

Моделирование систем массового обслуживания

- **Системы массового обслуживания** – это системы, в которых, с одной стороны, возникают массовые запросы (требования) на выполнение каких-либо услуг, с другой – происходит удовлетворение этих запросов.
- Методами теории массового обслуживания могут быть решены многие задачи исследования процессов, происходящих в экономике

- В борьбу за клиента в современной экономике вкладываются огромные средства.
- По оценкам западных экономистов, завоевание фирмой нового клиента обходится ей в 6 раз дороже, чем удержание существующих покупателей.
- А если клиент ушел неудовлетворенным, то на его возвращение приходится потратить в 25 раз больше средств.

- Во многих случаях неудовлетворенность клиента вызвана неудачной организацией его обслуживания (слишком долгое ожидание в очереди, отказ в обслуживании и т.д.).
- Использование теории массового обслуживания позволяет фирме избежать подобных неприятностей

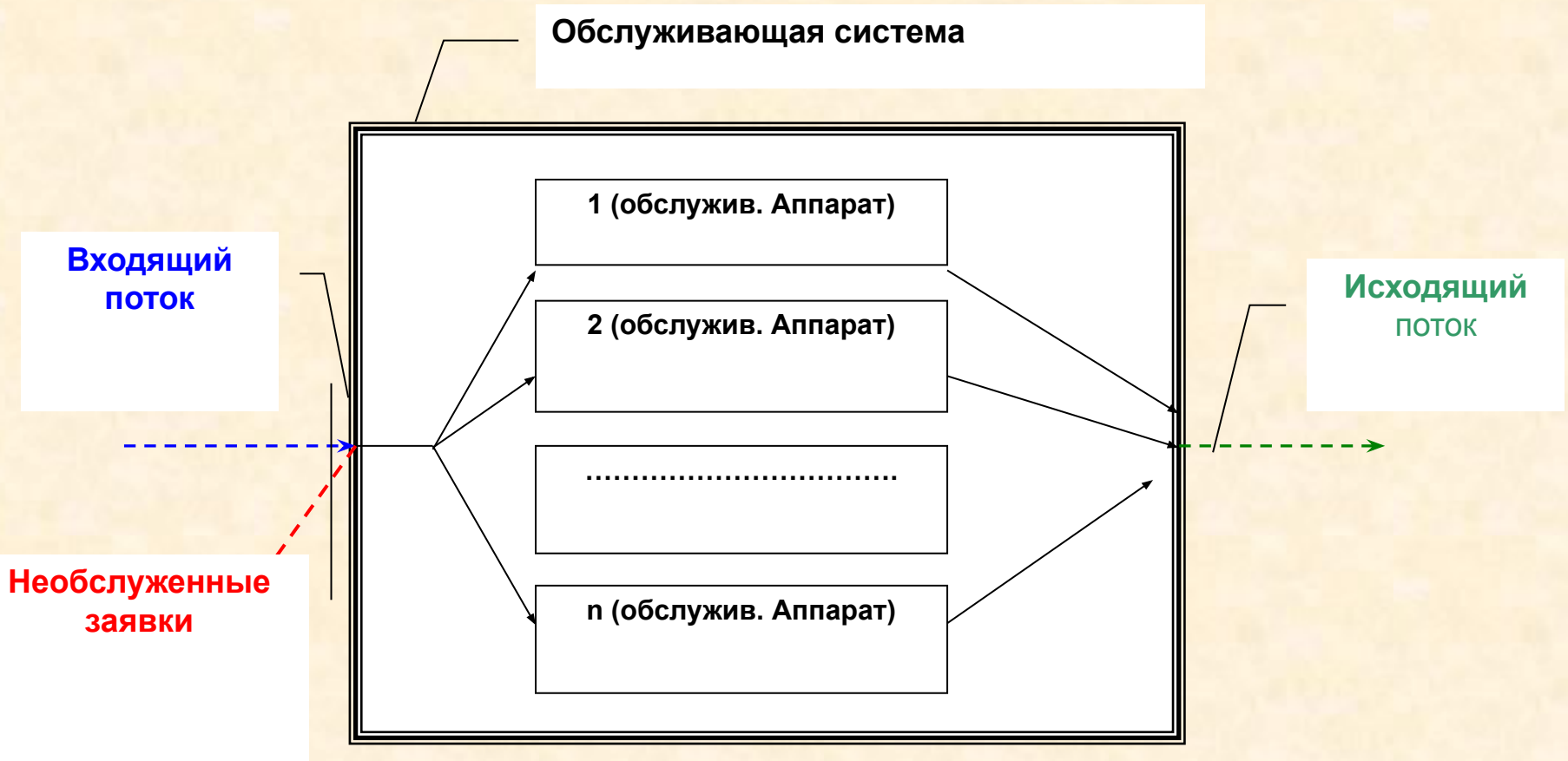
- Основоположником теории массового обслуживания считается датский ученый А. К. Эрланг.
- Являясь сотрудником Копенгагенской телефонной компании, он опубликовал в 1909 году работу «Теория вероятностей и телефонные переговоры», в которой решил ряд задач по теории систем массового обслуживания с отказами.

- Значительный вклад в создание и разработку общей теории массового обслуживания внес выдающийся советский математик Александр Яковлевич Хинчин (1984 – 1959), который предложил сам термин *теория массового обслуживания*.
- В зарубежной литературе чаще используется название *теория очередей*

Системы массового обслуживания, включают следующие элементы:

- Источник требований;
- Входящий поток требований;
- Очередь;
- Обслуживающие устройства (каналы обслуживания);
- Выходящий поток требований.

Структура СМО:



- **Заявками** могут быть производственные и торговые заказы, заявки на ремонт станков, посадку самолетов в аэропорту и заправку автомобилей на автозаправочной станции и т.д.
- **Канал обслуживания** может представлять собой совокупность устройств, этап производственного процесса, аэропорт и т.д.
- Интервалы между последовательными заявками и продолжительность их обслуживания являются **случайными величинами**.

Примеры задач систем МО

1. В торговле Определить оптимальное количество торговых точек данного профиля, численность продавцов, частоту завоза товаров и другие параметры.
2. Склады Установить оптимальное соотношение между числом поступающих на базу требований на обслуживание и числом обслуживающих устройств, при котором суммарные расходы на обслуживание и убытки от простоя транспорта были бы минимальными.

3. Расчет площади складских помещений
 - Складская площадь рассматривается как обслуживающее устройство, а - прибытие транспортных средств под выгрузку — как требование
4. Модель производственной фирмы,
 - Включает несколько цехов, которые последовательно участвуют в процессе производства некоторого изделия, заказы на изготовление изделия поступают в случайные моменты времени

5. Модель управленческого звена фирмы,
- Состоит из начальника и заместителей, которые принимают участие в приеме посетителей.
 - В процессе моделирования требуется обеспечить одинаковую занятость участников процесса
6. Модель бензоколонки
- Количество автомобилей - случайная величина

7. Модели в коммерческой деятельности предприятия.

- Коммерческая деятельность: погрузка товаров, перевозка, разгрузка, хранение, обработка, фасовка, реализация, а также операции с платежными документами, тарой, деньгами, автомашинами, клиентами и т.п.
- Для коммерческой деятельности характерны **массовость поступления товаров, денег, посетителей в случайные моменты времени. Время их обслуживания носит также случайный характер.**

- В качестве *характеристик* эффективности функционирования СМО можно выбрать **три** основные группы
- **1. Показатели эффективности использования СМО:**
 - 1.1. *Абсолютная пропускная способность* СМО – среднее число заявок, которое сможет обслужить СМО в единицу времени.
 - 1.2. *Относительная пропускная способность* СМО – отношение среднего числа заявок, обслуживаемых СМО в единицу времени, к среднему числу поступивших за это же время заявок.

- 1.3. Средняя продолжительность периода занятости СМО.
- 1.4. Коэффициент использования СМО – средняя доля времени, в течение которого СМО занята обслуживанием заявок, и т.п.

- **2. Показатели качества обслуживания заявок:**
- 2.1. Среднее время ожидания заявки в очереди.
- 2.2. Среднее время пребывания заявки в СМО.
- 2.3. Вероятность отказа заявке в обслуживании без ожидания.
- 2.4. Вероятность того, что вновь поступившая заявка немедленно будет принята к обслуживанию.

- 2.5. Закон распределения времени ожидания заявки в очереди.
- 2.6. Закон распределения времени пребывания заявки в СМО.
- 2.7. Среднее число заявок, находящихся в очереди.
- 2.8. Среднее число заявок, находящихся в СМО, и т.п.

- **3. Показатели эффективности функционирования пары «СМО – клиент»,** где под «клиентом» понимают всю совокупность заявок или некий их источник.
- К числу таких показателей относится, например, средний доход, приносимый СМО в единицу времени, и т.п.

- Случайный характер потока заявок и длительности их обслуживания порождает в СМО *случайный процесс*.
- **Случайным процессом** (или *случайной функцией*) называется соответствие, при котором каждому значению аргумента ставится в соответствие случайная величина.

Классификация СМО

1. По месту нахождения источника требований
 - **Разомкнутые** - источник требования находится вне системы
 - **Замкнутые** - источник находится в самой системе

- Примером разомкнутой системы может служить ателье по ремонту телевизоров (магазины, кассы вокзалов, портов ...)
- Здесь неисправные телевизоры — это требования, источник требований находятся вне системы, число требований можно считать **неограниченным**.
- Это система с **неограниченным** потоком требований.

- К замкнутым СМО относится, например, станочный участок, в котором станки являются источником требований на их обслуживание бригадой наладчиков.
- Каждый налаженный станок становится потенциальным источником требований на новую накладку.
- В подобных системах общее число требований **конечно** и чаще всего **постоянно**.

2. По характеру образования очереди

- **С ожиданием** - требование, застав все обслуживающие каналы занятыми, становится в очередь и ожидает, пока не освободится один из обслуживающих каналов
- ◆ **С ограничением на длину очереди** (с ограниченным числом требований в очереди)
- ◆ **С ограничением на время пребывания в очереди** (ограниченным сроком пребывания каждого требования в очереди)
- **С отказами** - требования, поступающие в момент, когда все каналы обслуживания заняты, получают отказ и теряются

- Примером **системы с отказами** является телефонная станция.
- Если вызываемый абонент занят, то требование на соединение с ним получает отказ и теряется.

3. По наличию приоритета

- **Без приоритета:**

- первым пришел - первым ушел,
- последним пришел – первым обслужен,
- случайный отбор

- **С приоритетом:**

- абсолютный приоритет,
- относительный приоритет,
- специальные правила приоритета

4. По количеству каналов

- **Многоканальные:**

- с однородными каналами,
- с неоднородными каналами,
- с параллельно расположенными каналами,
- с последовательно расположенными каналами.

- **Одноканальные.**

Потоки событий

- Под *поток* *событий* понимается *последовательность* *однородных* *событий*, следующих одно за другим в какие-то случайные моменты времени
- (например, *поток* *вызовов* на телефонной *станции*, *поток* *покупателей*, *поток* *заказных писем*, поступающих в почтовое отделение и т. п.)..

- Поток характеризуется **интенсивностью λ** – частотой появления событий или средним числом событий, поступающих в СМО в единицу времени.
- Поток событий называется *регулярным*, если события следуют одно за другим через определенные равные промежутки времени.

- Например, поток изделий на конвейере сборочного цеха (с постоянной скоростью движения) является регулярным.
- Такой поток сравнительно редко встречается в реальных системах, но представляет интерес как предельный случай.
- Типичным для системы массового обслуживания является **случайный поток заявок**.

- Наиболее разработаны методы решения, в которых входящий поток требований является **простейшим (пуассоновским)**.
- Для простейшего потока частота поступления требований в систему подчиняется закону Пуассона, т.е. вероятность поступления за время t k требований задается формулой:

$$P_k(t) = \frac{(\lambda \cdot t)^k}{k!} \cdot e^{-\lambda \cdot t},$$

λ —математическое ожидание числа требований, поступающих в систему в единицу времени

- Простейший поток обладает тремя основными свойствами:

1. ординарности,
2. стационарности
3. отсутствием последдействия

- *Ординарность* потока означает практическую невозможность одновременного поступления двух и более требований.
- Например, достаточно малой является вероятность того, что из группы станков, обслуживаемых бригадой ремонтников, одновременно выйдут из строя сразу несколько станков

- *Стационарным* называется поток, для которого математическое ожидание числа требований, поступающих в систему в единицу времени (λ), не меняется во времени.
- Таким образом, вероятность поступления в систему определенного количества требований в течение заданного промежутка времени Δt зависит от величины промежутка и не зависит от начала его отсчета.

- *Отсутствие последствия* означает, что число требований, поступивших в систему до момента t , не определяет того, сколько требований поступит в систему за промежуток времени от t до $t + \Delta t$.
- Например, если на ткацком станке в данный момент произошел обрыв нити, и он устранен ткачихой, то это не определяет, произойдет новый обрыв на данном станке в следующий момент или нет, тем более это не влияет на вероятность возникновения обрыва на других станках.

Пример.

- На автоматическую телефонную станцию поступает простейший поток вызовов с интенсивностью $\lambda=1,2$ вызовов в минуту.
- Найти вероятность того, что за две минуты: а) не придет ни одного вызова; б) придет ровно один вызов; в) придет хотя бы один вызов.

Решение

- а) Случайная величина X – *число вызовов за две минуты* – *распределена по закону Пуассона* с параметром $\lambda = 1,2 \cdot 2 = 2,4$.
- Вероятность того, что вызовов не будет ($k=0$)

$$P_k(t) = \frac{(\lambda \cdot t)^k}{k!} \cdot e^{-\lambda \cdot t} = P_0(2) =$$
$$= e^{-2,4} \approx 0,091$$

- б) Вероятность одного вызова ($k = 1$)

$$P_1(2) = 2,4 \cdot 0,091 \approx 0,218$$

в) Вероятность хотя бы одного вызова:

$$P(X \geq 1) = 1 - P(X = 0) = \\ 1 - P_0(2) = 1 - 0,091 \approx 0,909$$

- Важная характеристика СМО — **время обслуживания требований в системе.**
- Оно является, как правило, случайной величиной и может быть описано законом распределения. Обычно используется ***экспоненциальный закон распределения времени обслуживания.***
- Вероятность того, что время обслуживания не превосходит некоторой величины ***t***, определяется формулой

$$P(t) = 1 - e^{-\mu t},$$

- где **μ** — величина, обратная среднему времени обслуживания **$\mu = 1/t_{об}$**

СМО с отказами

- В качестве показателей эффективности СМО с отказами будем рассматривать:
- A – абсолютную пропускную способность СМО, т.е. среднее число заявок, обслуживаемых в единицу времени;
- Q – относительную пропускную способность, т.е. среднюю долю пришедших заявок, обслуженных системой (или вероятность того, что пришедшая заявка будет обслужена);

- $P_{отк}$ – вероятность отказа – вероятность того, что заявка покинет СМО необслуженной;
- k – среднее число занятых каналов (для многоканальной системы).

1.Одноканальная система с отказами.

$$Q = \frac{\mu}{\lambda + \mu}$$

$$P_{отк} = \frac{\lambda}{\lambda + \mu}$$

$$A = \lambda \cdot Q$$

Пример.

- В фирму поступает простейший поток заявок на телефонные переговоры с интенсивностью $\lambda = 90$ вызовов в час,
- а средняя продолжительность разговора по телефону = 2 мин.
Определить показатели эффективности работы СМО (телефонной связи) при наличии одного телефонного номера.

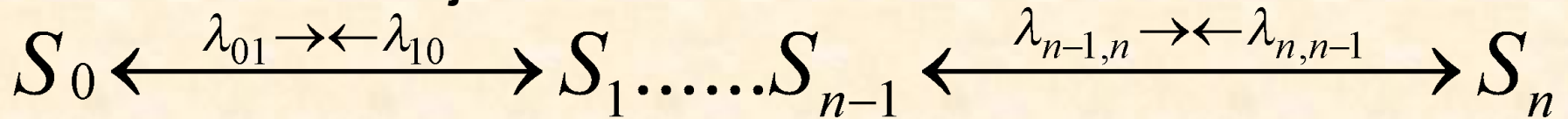
Решение

- Интенсивность потока обслуживаний $\mu = 1/t_{об} = 1/2 = 0,5$ в мин = 30 в час
- $Q = 30/(90+30) = 0,25$ т.е., в среднем около 25 % поступающих заявок осуществляют переговоры по телефону.
- Вероятность отказа составит $P_{отк} = 1 - 0,25 = 0,75$
- Абсолютная пропускная способность $A = 90 \cdot 0,25 = 22,5$ т.е. в среднем в час будут обслужены 22,5 заявки.
- Очевидно, что при наличии только одного телефонного номера СМО будет плохо справляться с потоком заявок.

2. Многоканальная система с отказами (задача Эрланга)

- Эта задача возникла из нужд телефонии и была решена в 1909 г. датским инженером-математиком А.К. Эрлангом.
- Задача ставится так: имеется n каналов (линий связи), на которые поступает поток заявок с интенсивностью λ . Поток обслуживаний каждого канала имеет интенсивность μ .
- Найти предельные вероятности состояний системы и показатели ее эффективности.

- Обозначим λ_{ji} - интенсивность потока событий, переводящий систему из состояния S_i в состояние S_j .



Для S_0 $\lambda_{01}p_0 = \lambda_{10}p_1$

следовательно $\lambda_{12}p_1 = \lambda_{21}p_2$

и $\lambda_{k-1,k}p_{k-1} = \lambda_{k,k-1}p_k$

кроме того $p_0 + p_1 + \dots + p_n = 1$

Решим систему

$$p_1 = \frac{\lambda_{01}}{\lambda_{10}} p_0 \quad p_2 = \frac{\lambda_{12}}{\lambda_{21}} p_1 = \frac{\lambda_{12} \lambda_{01}}{\lambda_{21} \lambda_{10}} p_0$$

$$p_3 = \frac{\lambda_{23} \lambda_{12} \lambda_{01}}{\lambda_{32} \lambda_{21} \lambda_{10}} p_0$$

$$p_k = \frac{\lambda_{k-1,k} \dots \lambda_{12} \lambda_{01}}{\lambda_{k,k-1} \dots \lambda_{21} \lambda_{10}} p_0$$

Таким образом, все вероятности состояний $p_0 p_1 \dots p_n$ выражены через p_0

$$\text{Так как } p_0 + p_1 + \dots + p_n = 1$$

$$p_0 \left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} + \dots + \frac{\lambda_{n-1,n} \dots \lambda_{12}\lambda_{01}}{\lambda_{n,n-1} \dots \lambda_{21}\lambda_{10}} \right) = 1$$

$$p_0 = \left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} + \dots + \frac{\lambda_{n-1,n} \dots \lambda_{12}\lambda_{01}}{\lambda_{n,n-1} \dots \lambda_{21}\lambda_{10}} \right)^{-1}$$

$$\text{или } p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2! \mu^2} + \dots + \frac{\lambda^k}{k! \mu^k} + \dots + \frac{\lambda^n}{n! \mu^n} \right)^{-1}$$

- Введем параметр $\alpha = \lambda/\mu$.
- λ — среднее число требований, поступающих за единицу времени, $1/\mu$ — среднее время обслуживания одним каналом одного требования, тогда $\alpha = \lambda/\mu$, — среднее число каналов, которое необходимо иметь, чтобы обслуживать в единицу времени все поступающие требования.
- $\alpha/n < 1$ означает, что число обслуживающих каналов больше среднего числа каналов, необходимых для того, чтобы обслужить все поступившие требования.

Важнейшие характеристики

1. Вероятность того, что все обслуживающие каналы свободны

$$p_0 = \left(1 + \alpha + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^k}{k!} + \dots + \frac{\alpha^n}{n!}\right)^{-1}$$

2. Вероятность того, что все обслуживающие каналы заняты

$$P_{отк} = \frac{\alpha^n}{n!} \cdot p_0$$

2 Вероятность того, что занято ровно k обслуживающих каналов при условии, что общее число требований, находящихся на обслуживании, не превосходит числа обслуживающих аппаратов

$$P_k = \frac{\alpha^k}{k!} P_0 \quad \text{при} \quad 1 \leq k \leq n$$

4. Вероятность того, что занято ровно k обслуживающих каналов при условии, что общее число требований, находящихся на обслуживании, превосходит число обслуживающих аппаратов

$$P_k = \frac{\alpha^k}{n! n^{k-n}} P_0 \quad \text{при} \quad k \geq n$$

5. Относительная пропускная способность – вероятность того, что заявка будет обслужена

$$Q = 1 - P_{отк} = 1 - \frac{\alpha^n}{n!} p_0$$

6. Абсолютная пропускная способность получается умножением интенсивности заявок на относительную пропускную способность

$$A = \lambda \cdot Q = \lambda \left(1 - \frac{\alpha^n}{n!} p_0 \right)$$

7. Среднее число каналов занятых обслуживанием

$$\bar{k} = A / \mu = \alpha \left(1 - \frac{\alpha^n}{n!} p_0 \right)$$

8. Среднее время ожидания начала обслуживания в системе

$$\bar{t}_{ож} = \frac{P_n}{\mu(n - \alpha)}$$

9. Среднее число свободных от обслуживания каналов:

$$\overline{N}_0 = n - \overline{k}$$

10. Коэффициент простоя каналов:

$$K_{np} = \frac{\overline{N}_0}{n}$$

11. Коэффициент загрузки каналов

$$K_3 = \frac{\overline{k}}{n}$$

Примеры

- 1. В условиях предыдущего примера определить оптимальное число телефонных номеров в фирме, если условием оптимальности считать удовлетворение из каждых 100 заявок на переговоры в среднем **не менее 90** заявок.

Решение

- $\mu = 1/t_{об} = 1/2 = 0,5$ в мин = 30 в час
- Интенсивность нагрузки канала $\alpha = 90/30 = 3$ т.е. за время среднего (по продолжительности телефонного разговора = 2 мин) поступает в среднем 3 заявки на переговоры.

- Будем постепенно увеличивать число каналов (телефонных номеров) $n = 2, 3, 4, \dots$ и определим для получаемой n - канальной СМО характеристики обслуживания.

Показатели эффективности	Обозначение	Число каналов (телефонных номеров)					
		1	2	3	4	5	6
Относительная пропускная способность	Q	0,25	0,47	0,65	0,79	0,90	0,95
Абсолютная пропускная способность	A	22,5	42,3	58,8	71,5	80,1	85,3

- По условию оптимальности $Q \geq 0,9$, следовательно, в фирме необходимо установить 5 телефонных номеров (в этом случае $Q = 0,90$).
- При этом в час будут обслуживаться в среднем **80** заявок ($A = 80,1$),
- а среднее число занятых телефонных номеров (каналов) $k = A \mu = 80,1 \cdot 30 \approx$
2,67 .

- *Содержание* каждого канала в единицу времени обходится в какую-то сумму. Вместе с тем, каждая обслуженная заявка приносит какой-то доход.
- Умножая этот доход на среднее число заявок, *обслуживаемых в единицу времени*, мы получим **средний доход от СМО в единицу времени.**

- Естественно, при увеличении числа каналов этот доход растёт, но растут и расходы, связанные с содержанием каналов.
- Что перевесит – увеличение доходов или расходов? Это зависит от условий операции, т.е. от «платы за обслуживание заявки» и от стоимости содержания канала.
- Зная эти величины, можно найти **оптимальное число каналов, наиболее экономически эффективное**

Пример 2

- Пусть филиал фирмы по ремонту радиоаппаратуры имеет $n = 5$ мастеров. В среднем в течение рабочего дня от населения поступает в ремонт $\lambda = 10$ радиоаппаратов.
- Общее число радиоаппаратов, находящихся в эксплуатации у населения, очень велико, и они независимо друг от друга в различное время выходят из строя. Поэтому есть основания полагать, что поток заявок на ремонт аппаратуры является случайным пуассоновским.

- Каждый аппарат в зависимости от характера неисправности требует различного случайного времени на ремонт. Статистика показала, что время ремонта подчиняется экспоненциальному закону;
- В среднем в течение рабочего дня каждый из мастеров успевает отремонтировать $\mu = 2,5$ радиоаппарата.
- Требуется оценить работу филиала фирмы по ремонту радиоаппаратуры, рассчитав ряд основных характеристик данной СМО.
- За единицу времени принимаем 1 рабочий день (7 часов).

- 1. Определим параметр
- $\alpha = \lambda/\mu = 10/2,5 = 4$, так как $\alpha < n$, то очередь не может расти безгранично.
- 2. Вероятность того, что все мастера свободны от ремонта аппаратуры
- $P = 1/[1 + 4 + 4^2/2! + 4^3/3! + 4^4/4! + 4^5/5!] = 0,008$.
- 3. Вероятность того, что все мастера заняты ремонтом

$$P_n = \frac{4^5 \cdot 0,008}{5!} = 0,068$$

Это означает, что
6,8% времени
мастера полностью
загружены работой

- 4. Среднее время обслуживания (ремонта) одного аппарата
- $t_{об} = 1/\mu = 7/2,5 = 2,8$ ч/аппарат
(при условии семичасового рабочего дня).
- 5. В среднем время ожидания начала ремонта равно

$$\bar{t}_{ож} = \frac{0,068 \cdot 2,8}{5 - 4} = 0,19 \text{ ч.}$$

Среднее число мастеров занятых обслуживанием

$$\bar{k} = \alpha \left(1 - \frac{\alpha^n}{n!} p_0\right) = 4 \left(1 - \frac{4^5}{5!} 0,008\right) = 3,72$$

Среднее число мастеров, свободных от работы

$$\bar{N}_o = n - \bar{k} = 5 - 3,72 = 1,28$$

Таким образом, в среднем в течение рабочего дня ремонтом заняты четыре мастера из пяти

СМО с ожиданием (с очередью)

- 1. Одноканальная СМО с ожиданием и ограничением на длину очереди**
 - Часто встречаются одноканальные СМО с очередью (врач, обслуживающий пациентов; кассир, выдающий зарплату и т.д.).
 - Определим основные характеристики одноканальной СМО с ожиданием:

Рассмотрим одноканальную СМО, на вход которой поступает простейший поток заявок с интенсивностью λ

- Предположим, что поток обслуживаний также простейший с интенсивностью μ . Это означает, что непрерывно занятый канал обслуживает в среднем μ заявок в единицу времени.
- Заявка, поступившая в СМО в момент, когда канал занят, становится в очередь и ожидает обслуживания.
- Интенсивность нагрузки канала: $\alpha = \lambda / \mu$

- Далее предполагаем, что в системе имеется ограничение на длину очереди, предполагаем, что в очереди могут находиться максимум ***m*** ($m \geq 1$) заявок.
- *Поэтому заявка, пришедшая на вход СМО, в момент, когда в очереди уже стоят ***m*** заявок, получает отказ и покидает систему необслуженной.*

Основные характеристики одноканальной СМО с ожиданием

$$P_0 = \frac{1 - \alpha}{1 - \alpha^{m+2}}$$

$$P_{отк} = P_{m+1} = \begin{cases} \frac{\alpha^{m+1}(1 - \alpha)}{1 - \alpha^{m+2}}, & \text{если } \alpha \neq 1 \\ \frac{1}{m+2}, & \text{если } \alpha = 1 \end{cases}$$

- **Относительная пропускная способность**, или доля обслуживаемых заявок, *совпадает со средней долей* не получивших отказ заявок, поскольку заявка попавшая в очередь будет обслужена.

$$Q = 1 - P_{отк} = \begin{cases} \frac{1 - \alpha^{m+1}}{1 - \alpha^{m+2}}, & \text{если } \alpha \neq 1 \\ \frac{m+1}{m+2}, & \text{если } \alpha = 1 \end{cases}$$

- Абсолютная пропускная способность системы $A = \lambda Q$
- Среднее число заявок в очереди $L_{оч}$ определяется как математическое ожидание случайной величины – числа заявок, стоящих в очереди:

$$L_{оч} = \begin{cases} \frac{\alpha^2 (1 - \alpha^m (m + 1 - \alpha m))}{(1 - \alpha)(1 - \alpha^{m+2})}, & \text{если } \alpha \neq 1 \\ \frac{m(m + 1)}{2m(m + 2)}, & \text{если } \alpha = 1 \end{cases}$$

- Важной характеристикой СМО с ожиданием является **среднее время ожидания заявки в очереди $T_{оч}$** , которая называется формулой Литтла

$$\overline{T_{оч}} = \frac{L_{оч}}{\lambda}$$

т.е. среднее время ожидания заявки в очереди равно среднему числу заявок в очереди ,деленному на интенсивность λ входящего потока заявок

Пример

- На АЗС имеется **одна** колонка. Площадка, на которой машины ожидают заправку, может вместить не более **трех** машин одновременно, и если она занята, то очередная машина, прибывшая к станции, в очередь не становится, а проезжает на соседнюю АЗС.
- В среднем машины прибывают на станцию каждые **2** мин. Процесс заправки одной машины продолжается в среднем **2,5** мин.
- Определить основные характеристики системы.

- **Решение.** Математической моделью данной АЗС является одноканальная СМО с ожиданием и ограничением на длину очереди ($m = 3$).
- *Предполагается, что поток машин, подъезжающих к АЗС для заправки, и поток обслуживаний – простейшие*
- Поскольку машины прибывают в среднем через каждые 2 мин, то интенсивность входящего потока равна $\lambda = 1 / 2 = 0,5$ (машин в минуту)

- Среднее время обслуживания одной машины $T_{об} = 2,5$ мин, следовательно, интенсивность потока обслуживаний $\mu = 1 / 2,5 = 0,4$ (машины в минуту).
- Определяем интенсивность нагрузки канала: $\alpha = \lambda / \mu = 0,5 / 0,4 = 1,25$
- Вычисляем вероятность отказа

$$P_{отк} = \frac{\alpha^4 (1 - \alpha)}{1 - \alpha^5} \approx 0,297$$

- Относительная пропускная способность

$$Q = 1 - P_{отк} \approx 1 - 0,297 = 0,703$$

- Абсолютная пропускная способность $A = \lambda Q \approx 0,5 \cdot 0,703 \approx 0,352$ способ

- Среднее число машин, ожидающих в очереди на заправку

$$L_{оч} = \frac{\alpha^2 (1 - \alpha^3 (m + 1 - 3\alpha))}{(1 - \alpha)(1 - \alpha^5)} \approx 1,559$$

Среднее время ожидания машины в очереди находим по формуле Литтла

$$\overline{T_{оч}} = \frac{L_{оч}}{\lambda} \approx \frac{1,559}{0,5} = 3,118$$

- Таким образом, из анализа работы СМО следует, что из каждых 100 подъезжающих машин 30 получают отказ ($P_{отк} \approx 29,7\%$), т.е. обслуживаются 2/3 заявок.
- Поэтому необходимо либо сократить время обслуживания одной машины (увеличить интенсивность потока обслуживаний), либо увеличить число колонок, либо увеличить площадку для ожидания.

- Оптимальное решение принимается с учетом затрат, связанных соответственно с увеличением штата обслуживающего персонала (увеличение производительности канала),
- с расширением площадки для ожидания или приобретением дополнительной колонки, и потерь, связанных с потерей заявок на обслуживание.

Многоканальная система с ограниченной очередью

Вероятность простоя каналов

$$p_0 = \left(1 + \alpha + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^n}{n!} + \frac{\alpha^{n+1}}{n n!} + \frac{\alpha^{n+2}}{n^2 n!} + \dots \frac{\alpha^{n+m}}{n^m n!} \right)^{-1}$$

Вероятность отказа в обслуживании заявки (отказ произойдет в случае, если все каналы заняты и в очереди находятся m

$$P_{отк} = P_{n+m} = \frac{\alpha^{n+m}}{n^m n!} P_0$$

$$P_k = \frac{\alpha^k}{k!} P_0 \quad k = 1, 2, \dots, n$$

m — длина
очереди

$$P_{n+i} = \frac{\alpha^{n+i}}{n^i n!} P_0 \quad i = 1, 2, \dots, m$$

- Относительная пропускная способность

$$Q = 1 - P_{\text{отк}}$$

Абсолютная пропускная способность

$$A = Q * \lambda$$

- ПРИМЕР
- Междугородный переговорный пункт имеет четыре телефонных аппарата. В среднем за сутки поступает 320 заявок на переговоры. Средняя длительность переговоров составляет 5 мин. Длина очереди не должна превышать 6 абонентов. Потoki заявок и обслуживаний простейшие.
- Определить характеристики обслуживания переговорного пункта в стационарном режиме (вероятность простоя каналов, вероятность отказа, среднее число занятых каналов, абсолютную пропускную способность, относительную пропускную способность, среднее время заявки под обслуживанием).

РЕШЕНИЕ.

- Имеем систему массового обслуживания (СМО) с четырьмя каналами (четыре аппарата), с ожиданием и ограниченной очередью (6 мест).
- Получаем параметры $n = 4$ (число каналов), $m = 6$ (число мест в очереди), $\Lambda = 320/60 \cdot 24 = 2/9$ (интенсивность входящего потока, заявок в минуту), $\mu = 1/5$ (интенсивность потока обслуживания, одна заявка за 5 минут). $\alpha = 2/9 : 1/5 = 1,1$
- Определим характеристики работы данной СМО в предельном режиме

Вероятность простоя каналов

$$p_0 = \left(1 + 1,1 + \frac{1,1^2}{2!} + \frac{1,1^3}{3!} + \frac{1,1^4}{4!} + \frac{1,1^5}{4*4!} + \frac{1,1^6}{4^2*4!} + \frac{1,1^7}{4^3*4!} + \frac{1,1^8}{4^4*4!} + \frac{1,1^9}{4^5*4!} + \frac{1,1^{10}}{4^6*4!} \right)^{-1} \approx 0,328$$

- Вероятность отказа в обслуживании

.

$$P_{отк} = P_{n+m} = \frac{1,1^{10}}{4^6 4!} 0,328 = 0,0000009$$

- Относительная пропускная способность
(вероятность обслуживания)

$$Q = 1 - 0,000009 = 0,99999$$

- Абсолютная пропускная способность

$$A = 0,99999 * 2/9 = 0,22222$$

- Среднее число занятых каналов

$$N = A / \mu = 0,2222 * 5 = 1,1111$$

- Среднее время заявки под обслуживанием

$$T = N / \lambda = 1,1111 / (2/9) = 4,9995 \text{ минут}$$

Одноканальная СМО с неограниченным ожиданием

- Если $\lambda > \mu$ ($\alpha > 1$), т.е. среднее число заявок, поступивших в систему за единицу времени, больше среднего числа обслуживаемых заявок за то же время при непрерывно работающем канале, то очевидно, что **очередь неограниченно растёт**.
- В этом случае предельный режим не устанавливается и предельных вероятностей состояний не существует (они равны нулю).

- В случае $\lambda = \mu$ ($\alpha = 1$) при условии, что входящий поток заявок и поток обслуживаний регулярные (заявки поступают через равные интервалы времени, и время обслуживания одной заявки является постоянным, равным интервалу времени между поступлениями заявок), **очереди не будет и канал будет обслуживать заявки непрерывно.**
- Но если входящий поток или поток обслуживаний становится случайным, очередь начинает расти до бесконечности.

- Поэтому далее при рассмотрении указанных систем будем предполагать, что $\lambda < \mu$, т.е. $\alpha < 1$.
- При этом условии с течением времени устанавливается предельный режим, и предельные вероятности состояний существуют.

- При отсутствии ограничений на очередь каждая заявка, поступившая в СМО, будет обслужена. Поэтому вероятность отказа равна нулю **$P_{отк}=0$**
- Следовательно, вероятность того, что поступившая заявка будет принята в систему, так же как и относительная пропускная способность *равна единице* **$Q = 1 - P_{отк} = 1$**

- Тогда для абсолютной пропускной способности A (*и интенсивности выходящего потока*) будем иметь:

$A = \lambda Q = \lambda$, т.е. *интенсивности входящего и выходящего потоков совпадают*

Среднее число заявок в очереди

$$L_{оч} = \frac{\alpha^2}{(1 - \alpha)}$$

- Среднее время ожидания заявки в очереди по формуле Литтла равно

$$\overline{T}_{оч} = \frac{L_{оч}}{\lambda} = \frac{\alpha}{\mu(1-\alpha)}$$

среднее время пребывания заявки в СМО складывается из среднего времени заявки в очереди и среднего времени обслуживания заявки :

$$\overline{T}_{СМО} = \overline{T}_{оч} + \overline{T}_{об} = \frac{\alpha}{(1-\alpha)}$$

- **Пример.** В парикмахерской работает только один мужской мастер. Среднее время стрижки одного клиента составляет 20 мин. Клиенты в среднем приходят каждые 25 мин.
- Средняя стоимость стрижки составляет 60 руб. Как в первую смену с 9 до 15, так и во вторую – с 15 до 21, работают по одному мастеру.
- . Определить ежедневный «чистый» доход каждого мастера, если он получает только 30% от выручки (остальное уходит на оплату аренды, налоги, и проч.).

- **Решение.** Интенсивность входящего потока $\lambda = 2,4$ клиента/ч,
- интенсивность потока обслуживаний $\mu = 1 / 20 \text{ мин} = 1 / (1/3) \text{ часа} = 3$ клиента в час
- интенсивность нагрузки (канала) мастера $\alpha = \lambda / \mu = 0,8$
- долю времени (вероятность) простоя мастера $P_0 = 1 - \alpha = 1 - 0,8 = 0,2$
- вероятность того, что мастер занят работой $P_{\text{зан}} = 1 - P_0 = 1 - 0,2 = 0,8$

- Среднее число клиентов в очереди

$$L_{оч} = \frac{\alpha^2}{(1-\alpha)} = \frac{0,8^2}{1-0,8} = 3,2 \text{ клиента}$$

Среднее время ожидания в очереди

$$\overline{T_{оч}} = \frac{L_{оч}}{\lambda} = \frac{3,2}{2,4} = 1,34 \text{ минуты}$$

Среднее время пребывания в парикмахерской

$$\overline{T_{CMO}} = \overline{T_{оч}} + \overline{T_{об}} = 1,34 + 20 = 21,34 \text{ мин}$$

- Система работает вполне удовлетворительно. Поскольку $\alpha < 1$, то режим работы системы устойчивый, 20% рабочего времени мастер не занят,
- а остальные 80% времени занят работой, длина очереди 3,2 клиента небольшая, а среднее время пребывания клиента в парикмахерской всего 21,34 мин.

- Каждый мастер занимается обслуживанием клиентов в среднем ежедневно в течение

$$0,8(15-9)=4,8 \text{ часа}=288 \text{ мин.}$$

- За это время он обслужит $288 : 20 = 14,4$ клиента, поэтому ежедневная выручка в среднем составит $14,4 \cdot 60 = 864$ руб.
- Ежедневный «чистый» доход каждого мастера в среднем составляет $864 \cdot 0,3 = 259,2$ руб.

Многоканальная СМО с неограниченной очередью

- Предположим, что $\alpha/n < 1$ выполнено и предельные вероятности существуют.
- 1. Вероятность того, что все обслуживающие каналы свободны

$$p_0 = \left(1 + \alpha + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^k}{k!} + \dots + \frac{\alpha^n}{n!} + \frac{\alpha^{n+1}}{n!(n-\alpha)} \right)^{-1}$$

2 Вероятность того, что занято ровно k обслуживающих каналов при условии, что общее число требований, находящихся на обслуживании, не превосходит числа обслуживающих аппаратов

$$P_k = \frac{\alpha^k}{k!} P_0 \quad \text{при} \quad 1 \leq k \leq n$$

4. Вероятность того, что занято ровно k обслуживающих каналов при условии, что общее число требований, находящихся на обслуживании, превосходит число обслуживающих аппаратов

$$P_{n+r} = \frac{\alpha^{n+r}}{n! n^r} P_0 \quad \text{при} \quad k \geq n$$

Среднее число занятых каналов

$$\bar{k} = \lambda / \mu = \alpha$$

Среднее число заявок в очереди

$$L_{оч} = \frac{\alpha^{n+1} P_0}{n \cdot n! (1 - \alpha/n)^2}$$

- Среднее число заявок в системе

$$L_{\text{сист}} = L_{\text{оч}} + \alpha$$

- Среднее время заявки в очереди

$$W_{\text{оч}} = 1/\lambda L_{\text{оч}}$$

- Среднее время заявки в системе

$$W_{\text{сист}} = 1/\lambda L_{\text{сист}}$$

Пример

- Железнодорожная касса с двумя кассирами (очередь одна $n=2$), $\lambda=0,9$ (пассажира в минуту), кассир тратит на обслуживание одного пассажира в среднем 2 минуты.
- **РЕШЕНИЕ** $\mu=1/2=0,5$ $\alpha=\lambda/\mu=1,8$.
- Т.к. $\alpha/2=1,8/2=0,9<1$, то финальные (предельные) вероятности существуют.

$$p_0 = (1 + 1,8 + \frac{1,8^2}{2!})^{-1} \approx 0,0525$$

$$L_{оч} = 7,68$$

$$W_{оч} \approx 8,54 \text{ мин}$$