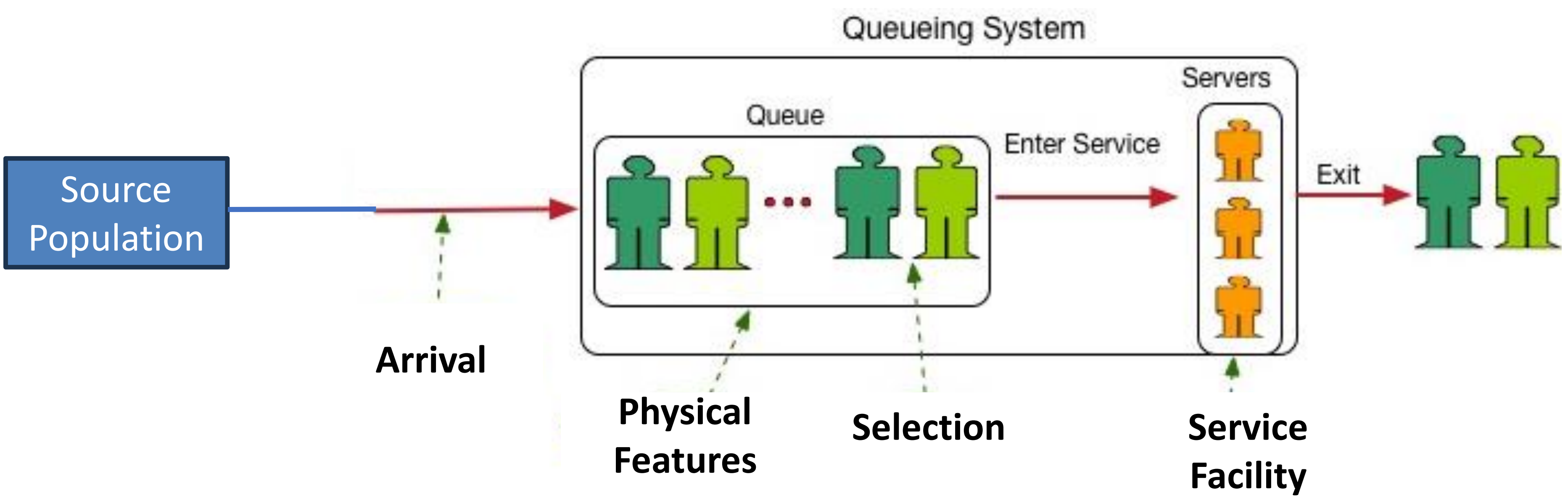
VIth SEMESTER
OPERATIONS RESEARCH
4. QUEUEING MODEL (WAITING LINE MODELS)

Introduction

- A queue is formed at a production/operation system when either customers (human beings or physical entities) requiring service wait because number of customers exceeds the number of service facilities, or service facilities do not work efficiently/take more time than prescribed to serve a customer.
- Queuing theory can be applied to a variety of situations where it is not possible to accurately predict the arrival rate (or time) of customers and service rate (or time) of service facility or facilities.
- In particular, it can be used to determine the level of service (either the service rate or the number of service facilities) that balances the following two conflicting costs: (i) cost of offering the service (ii) cost incurred due to delay in offering service.
- The first cost is associated with the service facilities and their operation, and the second represents the cost of customers waiting for service.

Introduction

Structure of the Queueing System



Introduction

Structure of the Queueing System

- *Source Population*
  - *Finite population model*
    - ✓ If the arrival rate depends on the number of customers being served and waiting.
    - ✓ Where new customers can only be added to the line once others move out of the line.
    - ✓ Example: machine operator responsible only to handle 5 machines, etc.

Introduction

Structure of the Queueing System

- *Source Population*
  - *Infinite population model*
    - ✓ The rate of arrival of customers is not affected by the number of customers that have already joined the queuing system.
    - ✓ Where new customers are not affected by the number of customers already in the system.
    - ✓ Examples: Banks, supermarkets, petrol pumps, ticket counters, restaurants, etc.

Introduction

Structure of the Queueing System

- *Arrival Characteristics*
  - *Size of the arrival*
    - ✓ Customers may arrive in batches (such as the arrival of a family at a restaurant) or individually.
    - ✓ These customers may arrive at a service facility either on a scheduled time (by prior information) or on unscheduled time (without information).

Introduction

Structure of the Queueing System

- *Arrival Characteristics*
  - *Pattern of arrival*
    - ✓ Controllable: Examples: movie theatres offering Monday specials, Airlines offering off-season rates, etc.
    - ✓ Uncontrollable: Examples: Emergency operations, fire department, etc.

Introduction

Structure of the Queueing System

- *Arrival Characteristics*
  - *Distribution of the arrival*
    - ✓ Constant arrival pattern (example: A component/part of the assembly arrive at every 30 minutes)
    - ✓ Variable arrival pattern (example: people arriving in a bank)

Variable Pattern

Constant Pattern

Introduction

Structure of the Queueing System

- *Arrival Characteristics*
  - *Degree of patience*
    - ✓ Patient customer - The customer arrives at the service system, waits in the queue until served, and does not switch between waiting lines.
    - ✓ Impatient customer - The customer, who waits for a certain time in the queue and leaves the service system without getting service due to certain reasons.

Introduction

Structure of the Queueing System

- *Arrival Characteristics*
  - *Degree of patience*
    - ✓ Impatient customer (Queue Behavior examples)
      - Balking**:** Customers do not join the queue either by seeing the number of customers already in service system or by estimating the excessive waiting time for the desired service.
      - Reneging: Customers, after joining the queue, wait for sometime in the queue but leave before being served on account of certain reasons.
      - Jockeying: Customers move from one queue to another hoping to receive service more quickly.

Introduction

Structure of the Queueing System

- *Physical features*
  - *Length of the queue*
    - ✓ Finite queue length (restriction on the length of the queue)
      - Restrict queue length to a certain limit after which people who come into the system do not join the system. (Example: Garage)
    - ✓ Infinite queue length (no restriction on the length of the queue)

Introduction

Structure of the Queueing System

- *Physical features*
  - *Number of lines (or server)*
    - ✓ Single server queueing model (where there is only server)
    - ✓ Multiple server queueing model (where there are multiple servers)

Introduction

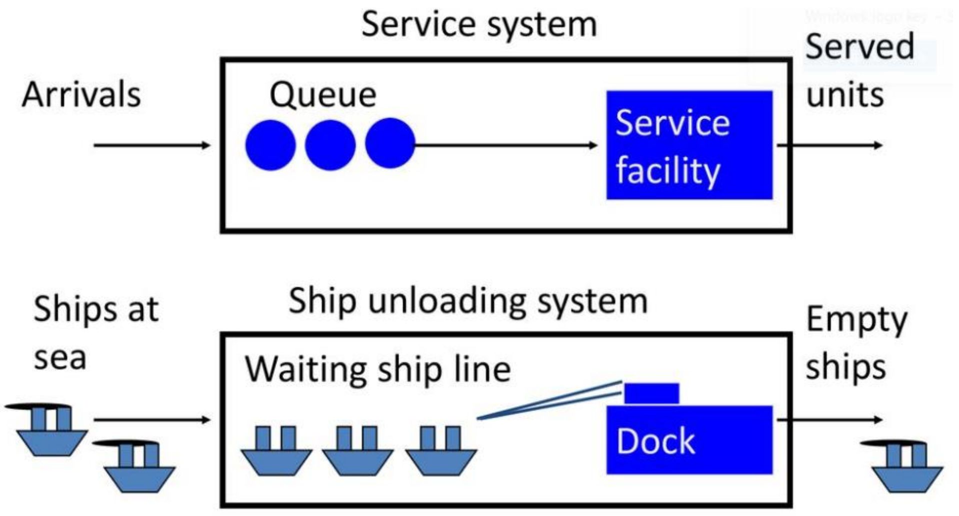
Structure of the Queueing System

- *Selection – from the waiting line*
  - *Queue discipline: the logical ordering of customers in a queue that determines which customer is chosen for service when a server becomes free*
    - ✓ First-in-first-out (FIFO), Last-in-first-out (LIFO), Service in random order (SIRO), Shortest processing time first (SPT), Service according to priority (PR), Limited needs first.

Introduction

Structure of the Queueing System

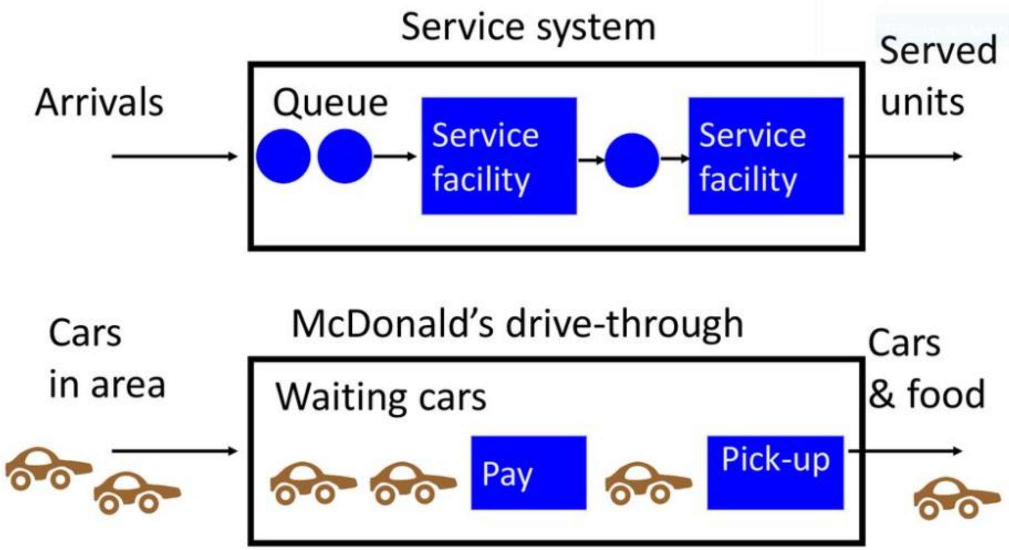
- *Service facility*
  - *Structure*
    - ✓Single channel, single phase



Introduction

Structure of the Queueing System

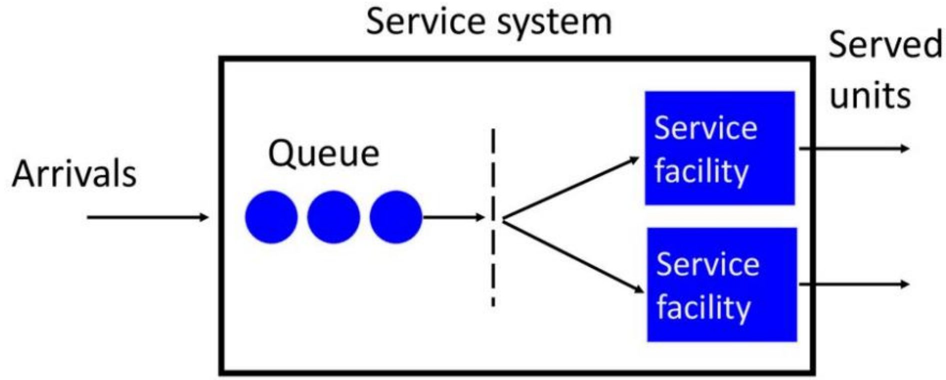
- *Service facility*
  - *Structure*
    - ✓ Single channel, multiple phase



Introduction

Structure of the Queueing System

- *Service facility*
  - *Structure*
    - ✓ Multiple channel, single phase

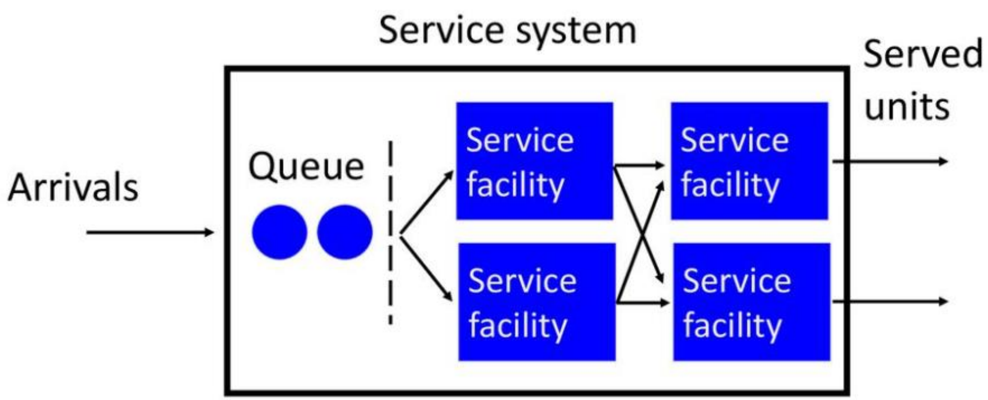


Example: Bank customers wait in ***single line*** for one of several tellers.

Introduction

Structure of the Queueing System

- *Service facility*
  - *Structure*
    - ✓ Multiple channel, multiple phase



Example: At a laundromat, customers use one of several washers, then one of several dryers.

Introduction

Operating Characteristics of a Queueing System

- *Analysis of queueing system involves a study of its different operating characteristics*
  - *Queue length (Lq):* the average number of customers in the queue waiting to get service. This excludes the customer being served.
  - *System length (Ls):* the average number of customers in the system including those waiting as well as those being served.
  - *Waiting time in the queue (Wq):* the average time for which a customer has to wait in the queue to get service.

Introduction

Operating Characteristics of a Queueing System

- *Analysis of queueing system involves a study of its different operating characteristics*
  - *Total time in the system (Ws):* the average total time spent by a customer in the system from the moment he arrives till he leaves the system. It is taken to be the waiting time plus the service time.
  - *Utilization factor (ρ):* it is the proportion of time a server actually spends with the customers. It is also called "traffic intensity".

Introduction

Mathematical Analysis of Queueing Theory

- *Length of the system = Length of the queue + Number of customers receiving services*

  Ls = Lq + Number of customers receiving services

- *Waiting time in the system (Ws) = Waiting time in the queue + Service time*

- *Service utilization factor*

$$\rho = \frac{\text{Mean arrival rate } (\lambda)}{\text{Mean service rate } (\mu)}$$

  – *Mean arrival rate is number of customers arriving to receive service per unit time*

  – *Mean service rate is number of customers served per unit time*

Introduction

Measure of Performance of Queueing System

- *Model 1: Single channel, single phase, infinite population (Poisson distribution)*
  - *User oriented statistics (measures what user experiences)*

$$Ws = \frac{1}{\mu - \lambda}$$

$$Wq = \frac{\lambda}{\mu(\mu - \lambda)}$$

Mean arrival rate (**λ**)
Mean service rate (μ)
Waiting time in the system (Ws)
Waiting time in the queue (Wq)

Introduction

Measure of Performance of Queueing System

- *Model 1: Single channel, single phase, infinite population (Poisson distribution)*
  - *System oriented statistics (measures the characteristics of the system)*

$$Ls = \frac{\lambda}{(\mu - \lambda)}$$

$$Lq = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Mean arrival rate (**λ**)
Mean service rate (μ)
Length of the system (Ls)
Length of the queue (Lq)

Introduction

Measure of Performance of Queueing System

- *Model 1: Single channel, single phase, infinite population (Poisson distribution)*
  - *The probability that the server is busy (traffic density or system utilization)*

$$\rho = \frac{\lambda}{\mu}$$

Mean arrival rate (λ)
Mean service rate (μ)
Service utilization factor (ρ)

Any Questions