.25	0.65	41–65
4	0.20	0.85	66–85
5	0.15	1.00	86–00

Server B:

Service time	Prob.	Cum. prob.	Random digits
2	0.35	0.35	01–35
3	0.27	0.62	36–62
4	0.20	0.82	63–82
5	0.18	1.00	83-00

Example

Customer	Random digits	Time betw. arrivals	Clock time of arrival	Random digits (service)	Server	A start	A service time	A end	B start	B service time	B end	Time in queue
1	0	4	0	75	B				0	4	4	0
2	56	2	2	29	A	2	2	4				0
3	27	2	4	9	B				4	2	8	0
4	60	2	6	62	A	6	3	9				0
5	18	1	7	79	B				8	4	12	1
6	2	1	8	37	A	9	2	11				1
7	92	4	12	28	A	12	2	14				0
8	73	3	15	61	B				15	3	19	0
9	93	4	19	19	A	19	2	21				0
10	78	3	22	45	B				22	3	26	0

Example

Customer	Random digits	Time betw. arrivals	Clock time of arrival	Random digits (service)	Server	A start	A service time	A end	B start	B service time	B end	Time in queue
11	36	2	24	10	A	24	1	25				0
12	32	2	26	6	A	26	1	27				0
13	55	2	28	47	B				28	3	32	0
14	56	2	30	93	A	30	5	35				0
15	45	2	32	25	B				32	2	36	0
16	74	3	35	43	A	35	3	38				0
17	78	3	38	68	B				38	4	42	0
18	43	2	40	14	A	40	2	42				0
19	20	1	41	68	B				42	4	46	1
20	26	2	43	64	A	43	3	46				0

Example

Customer	Random digits	Time betw. arrivals	Clock time of arrival	Random digits (service)	Server	A start	A service time	A end	B start	B service time	B end	Time in queue
21	50	2	45	48	B				46	3	50	1
22	17	1	46	27	A	46	2	48				0
23	26	2	48	10	A	48	1	49				0
24	15	1	49	94	A	49	5	54				0
25	9	1	50	17	B				50	2	54	0
26	11	1	51	59	B				54	3	58	3
27	76	3	54	73	A	54	4	58				0
28	56	2	56	39	A	58	2	60				2
29	81	3	59	29	B				59	2	63	0

Example

- Both server work about the same amount of time (A 40 time units, B 39 time units).
- Server A is 63% of the time busy, server B also 63%.
- Only 6 customers had to wait, and the total waiting time is very low with 9 time units. The average waiting time for those who had to wait was just 0.67 time units.
- The system looks well-balanced. One server would not be able to do the job, and, unless the costs for waiting are very high, a third server is not justified.