Recitation 14

Part 1. Shift Scheduling.

We know how to determine the "optimal" (desired) staffing for constant arrival rate.

What to do if the rate is changing in time?

For example, divide the whole working day into a number of equal intervals $i = 1,..,I$, and assume that the arrival rate may be different for each of the intervals.

How to determine the optimal staffing for the whole working day?

- Assume that we calculated the optimal staffing level N_i for each interval.
- The problem is that one can not employ a server only for a small time interval, but for a certain period of time, called shift.
- A system of shift scheduling needs to be considered:

1. A shift denotes a set of time intervals during which a server works over the course of the day.

2. A schedule is a set of daily shifts to which an employee is assigned over the course of a day, week or month.

For example, a feasible shift may start on the half-hour, last 9 hours, including an hour total of break time. One half hour of this break must be devoted to lunch, which must begin sometime between two and three hours after the shift begins, and the other to a morning or afternoon pause.

A feasible schedule may require an employee to work five, 9-hour shifts each week of the month, on Sunday, Monday, Tuesday, Friday, and Saturday. Another may require a CSR to work a different set of shifts each week of the month.

Assume there is a collection of $j = 1,\ldots, J$ feasible schedules to which employees may be assigned. A collection of feasible shifts may be described in terms of the $I \times J$ matrix $A = (a_{ij})$, where

$$a_{ij} = \begin{cases} 1, & \text{if an agent working according to schedule j is available during interval i;} \\ 0, & \text{otherwise.} \end{cases}$$

time		j = 1	2	3	4	5	6	7	8	9	10
8:00-8:29am	i = 1	1					1				
8:30-8:59	2	1	1				1	1			
9:00-9:29	3	1	1	1			1	1	1		
9:30-9:59	4		1	1	1		1	1	1	1	
10:00-10:29	5	1		1	1	1		1	1	1	1
10:30-10:59	6	1	1		1	1	1		1	1	1
11:00-11:29	7	1	1	1		1	1	1		1	1
11:30-11:59	8	1	1	1	1		1	1	1		1
12:00-12:29pm	9		1	1	1	1	1	1	1	1	
12:30-12:59	10			1	1	1		1	1	1	1
1:00-1:29	11	1			1	1			1	1	1
1:30-1:59	12	1	1			1	1			1	1
2:00-2:29	13	1	1	1			1	1			1
2:30-2:59	14	1	1	1	1		1	1	1		
3:00-3:29	15		1	1	1	1		1	1	1	
3:30-3:59	16	1		1	1	1	1		1	1	1
4:00-4:29	17	1	1		1	1	1	1		1	1
4:30-4:59	18	1	1	1		1	1	1	1		1
5:00-5:29	19		1	1	1			1	1	1	
5:30-5:59	20			1	1	1			1	1	1
6:00-6:29	21				1	1				1	1
6:30-6:59	22					1					1

Figure 7: An Example A-Matrix for the Scheduling Problem

Above we see the complete A-matrix for schedules that cover one 11-hour day (for simplicity, rather than 30-days). To enhance readability, only the matrix's ones are shown, not the zeros. Each of the 22 rows represents a different 1/2-hour interval, and each of the 10 columns represents a different schedule to which employees may be assigned. Inspection reveals that the first 5 columns all have the same structure; the only difference among them is the time that an employee assigned to the schedule would start. Similarly, the second set of 5 columns share the same structure. Every one of the 10 schedules has an employee take calls for 7 hours of a 9 hour day.

Assume that the monthly (or daily, see the problem itself) cost of assigning an agent to schedule j equals c_j.

How to determine the optimal set of schedules? **Integer Programming** is needed…

Let the decision variables X_j, $j = 1,..,J$, represent the numbers of agents assigned to the various schedules, and let N_i, $i = 1,..,I$, denote the staffing requirements for each interval i. Then solve

Integer Programming Formulation

Min $\sum_{j=1}^{J} c_j X_j$

s.t.

$$\sum_{j=1}^{J} a_{ij} X_j \geq N_i, \quad i = 1,..,I$$

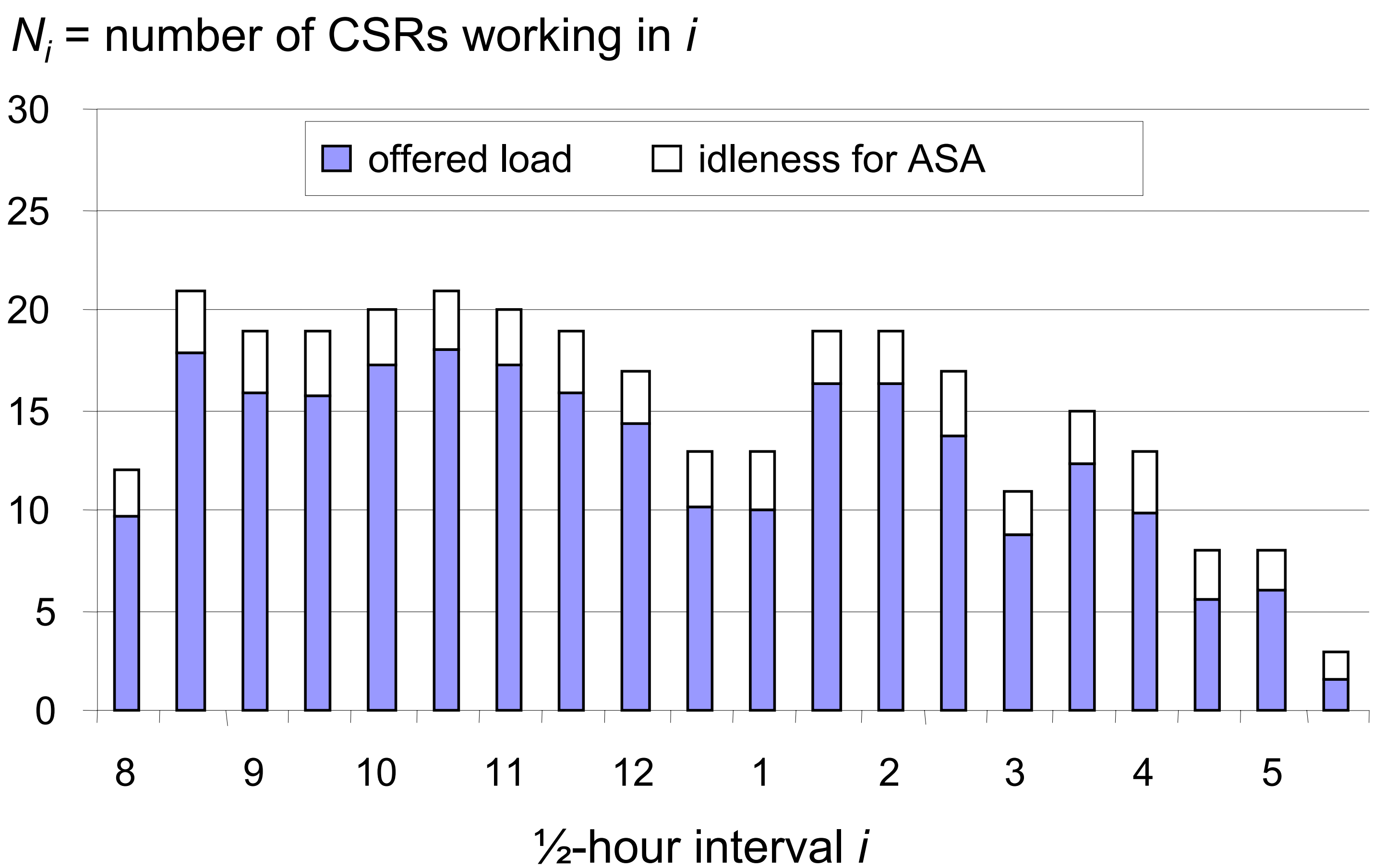
$$X_j \geq 0, \text{ integer.}$$

Example:

- Day-of-work: 8:00-18:00.
- Four-hour shifts without breaks which can start every half-hour, 8:00-14:00.
- **Performance goal:** ASA less than 30 seconds for every half-hour interval.
- Same cost for all schedules. Hence, minimize overall number of servers.

Tutorial, basic model...

Example with 30-second ASA:



Determining staffing requirements for a day or week:

- Data

  N_i = num. CSRs required for interval i

  c_j = cost of putting a CSR on schedule j

  $a_{ij} = \begin{cases} 1 & \text{if sched. } j \text{ take calls during int. } i, \\ 0 & \text{otherwise.} \end{cases}$

- Decision variables

  x_j = num. CSRs to work on schedule j

- Math program to choose schedules

$$\min\{c'x \,|\, Ax \geq N; x \geq 0; x \text{ integer}\}$$

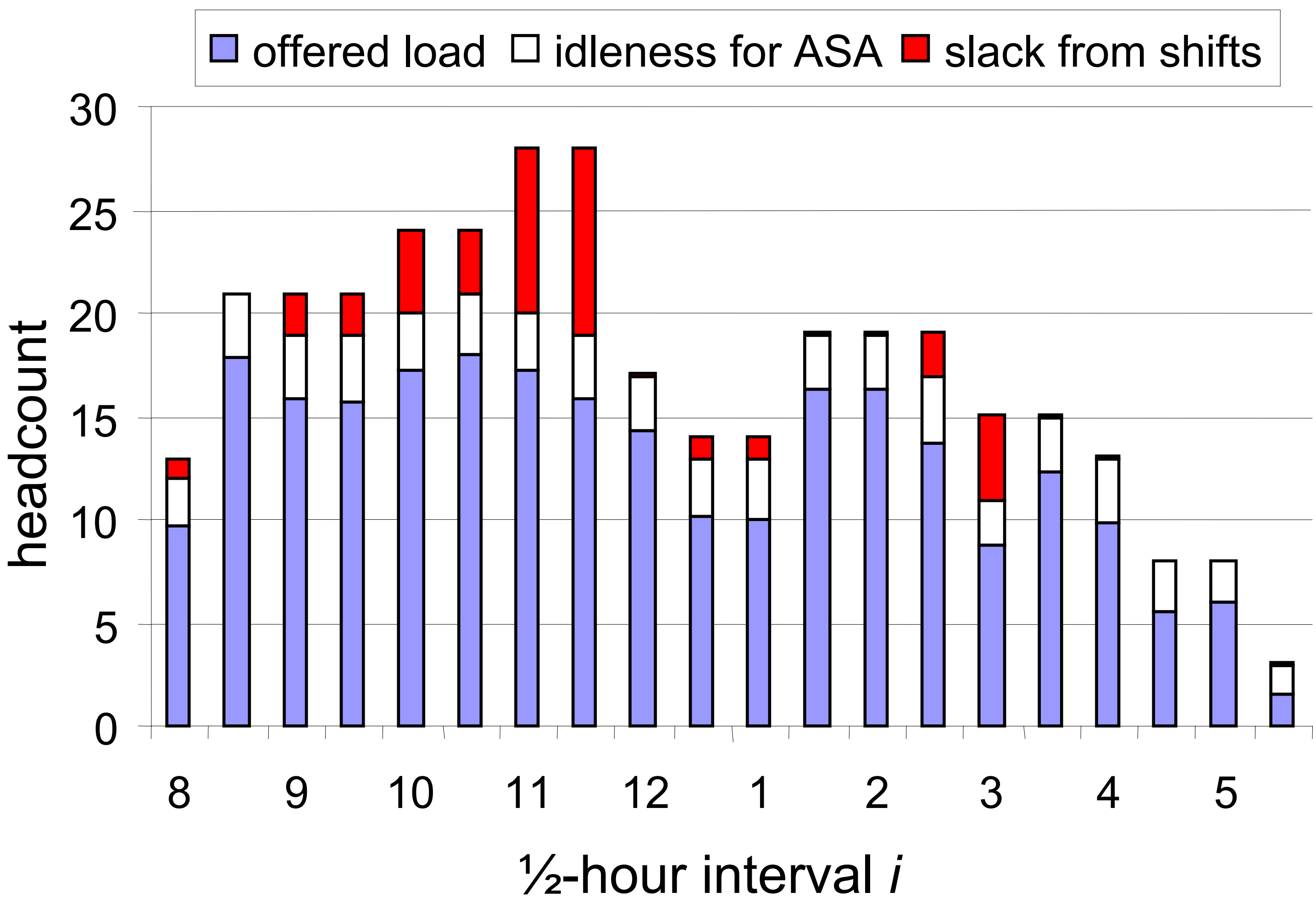
Tutorial, basic model...

Example with 4-hour shifts on the 1/2-hour:

Shifts = x_j

	8	**8:30**	**9**	**9:30**	**10**	**10:30**	**11**	**11:30**	**12**	**12:30**	**1**	**1:30**	**2**		
43	13	8	0	0	3	0	4	0	2	5	0	5	3		
														total	**N_i**
8:00-8:29	1													13	12
8:30-8:59	1	1												21	21
9:00-9:29	1	1	1											21	19
9:30-9:59	1	1	1	1										21	19
10:00-10:29	1	1	1	1	1									24	20
10:30-10:59	1	1	1	1	1	1								24	21
11:00-11:29	1	1	1	1	1	1	1							28	20
11:30-11:59	1	1	1	1	1	1	1	1						28	19
12:00-12:29		1	1	1	1	1	1	1	1					17	17
12:30-12:59			1	1	1	1	1	1	1	1				14	13
1:00-1:29				1	1	1	1	1	1	1	1			14	13
1:30-1:59					1	1	1	1	1	1	1	1		19	19
2:00-2:29						1	1	1	1	1	1	1	1	19	19
2:30-2:59							1	1	1	1	1	1	1	19	17
3:00-3:29								1	1	1	1	1	1	15	11
3:30-3:59									1	1	1	1	1	15	15
4:00-4:29										1	1	1	1	13	13
4:30-4:59											1	1	1	8	8
5:00-5:29												1	1	8	8
5:30-5:59													1	3	3

Excess capacity resulting from the scheduling IP:



Part 2. Introduction to Operational Regimes

Example from exam (Spring, 2006)

Consider 3 examples of ACD reports from an Israeli Telecom call center:

Report 1

Time	Arrivals	Average Service Time (sec)	Offered Load	Number of Agents	Fraction Abandoning (%)	Average Wait (sec)	Served Immediately (%)
22:00	476	160.1	42.3	34.3	21.8	104.4	0.8
22:30	**416**	**196.5**	**45.4**	**28.9**	**38.7**	**163.2**	**0.4**
23:00	346	194.0	37.3	24.8	41.0	171.1	0.9

Report 2

Time	Arrivals	Average Service Time (sec)	Offered Load	Number of Agents	Fraction Abandoning (%)	Average Wait (sec)	Served Immediately (%)
13:30	859	70.9	33.8	53.6	0.0	0.0	100.0
14:00	**768**	**75.1**	**32.0**	**57.3**	**0.0**	**0.0**	**100.0**
14:30	700	80.2	31.2	52.8	0.0	0.0	100.0
15:00	596	84.5	28.0	39.0	0.0	1.6	88.1

Report 3

Time	Arrivals	Average Service Time (sec)	Offered Load	Number of Agents	Fraction Abandoning (%)	Average Wait (sec)	Served Immediately (%)
13:00	176	202.2	19.8	18.8	5.7	30.6	23.5
13:30	**175**	**207.1**	**20.1**	**20.1**	**3.5**	**17.1**	**54.2**
14:00	169	196.3	18.4	20.9	1.8	5.4	74.7

Question 1. Explain how the Offered Load has been calculated.

Solution: $R = \frac{\lambda \cdot E(S)}{1800}$, where $E(S)$ is the average service times, 1800 – number of seconds in half-hour interval.

Question 2. Explain how non-integer number of agents could arise.

Solution: Number of agents changed during half-hour interval.

Question 3. Write an operational regime that corresponds to each report.

Solution:

Report 1: ED

Report 2: QD

Report 3: QED

Question 4. For each report, write the most relevant staffing rule that establishes the characterizing relation between the offered load R and the number of agents n. Calculate parameter values (QoS grades) for the second line of each report.

Solution:

Report 1: ED. $n = R - \gamma R,\ \ \gamma > 0$.

For the second line $\gamma = 1 - n/R = 0.363$.

Report 2: QD. $n = R + \delta R,\ \ \delta > 0$.

For the second line $\delta = n/R - 1 = 0.791$.

Report 3: QED. $n = R + \beta\sqrt{R},\ \ -\infty < \beta < \infty$.

For the second line $\beta = (n-R)/\sqrt{R} = 0$.

<u>Question 5.</u> For all intervals, presented in the reports, assume that the nine-fold increase of the arrival rate and the number of agents took place. For the second line of each report, estimate how the average wait will change in this case.

Solution:

Report 1: Average wait will remain 163 seconds.

Report 2: Average wait will remain zero.

Report 3: A three-fold decrease of average wait will take place (to 5.7 seconds).

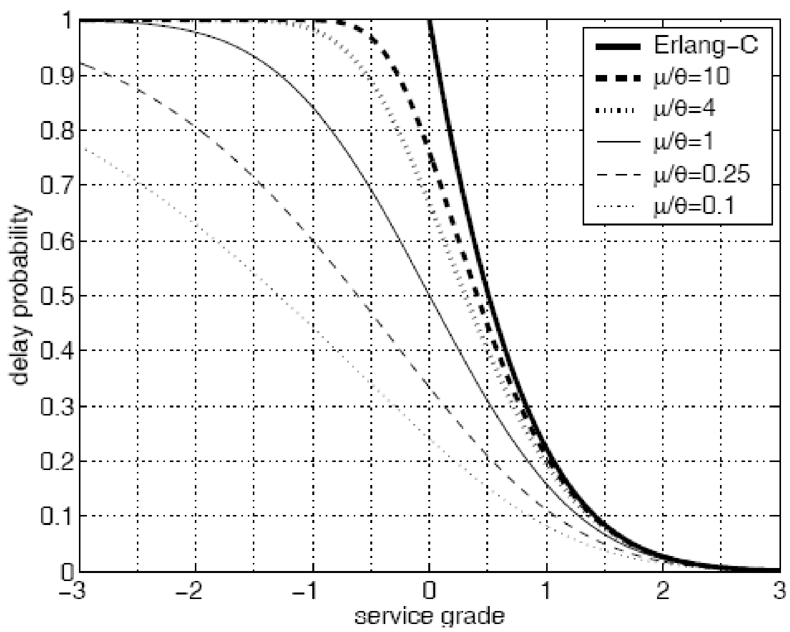
<u>Question 6.</u> (This question is not related directly to the previous ones.)

Assume that in a telephone call center the arrival rate is 100 calls per minute, the average service time is equal to 1 minute and the average patience is equal to 4 minutes. There are 105 agents answering calls. Using the graph below, estimate the fraction of customers that are served without delay.

Solution: The system is operating in the QED regime. The offered load R=100.

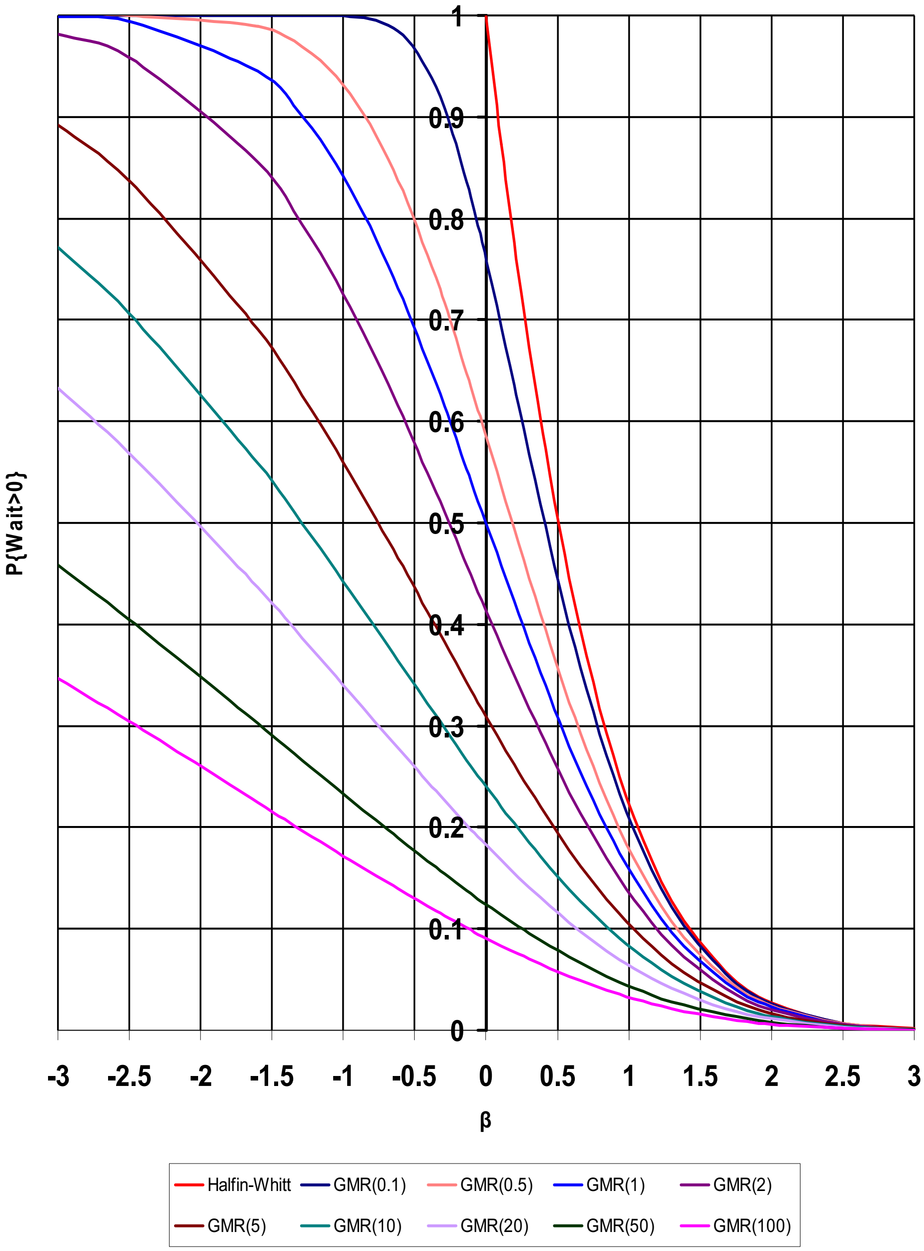
The QoS grade $\beta = \frac{105-100}{\sqrt{100}} = 0.5$.

The ratio $\mu/\theta = 4$. According to the graph, the delay probability is 0.4. Hence, 60% of customers are served without delay.



Erlang-A: The Garnett Delay-Functions

$P\{W_q > 0\}$ vs. the QOS parameter β, for varying patience θ/μ.



GMR(x) describes the asymptotic probability of delay as a function of β when $\theta/\mu = x$. Here, θ and μ are the abandonment and service rate, respectively.

Note: **Erlang-C** = limit of **Erlang-A**, as patience ↑ indefinitely.