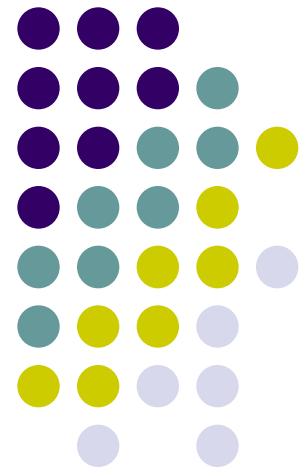


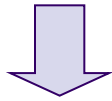
3. Аналитическое моделирование непрерывно- стохастических систем

**На примере систем
массового обслуживания
(Q -схемы)**





Системы массового обслуживания
(англ. *queuing system*)



Q-схемы.

Модели разработаны в теории массового обслуживания (ТМО) для формализации процессов функционирования систем, которые по своей сути являются процессами обслуживания.



ТМО возникла как раздел теории вероятностей.

Первые работы – в начале XX в.

Были вызваны потребностями практики, в частности широким развитием телефонных сетей.



В работах по ТМО часто используется терминология, заимствованная из телефонии: требования, вызовы, заявки, каналы связи, и т. п.



Позднее – понимание:

общие математические модели, исследуемые как модели телефонии, могут описывать и другие жизненные явления.

Задачи ТМО связаны с исследованием любых операций, состоящих из многих однородных элементарных операций, на осуществление которых влияют случайные факторы.



Примеры. **В значительной мере создали базу ТМО**

1. В вычислительную систему, управляющую технологическим процессом, в случайные моменты времени поступают сигналы от датчиков, связанных с управляемым объектом. Каждый сигнал требует обработки в течение некоторого случайного времени (может зависеть от содержания сигнала).

Работу системы можно рассматривать как операцию массового обслуживания, состоящую из элементарных операций – обработки отдельных сигналов.

Вопрос: способна ли вычислительная система с заданными параметрами справиться с обработкой всех поступающих сигналов?



2. Автоматическая телефонная станция (АТС) обслуживает некоторое число абонентов, выходящих на связь в случайные моменты времени. Абонент, пытающийся позвонить (послать вызов) в момент занятости всех линий, получает отказ.

В данном случае

операция массового обслуживания – процесс функционирования АТС в течение длительного времени;

ее основная характеристика – вероятность отказа при вызове;

элементарная операция – отдельный телефонный разговор.



3. В сборочный цех поступают для сборки детали различных видов. При нехватке хотя бы одного вида деталей производство останавливается; избыточные детали поступают в бункера определенной вместимости.

И на процесс поступления деталей, и на время сборки изделия влияют случайные факторы.

Элементарная операция – сборка одного изделия из готового комплекта деталей.

Вопросы: какова вероятность простоя
производственной линии?
какова вероятность переполнения бункеров?



4. В морской порт прибывают суда не строго по графику, а со случайными отклонениями. Имеется несколько погрузочно-разгрузочных площадок с соответствующим оборудованием.

Элементарная операция – процесс разгрузки и погрузки одного судна.

Вопрос: чему равно среднее время от момента прибытия судна до окончания его разгрузки и погрузки?

Для иллюстрации влияния различных случайных факторов на выполнение элементарных операций – данные из статьи Б.В. Гнеденко и М.Н. Зубкова «Об определении оптимального числа причалов».

График прихода судов и его выполнение (один и тот же месяц 1962 и 1963 г.).



Название судна	Дата прихода по графику	Дата фактического прихода
1962 г.		
Авамери	От 8 до 20	23
Либерейтор	5 или 6	Не пришло
Маруба	От 25 до 8 след. месяца	Не пришло
Атлантида	28	Не пришло
Иван Ползунов	12	25
Пулково	Не ожидалось	27
Алапаевск	От 12 до 15	Не пришло
1963 г.		
Арагви	4	5
Вормси	6	10
Коломна	7	30
Грибоедов	13	Не пришло
Медногорск	15	Не пришло
Ижевск	22	18
Кировск	24	Не пришло
Репино	Не ожидалось	30



Подобная картина

не только в этом порту;
не только в этом месяце.

Например: в другой порт за месяц
не прибыло 17 судов, приход которых намечался
графиком;
прибыло 22 судна, не предусмотренных
графиком.

При этом: общее число судов, приходящих в порт,
близко к плановому.

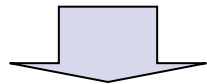


График – не закон движения судов, а примерное
указание на интенсивность их возможного движения в
рассматриваемый период.



Фактические наблюдения о длительности погрузочно-разгрузочных операций.

Время, затраченное на обработку судна в порту (в часах).

Длительность погрузочно-разгрузочных операций	Число случаев	Длительность погрузочно-разгрузочных операций	Число случаев
0 – 5	17	80 – 85	8
5 – 10	23	85 – 90	6
10 – 15	17	90 – 95	12
15 – 20	24	95 – 100	4
20 – 25	25	100 – 110	12
25 – 30	17	110 – 120	10
30 – 35	14	120 – 130	11
35 – 40	13	130 – 140	6
40 – 45	9	140 – 150	10
45 – 50	20	150 – 160	5
50 – 55	17	160 – 180	10
55 – 60	19	180 – 200	7
60 – 65	14	200 – 250	10
65 – 70	16	250 – 300	7
70 – 75	15	300 – 400	2
75 – 80	8	400 – 600	1



При планировании работы порта необходимо учитывать

- случайный разброс в моментах прихода судов;
- случайный разброс длительности погрузочно-разгрузочных операций.

При этом:

для решения экономических и технических проблем, связанных с эксплуатацией порта, недостаточно элементарных арифметических приемов (исходя из средних значений).



Характерные особенности математических моделей, исследуемых в ТМО

- Наличие некоторого *потока* (протяженного во времени) однородных абстрактных объектов (заявок, требований, событий).

Существенными являются моменты появления этих объектов.



- ❑ Наличие некоторых правил – *дисциплины обслуживания*.

Включает:

- определение числа объектов, которые могут одновременно обслуживаться в системе;
- определение числа объектов, которые могут ожидать начала обслуживания;
- определение порядка, в котором ожидающие обслуживания объекты поступают на обслуживание;
- определение порядка, в котором объекты покидают систему и др.

Структура
системы

- ❑ Моменты появления объектов и продолжительность обслуживания являются случайными величинами.



Математическая модель системы массового обслуживания (СМО) должна включать

- описание свойств входящего потока однородных событий,
- описание структуры исследуемой системы,
- описание дисциплины и характеристик процесса обслуживания.



Основные элементы модели СМО

Система массового обслуживания – это система, в которой выполняется последовательность (элементарных) операций.

Операции могут быть *реальными* или *фиктивными*.



Реальные операции – это операции, которые действительно выполняются и требуют определенных затрат работы.

Фиктивные операции в действительности не существуют и вводятся в математическую модель СМО для удобства ее построения.

Например: операция ожидания требования, окончание которой означает поступление в систему некоторого требования.

Всегда предполагается: число различных типов операций для данной СМО конечно.



Реальные операции выполняются *приборами* (*каналами*) обслуживания.

Если не оговорено противное, то предполагается: обслуживающий прибор может одновременно выполнять только одну операцию.

Количество приборов (каналов) в СМО конечно (реже – счетно).

СМО, содержащая один прибор (канал), называется *одноканальной* (*однолинейной*);

СМО, содержащая не менее двух приборов (каналов), – *многоканальной* (*многолинейной*).



Для каждого прибора указывается один или несколько типов операций, которые может выполнять данный прибор.

Если прибор выполняет операцию, то говорят, что он *занят* (обслуживанием);
в противном случае прибор *свободен* (от обслуживания).



Требования на обслуживание могут быть внешними (входящими) и внутренними.

Внешнее требование поступает извне системы в момент каждого события входящего потока требований.

Внутреннее требование может возникать в момент окончания реальной или фиктивной операции.

Множество моментов поступления в систему требований называется *входным потоком* данной СМО.

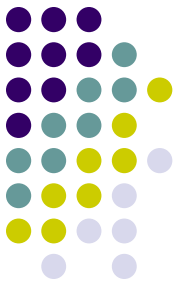


При построении автономной модели СМО (воздействия внешних факторов не включаются в модель) вводится дополнительный объект – *источник требований* (заявок).

Источник требований (потока требований) – это прибор, постоянно выполняющий фиктивные операции «ожидания требования».

В момент окончания каждой такой операции источник посылает требование.

Источник может иметь конечную или бесконечную мощность.



Источник конечной мощности ограничивает число требований на обслуживание, поступающих в СМО.

Пример.

В цехе, располагающем **N** станками, суммарное количество заявок на их ремонт не превышает **N** .

Источник бесконечной мощности не ограничивает число требований на обслуживание, поступающих в СМО.

Пример.

Звонки, поступающие на автоматическую телефонную станцию (АТС).



Очередью называется совокупность требований, ожидающих обслуживания в момент, когда приборы заняты обслуживанием других требований. Требования, ожидающие обслуживания, находятся в *накопителе*.

Накопитель характеризуется *емкостью* – максимальным числом требований, которые могут присутствовать в нем одновременно.

Емкость накопителя может быть

- конечной,
- бесконечной.



Дисциплина очереди – принцип, определяющий порядок, в соответствии с которым из очереди выбирается требование (заявка) для обслуживания.

Наиболее известные принципы определяются следующими правилами.

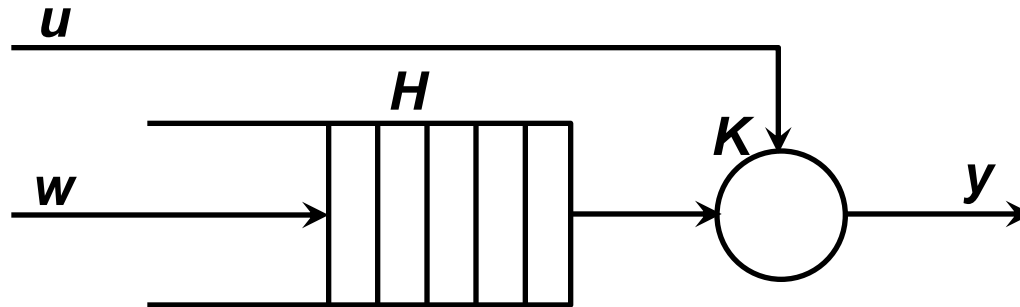
- «Первым пришел – первым обслуживаешься» (*FIFO – First-In-First-Out*).
- «Последним пришел – первым обслуживаешься» (*LIFO – Last-In-First-Out*).
- Случайный выбор требований (*SIRO – Service-In-Random-Out*).
- Выбор требований в соответствии с заданным приоритетом.



Предмет ТМО – установление зависимости между
характером входного потока требований,
количеством каналов обслуживания,
характеристиками каждого канала
и
эффективностью (показателями качества)
обслуживания.



Графическое изображение СМО



H – накопитель;

K – канал обслуживания;

w – поток заявок (требований);

u – поток обслуживаний;

y – выходной поток заявок

(заявки, обслуженные каналом K , и заявки, покинувшие СМО не обслуженными).



Несколько СМО могут быть объединены в *сеть массового обслуживания* (СеМО), когда после обслуживания в одной СМО требование нуждается в обслуживании в другой СМО.

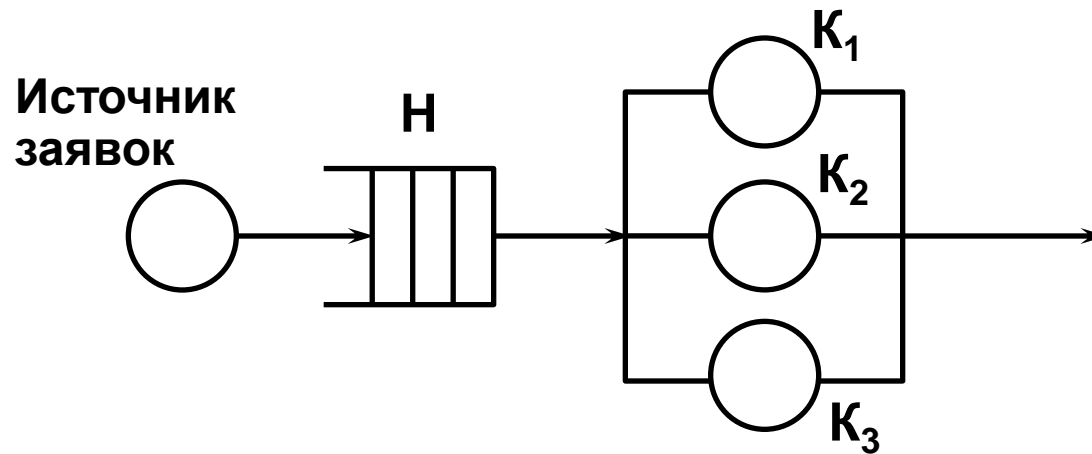
В таких случаях употребляется также термин «многофазная СМО»

В такой ситуации имеют значения свойства не только входного, но и выходного потока из СМО: поток, выходной для одной СМО, может оказаться входным (или частью входного потока) для другой СМО сети.



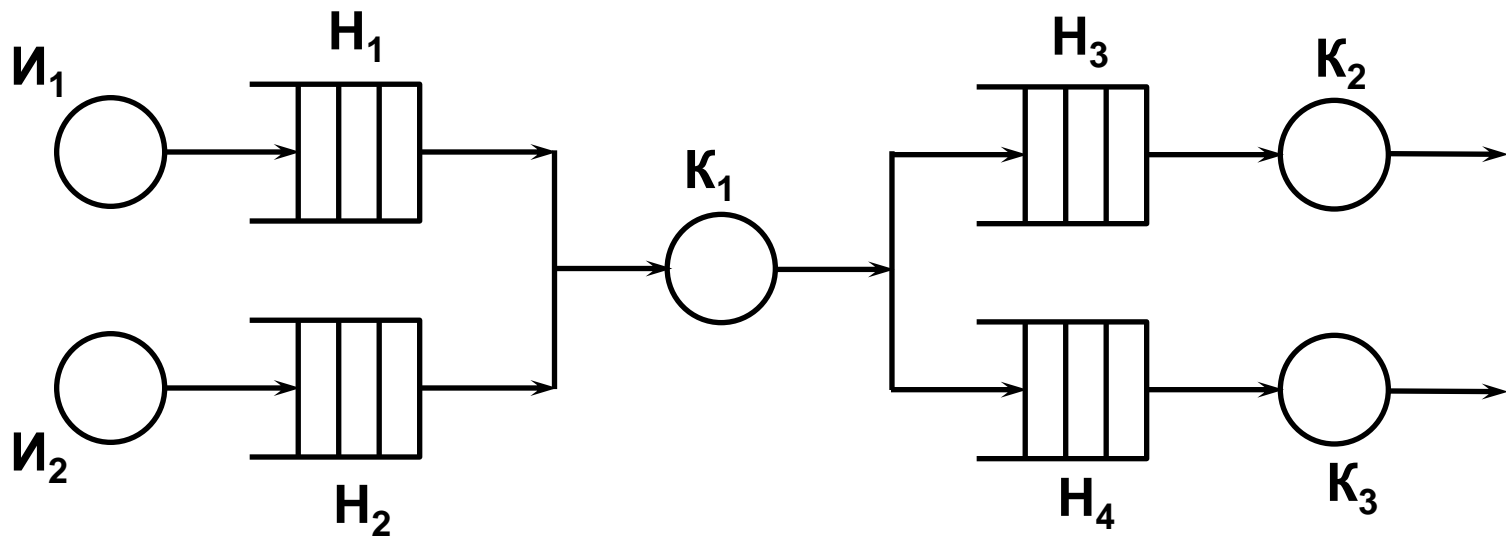
Примеры.

1. Трехканальная однофазная СМО.





2. Двухфазная СМО.





Случайный процесс со счетным множеством состояний

Случайный процесс, протекающий в СМО, состоит в том, что система в случайные моменты времени переходит из одного состояния в другое:

изменяется

число занятых каналов,

число требований (заявок), ожидающих обслуживания, и т. д.

СМО — дискретная система с конечным или счетным множеством состояний.



Пусть исследуемая система в любой момент времени t может находиться в одном из состояний

$$z_1, z_2, \dots, z_n, \dots$$

Счетное множество

Обозначим

$p_k(t)$ – вероятность того, что в момент времени t система находится в состоянии z_k ,
 $k = 1, 2, \dots, n, \dots$

Для любого t

$$\sum_k p_k(t) = 1.$$



Случайные процессы со счетным множеством состояний:

□ *с дискретным временем*

переходы из одного состояния в другое могут происходить только в определенные, разделенные конечными интервалами моменты времени t_1, t_2, \dots ;

□ *с непрерывным временем*

переход системы из одного состояния в другое возможен в любой момент времени t .

Случайные процессы
в СМО (как правило)



Основные факторы, обуславливающие процессы, протекающие в СМО:

- потоки требований (заявок),
- потоки обслуживания.

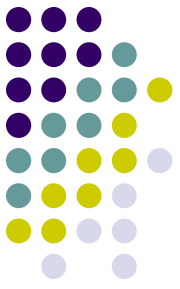


Потоки событий

Поток событий – последовательность событий, происходящих одно за другим в некоторые моменты времени.

Потоки событий:

- *однородные,*
- *неоднородные.*



Поток событий называется *однородным*, если он характеризуется только моментами поступления этих событий.

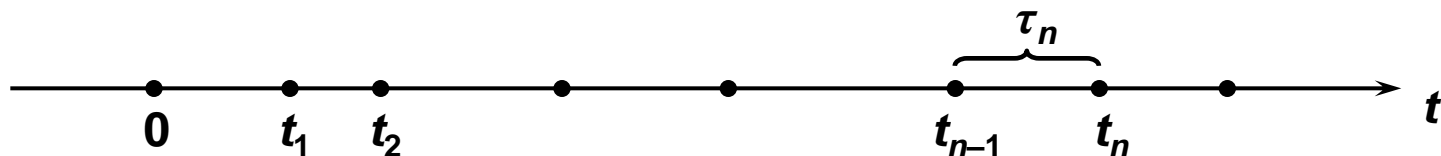
Однородный поток может быть задан

- последовательностью

$$\{ t_n \} = \{ 0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq \dots \},$$

где t_n – момент наступления n -го события (неотрицательное вещественное число);

- последовательностью $\{ \tau_n \}$ промежутков времени между $(n-1)$ -м и n -м событиями, где $\tau_1 = t_1$, $\tau_n = t_n - t_{n-1}$, $n \geq 2$.





Потоком *неоднородных событий* называется последовательность $\{ (t_n, f_n) \}$,

где t_n – моменты наступления событий,

f_n – набор признаков события.

Принадлежность к тому или иному источнику заявок, наличие приоритета, возможность обслуживания тем или иным типом канала и т. п.



Если интервал времени между соседними событиями в потоке является случайной величиной T , то поток называется *случайным*;
в противном случае — *детерминированным*.



Простейший поток и его свойства

Поток однородных событий называется *простейшим* (*стационарным пуассоновским потоком*), если он обладает следующими тремя свойствами.

1. *Стационарность.*

Каковы бы ни были $\tau > 0$ и целое $k \geq 0$, вероятность того, что в течение промежутка времени $(t, t + \tau)$ произойдет ровно k событий, одна и та же для всех $t \geq 0$ (зависит только от k и τ);

эту вероятность будем обозначать $v_k(\tau)$;

при любом τ
$$\sum_{k=0}^{\infty} v_k(\tau) = 1.$$

Среднее число событий в единицу времени постоянно



2. *Отсутствие последствия.*

Условная вероятность наступления k событий в течение промежутка времени $(t, t + \tau)$, вычисленная при любом предположении о наступлении событий до момента времени t , равна безусловной вероятности $v_k(\tau)$ этого же события.

**События наступают
независимо друг от друга**

Это главное свойство простейшего потока.



3. Ординарность.

Пусть для данного стационарного потока $P_{k \geq 2}(\tau)$ – вероятность наступления по крайней мере двух событий в течение промежутка времени $(t, t + \tau)$ (при любом $t \geq 0$);

$$P_{k \geq 2}(\tau) = 1 - v_0(\tau) - v_1(\tau).$$

Тогда

$$P_{k \geq 2}(\tau) = o(\tau) \quad \text{при } \tau \rightarrow 0,$$

т. е.

$$\frac{P_{k \geq 2}(\tau)}{\tau} \rightarrow 0 \quad \text{при } \tau \rightarrow 0.$$

События происходят поодиночке,
а не парами, тройками, и т. д.



Обозначим

$P_{k \geq 1}(\tau)$ – вероятность наступления по крайней мере одного события в течение промежутка времени $(t, t + \tau)$.

$$P_{k \geq 1}(\tau) = 1 - v_0(\tau) = v_1(\tau) + P_{k \geq 2}(\tau),$$

$$\lim_{\tau \rightarrow 0} \frac{P_{k \geq 1}(\tau)}{\tau} = \lim_{\tau \rightarrow 0} \left(\frac{v_1(\tau)}{\tau} + o(\tau) \right) = \lambda.$$

Существование предела
выводится из свойств
простейшего потока

λ – *интенсивность потока*.

Интерпретируется как среднее число событий в единицу времени.

Полностью характеризует простейший поток.



Для простейшего потока интенсивности λ можно показать:

- число X событий, попадающих на промежуток времени длины τ , имеет распределение Пуассона с параметром $a = \lambda\tau$:

$$P(X = k) = \frac{(\lambda\tau)^k}{k!} \cdot e^{-\lambda\tau}, \quad k = 0, 1, \dots,$$

$$M(X) = \lambda\tau;$$

- промежуток времени T между соседними событиями в потоке имеет показательное распределение с параметром λ :

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0.$$



Простейший поток играет среди потоков событий особую роль (аналогичную роли нормального закона среди законов распределения):

при суммировании (взаимном наложении) n ординарных стационарных независимых потоков с почти любым последствием при неограниченном увеличении n получается поток, сколь угодно близкий к простейшему.

Условие: складываемые потоки должны оказывать на сумму равномерно малое влияние.



Режимы функционирования СМО

При рассмотрении общих СМО предполагается:
система функционирует в течение достаточно
большого интервала времени;
при этом

- на начальном этапе функционирования в СМО (как и в любой динамической системе) возникает «переходный», нестационарный процесс;
- по истечении некоторого времени переходный процесс может затухать, а СМО – перейти на стационарный (установившийся) режим, вероятностные характеристики которого не зависят от времени.



Во многих прикладных задачах – основной вопрос:
существует ли в СМО стационарный режим;
если да, то каковы основные функциональные
характеристики СМО в стационарном режиме.



СМО с отказами и с ожиданием

- В *СМО с отказами* заявка, поступившая в систему в момент, когда все каналы обслуживания заняты, немедленно получает отказ, покидает систему и в дальнейшем процессе обслуживания не участвует.
- В *СМО с ожиданием* заявка, поступившая в систему в момент, когда все каналы обслуживания заняты, не покидает систему, а находится в очереди в состоянии ожидания обслуживания, пока не освободится какой-либо канал.



Если ожидание заявки в очереди ничем не ограничено, то такая СМО называется «*чистой СМО с ожиданием*».

Если ожидание ограничено какими-то условиями, то СМО называется «*системой смешанного типа*».

Для прикладных задач СМО смешанного типа представляют наибольший интерес.



Типы ограничений на ожидание.

❑ Ограничение на время ожидания в очереди.

Время пребывания заявки в очереди ограничено сверху значением $T_{ож}$, которое может быть как строго определенным, так и случайным.

Ограничивается только время пребывания в очереди; начатое обслуживание доводится до конца независимо от времени, проведенного в очереди.



- ❑ Ограничение на общее время пребывания заявки в системе.

Пример: воздушная цель может пробыть в контролируемой зоне лишь ограниченное время и покидает эту зону вне зависимости от того, закончилось ли «обслуживание».

- ❑ Ограничение на число заявок в очереди.
Наличие в СМО накопителя ограниченной емкости.



Функциональные характеристики стационарных СМО

Набор показателей качества обслуживания – исходя из практического использования конкретной системы.

Чаще всего используются следующие показатели:

- среднее число заявок, находящихся в СМО;
- среднее число заявок в очереди;
- среднее время пребывания заявки в СМО;
- среднее время пребывания заявки в очереди;
- среднее число занятых приборов (каналов) обслуживания;
- коэффициенты загрузки каналов обслуживания;
- вероятность потери заявки в СМО с отказами.



Сокращенные обозначения типов СМО

- Обозначения Кендалла (D.G. Kendall), 1953 г.

Тип СМО описывается кодом вида

$$\mathbf{A} / \mathbf{B} / m$$

где

A – обозначение закона распределения промежутков времени между поступлениями заявок в систему;

B – обозначение закона распределения времени обслуживания заявок;

m – число каналов обслуживания.



Стандартные обозначения законов распределения
(символы **A** и **B**):

M – показательное распределение (простейший поток событий);

D – детерминированный интервал времени между событиями в потоке;

E_k – распределение Эрланга **k**-го порядка;

G, **GI** – произвольный закон распределения
(**I** означает, что последовательные промежутки независимы).



- Добавления к обозначениям Кендалла.
 - Тип СМО описывается кодом вида

$A / B / m / K / n$

где

K – емкость накопителя;

n – мощность (конечная или бесконечная)
источника заявок.

Отсутствие одного (двух) последних символов означает: их значение может быть сколь угодно велико.

Пример:

$D / M / 2 / 20$ – двухканальная СМО с детерминированным временем между поступлениями заявок, показательным временем обслуживания и накопителем емкостью 20.



- Тип СМО описывается кодом вида

$$(A / B / m) : (d / K / n)$$

где

d – дисциплина очереди;

K – максимальная емкость системы (максимальное количество заявок, которое может находиться в СМО).

Стандартные обозначения дисциплины очереди
(символ ***d***):

FCFS – «первым пришел – первым обслуживаешься»;

LCFS – «последним пришел – первым обслуживаешься»;

SIRO – случайный выбор заявок;

GD – произвольный тип дисциплины.



Пример:

$(M / D / 10) : (GD / K / \infty) -$

десятиканальная СМО с простейшим входным потоком заявок (показательный закон распределения интервалов между поступлениями заявок), фиксированным временем обслуживания; при этом дисциплина очереди не регламентирована, максимальное число заявок, находящихся в системе, равно K , источник заявок имеет неограниченную емкость.



Формула Литтла

Рассмотрим СМО общего вида: **$G/G/m$** .

Обозначим:

$\alpha(t)$ – число требований, поступивших в СМО на промежутке $(0, t)$;

$\gamma(t)$ – общее время, проведенное всеми требованиями в системе за время $(0, t)$;

$\lambda_t = \frac{\alpha(t)}{t}$ – интенсивность поступления требований в СМО на промежутке $(0, t)$ (среднее число требований в единицу времени);

$T_t = \frac{\gamma(t)}{\alpha(t)}$ – время, проведенное одним требованием в системе, усредненное по всем требованиям, поступившим в СМО за время $(0, t)$;



$\bar{N}_t = \frac{\gamma(t)}{t}$ – среднее число требований в системе на промежутке $(0, t)$.

Ясно, что $\bar{N}_t = \lambda_t \cdot T_t$.

Пусть рассматриваемая СМО такова, что существуют пределы

$$\lambda = \lim_{t \rightarrow \infty} \lambda_t, \quad T = \lim_{t \rightarrow \infty} T_t.$$



Тогда

существует также предел

$$\bar{N} = \lim_{t \rightarrow \infty} \bar{N}_t ,$$

причем

$$\bar{N} = \lambda \cdot T . \quad (3.1)$$

Этот результат называется *формулой Литтла*.

Устанавливает:

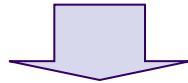
среднее число требований, находящихся в системе,
равно произведению интенсивности поступления
требований в систему на среднее время пребывания
требования в системе.



При обосновании этого результата не предполагалось никаких ограничений

- ни на вид распределения входящего потока,
- ни на вид распределения времени обслуживания,
- ни на число обслуживающих приборов,
- ни на конкретный характер дисциплины обслуживания.

Предполагается только, что в число $\alpha(t)$ входят те требования, которые действительно попали в систему (не получили отказ)



$\alpha(t)$ совпадает с плотностью входного потока заявок только в СМО с ожиданием.



Формула (3.1) может представлять как СМО в целом, так и ее отдельные компоненты.

- Пусть рассматривается только очередь (без обслуживающего прибора).

Тогда формула (3.1) примет вид

$$\bar{N}_q = \lambda \cdot W_q, \quad (3.2)$$

где

\bar{N}_q – среднее число требований в очереди;

W_q – среднее время ожидания в очереди.



- Пусть рассматривается только обслуживающий прибор (приборы).

Тогда формула (3.1) примет вид

$$\bar{N}_s = \lambda \cdot W_s, \quad (3.3)$$

где

\bar{N}_s – среднее число требований в обслуживающем приборе (приборах);

W_s – среднее время, проведенное требованием в обслуживающем приборе (время обслуживания).

Имеют место равенства:

$$T = W_q + W_s, \quad \bar{N} = \bar{N}_q + \bar{N}_s.$$



Марковские модели массового обслуживания

Марковская модель массового обслуживания — это описание операции массового обслуживания с помощью марковского процесса с дискретным множеством состояний (с помощью цепи Маркова).



Обозначим:

$p_i(t) = P(\xi_t = z_i)$ – вероятность того, что в момент времени t процесс ξ_t находится в состоянии z_i .

Определив значения $p_i(t)$ (для любого t), можно получить различные характеристики качества обслуживания.



СМО $M / M / n / 0$

Это n -канальная СМО марковского типа с отказами.

Обеспечивается свойствами потока заявок и потока обслуживаний (оба потока – простейшие)

Пусть

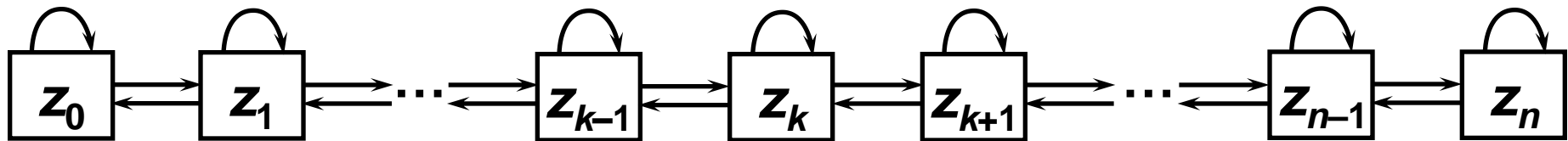
- входной поток требований (заявок) – простейший с интенсивностью λ ;
- поток обслуживаний – простейший с интенсивностью μ
(время обслуживания распределено по показательному закону с параметром μ).



СМО имеет конечное множество состояний:

- z_0 — ни один канал не занят,
- z_1 — занят ровно один канал,
-
- z_k — занято ровно k каналов,
-
- z_n — заняты все n каналов.

Схема возможных переходов:



В силу свойств простейшего потока вероятностью «перескока» через состояние можно пренебречь.



Используя предположения о характере входного потока заявок и потока обслуживаний, а также теорему сложения вероятностей, можно показать, что вероятности $p_i(t)$ удовлетворяют соотношениям:

$$p_0(t + \tau) = p_0(t)(1 - \lambda\tau) + \mu p_1(t)\tau + o(\tau);$$

для любого k , $0 < k < n$

$$p_k(t + \tau) = p_k(t)(1 - (\lambda + k\mu)\tau) + p_{k-1}(t)\lambda\tau + p_{k+1}(t)(k+1)\mu\tau + o(\tau);$$

$$p_n(t + \tau) = p_n(t)(1 - n\mu\tau) + p_{n-1}(t)\lambda\tau + o(\tau).$$

Перенеся $p_k(t)$, $0 \leq k \leq n$, в левые части, разделив обе части равенств на τ и переходя к пределу при $\tau \rightarrow 0$, получим систему дифференциальных уравнений:



$$\left\{ \begin{array}{l} \frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t), \\ \frac{dp_1(t)}{dt} = \lambda p_0(t) - (\lambda + \mu)p_1(t) + 2\mu p_2(t), \\ \dots\dots\dots \\ \frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)p_k(t) + (k+1)\mu p_{k+1}(t) \\ \quad (1 \leq k \leq n-1), \\ \dots\dots\dots \\ \frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - n\mu \cdot p_n(t). \end{array} \right. \quad (3.4)$$

Начальные условия: $p_0(0) = 1$, $p_1(0) = \dots = p_n(0) = 0$.

Уравнения (3.4) называются *уравнениями Эрланга*.



Система (3.4) может быть относительно легко проинтегрирована при любом конкретном n .

Вероятности $p_k(t)$ характеризуют среднюю загрузку СМО и ее изменение с течением времени.

В частности, $p_n(t)$ есть вероятность потери заявки (заявка, заставшая все каналы занятыми, получает отказ):

$$P_{\text{отказа}} = p_n(t).$$

Величина $1 - p_n(t)$ характеризует *относительную пропускную способность системы* – отношение среднего числа обслуженных в единицу времени заявок к общему числу заявок на обслуживание.



Существование стационарного режима.

В стационарном (установившемся) режиме вероятности $p_i(t) = p_i$ (не зависят от t).

При $p_i > 0$ – стационарное
(равновесное, предельное)
распределение

Тогда

$$\frac{d}{dt} p_i(t) = 0, \quad i = 0, 1, \dots, n. \quad (3.5)$$

При подстановке условия (3.5) в систему (3.4) получается *система уравнений равновесия* (СУР).



Предположим, что предельные вероятности p_0, p_1, \dots, p_n существуют.

Эти вероятности должны удовлетворять СУР и условию нормировки $\sum_{k=0}^n p_k = 1$.

СУР имеет вид:

$$\left\{ \begin{array}{l} -\lambda p_0 + \mu p_1 = 0, \\ \lambda p_0 - (\lambda + \mu)p_1 + 2\mu p_2 = 0, \\ \dots\dots\dots \\ \lambda p_{k-1} - (\lambda + k\mu)p_k + (k+1)\mu p_{k+1} = 0 \\ \quad (1 \leq k \leq n-1), \\ \dots\dots\dots \\ \lambda p_{n-1} - n\mu p_n = 0. \end{array} \right. \quad (3.6)$$



Из первого уравнения

$$p_1 = \frac{\lambda}{\mu} p_0 ,$$

из второго

$$\begin{aligned} p_2 &= \frac{1}{2\mu} (-\lambda p_0 + (\lambda + \mu) p_1) = \frac{1}{2\mu} \left(-\lambda p_0 + \frac{\lambda(\lambda + \mu)}{\mu} p_0 \right) = \\ &= \frac{\lambda^2}{2\mu^2} p_0 , \end{aligned}$$

.....

Для любого $k \leq n$

$$p_k = \frac{\lambda^k}{k! \cdot \mu^k} p_0 .$$



Обозначим

$$\alpha = \frac{\lambda}{\mu}.$$

Величина α называется *приведенной плотностью потока заявок*.

Интерпретация: это среднее число заявок, приходящееся на среднее время обслуживания одной заявки.

Тогда

$$p_k = \frac{\alpha^k}{k!} p_0, \quad k = 1, 2, \dots, n.$$



Из условия $\sum_{k=0}^n p_k = 1$:

$$\sum_{k=0}^n p_k = \sum_{k=0}^n \frac{\alpha^k}{k!} p_0 = p_0 \sum_{k=0}^n \frac{\alpha^k}{k!} = 1,$$

$$p_0 = \frac{1}{\sum_{k=0}^n \frac{\alpha^k}{k!}}.$$

Окончательно:

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^n \frac{\alpha^k}{k!}}, \quad k = 0, 1, \dots, n. \quad (3.7)$$



Формулы (3.7) называются *формулами Эрланга*.

Таким образом:

в системе **$M / M / n / 0$** при любом **α**
(т. е при любых значениях параметров **λ** и **μ**)
существуют предельные вероятности **p_k**
(а значит, и стационарный режим), которые могут
быть найдены по формулам (3.7).



Характеристики функционирования СМО в стационарном режиме:

- вероятность потери заявки

$$P_{отказа} = p_n = \frac{\frac{\alpha^n}{n!}}{\sum_{k=0}^n \frac{\alpha^k}{k!}} ;$$

Формула потерь
Эрланга

- относительная пропускная способность системы

$$q = 1 - P_{отказа} = 1 - \frac{\frac{\alpha^n}{n!}}{\sum_{k=0}^n \frac{\alpha^k}{k!}} ;$$



- среднее число заявок в системе

$$\bar{N} = \sum_{k=0}^n k \cdot p_k = \frac{\sum_{k=1}^n \frac{\alpha^k}{(k-1)!}}{\sum_{k=0}^n \frac{\alpha^k}{k!}} = \frac{\alpha \sum_{k=1}^n \frac{\alpha^{k-1}}{(k-1)!}}{\sum_{k=0}^n \frac{\alpha^k}{k!}} ;$$

- среднее время обслуживания и одновременно среднее время пребывания заявки в системе

$$T = \frac{1}{\mu} .$$



СМО $M / M / n$ с ограничением на время ожидания

Это обобщение разобранный задачи Эрланга для СМО с отказами.



Рассмотрим смешанную ***n***-канальную СМО при следующих условиях:

- входной поток требований (заявок) – простейший с интенсивностью λ ;
- поток обслуживаний – простейший с интенсивностью μ ;
- заявка, заставшая все каналы занятыми, становится в очередь и ожидает обслуживания;
- время ожидания ограничено величиной $T_{ож}$, имеющей показательное распределение с параметром ν .



По аналогии с параметрами λ и μ , параметр ν можно интерпретировать как интенсивность «потока уходов» из очереди заявок, у которых превышено время ожидания.

При $\nu \rightarrow \infty$ СМО смешанного типа превращается в систему с отказами.

При показательном распределении величины $T_{ож}$ функциональные характеристики СМО не зависят от дисциплины очереди:

для каждой заявки закон распределения оставшегося времени ожидания не зависит от того, сколько времени заявка уже стояла в очереди.

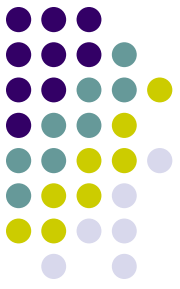


Возможные состояния системы:

- $\mathbf{z_0}$ – ни один канал не занят (очереди нет),
- $\mathbf{z_1}$ – занят ровно один канал (очереди нет),
-
- $\mathbf{z_n}$ – заняты все \mathbf{n} каналов (очереди нет),
- $\mathbf{z_{n+1}}$ – заняты все \mathbf{n} каналов, одна заявка стоит в очереди,
-
- $\mathbf{z_{n+s}}$ – заняты все \mathbf{n} каналов, \mathbf{s} заявок стоят в очереди,
-

**Бесконечное (счетное)
множество состояний**

Нумерация состояний – по числу заявок, находящихся в системе.



Система дифференциальных уравнений, связывающая вероятности $p_i(t)$, в данном случае будет иметь бесконечное число уравнений.

Первые n уравнений – это соответствующие уравнения Эрланга:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t),$$

$$\frac{dp_1(t)}{dt} = \lambda p_0(t) - (\lambda + \mu)p_1(t) + 2\mu p_2(t),$$

.....

$$\frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)p_k(t) + (k+1)\mu p_{k+1}(t) \\ (1 \leq k \leq n-2),$$

.....

$$\frac{dp_{n-1}(t)}{dt} = \lambda p_{n-2}(t) - (\lambda + (n-1)\mu)p_{n-1}(t) + n\mu \cdot p_n(t).$$



Остальные уравнения системы имеют вид:

$$\frac{d}{dt} p_n(t) = \lambda \cdot p_{n-1}(t) - (\lambda + n\mu) \cdot p_n(t) + (n\mu + \nu) \cdot p_{n+1}(t),$$

$$\begin{aligned} \frac{d}{dt} p_{n+s}(t) = \\ = \lambda \cdot p_{n+s-1}(t) - (\lambda + n\mu + s\nu) \cdot p_{n+s}(t) + (n\mu + (s+1)\nu) \cdot p_{n+s+1}(t), \\ s = 1, 2, \dots \end{aligned}$$

В итоге – система уравнений:

$$\left\{ \begin{array}{l} \frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t), \\ \frac{dp_1(t)}{dt} = \lambda p_0(t) - (\lambda + \mu)p_1(t) + 2\mu p_2(t), \\ \dots\dots\dots \\ \frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)p_k(t) + (k+1)\mu p_{k+1}(t) \\ \quad (1 \leq k \leq n-1), \\ \dots\dots\dots \\ \frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - (\lambda + n\mu)p_n(t) + (n\mu + \nu)p_{n+1}(t), \\ \dots\dots\dots \\ \frac{dp_{n+s}(t)}{dt} = \lambda p_{n+s-1}(t) - (\lambda + n\mu + s\nu)p_{n+s}(t) + (n\mu + (s+1)\nu)p_{n+s+1}(t), \\ \quad (s \geq 1), \\ \dots\dots\dots \end{array} \right. \quad (3.8)$$





Уравнения (3.8) являются обобщением уравнений Эрланга на случай СМО смешанного типа с ограниченным временем ожидания.

Предположим, что существуют предельные вероятности $p_0, p_1, \dots, p_n, \dots$

Эти вероятности должны удовлетворять СУР и условию нормировки

$$\sum_{k=0}^{\infty} p_k = 1.$$



СУР получается подстановкой условия (3.5) в уравнения (3.8):

$$\left\{ \begin{array}{l} -\lambda p_0 + \mu p_1 = 0, \\ \lambda p_0 - (\lambda + \mu)p_1 + 2\mu p_2 = 0, \\ \dots\dots\dots \\ \lambda p_{k-1} - (\lambda + k\mu)p_k + (k+1)\mu p_{k+1} = 0 \\ \quad (1 \leq k \leq n-1), \\ \dots\dots\dots \\ \lambda p_{n-1} - (\lambda + n\mu)p_n + (n\mu + \nu)p_{n+1} = 0, \\ \dots\dots\dots \\ \lambda p_{n+s-1} - (\lambda + n\mu + s\nu)p_{n+s} + (n\mu + (s+1)\nu)p_{n+s+1} = 0, \\ \quad (s \geq 1), \\ \dots\dots\dots \end{array} \right. \quad (3.9)$$



Так же как для системы (3.6), из первых $n+1$ уравнений получим:

$$\text{для любого } k \leq n \quad p_k = \frac{\lambda^k}{k! \cdot \mu^k} p_0 .$$

При $k = n+s$, $s = 1, 2, \dots$

$$p_{n+1} = \frac{(\lambda + n\mu)p_n - \lambda p_{n-1}}{n\mu + \nu} = \frac{\lambda^{n+1}}{n! \cdot \mu^n (n\mu + \nu)} p_0 ,$$

$$p_{n+2} = \frac{(\lambda + n\mu + \nu)p_{n+1} - \lambda p_n}{n\mu + 2\nu} = \frac{\lambda^{n+2}}{n! \cdot \mu^n (n\mu + \nu)(n\mu + 2\nu)} p_0 ,$$

.....



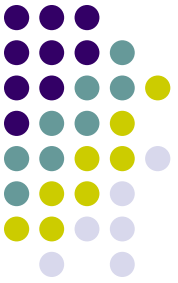
Итог:

для любого $s \geq 1$

$$p_{n+s} = \frac{\lambda^{n+s}}{n! \cdot \mu^n \prod_{l=1}^s (n\mu + l\nu)} p_0.$$

Вероятность p_0 , как и ранее, определим из условия

$$\sum_{k=0}^{\infty} p_k = 1.$$



$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^n \frac{\lambda^k}{k! \mu^k} p_0 + \sum_{s=1}^{\infty} \frac{\lambda^{n+s}}{n! \mu^n \prod_{l=1}^s (n\mu + l\nu)} p_0 = 1,$$

$$p_0 \left(\sum_{k=0}^n \frac{\lambda^k}{k! \mu^k} + \sum_{s=1}^{\infty} \frac{\lambda^{n+s}}{n! \mu^n \prod_{l=1}^s (n\mu + l\nu)} \right) = 1,$$

$$p_0 = \frac{1}{\sum_{k=0}^n \frac{\lambda^k}{k! \mu^k} + \sum_{s=1}^{\infty} \frac{\lambda^{n+s}}{n! \mu^n \prod_{l=1}^s (n\mu + l\nu)}}.$$



Для упрощения полученных выражений обозначим

$$\alpha = \frac{\lambda}{\mu}, \quad \beta = \frac{\nu}{\mu} \quad -$$

приведенные плотности потоков прихода заявок и ухода заявок, стоящих в очереди.

Интерпретация: **α** – среднее число заявок, приходящееся на среднее время обслуживания одной заявки;
 β – среднее число уходов заявок из очереди, приходящееся на среднее время обслуживания одной заявки.



Тогда

$$p_k = \frac{\alpha^k}{k!} \cdot p_0, \quad 1 \leq k \leq n,$$

$$p_{n+s} = \frac{\frac{\alpha^{n+s}}{n!}}{\prod_{l=1}^s (n+l\beta)} \cdot p_0, \quad s \geq 1,$$

(3.10)

где

$$p_0 = \frac{1}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n+l\beta)}}.$$

(3.11)

После подстановки (3.11) в (3.10) окончательно получим:



$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)}}, \quad 0 \leq k \leq n, \quad (3.12)$$

$$p_{n+s} = \frac{\frac{\alpha^n}{n!} \frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)}}, \quad s \geq 1. \quad (3.13)$$

В сходимости ряда

$$\sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)}$$



можно убедиться, например, используя признак д'Аламбера:

$$\lim_{s \rightarrow \infty} \frac{u_{s+1}}{u_s} = \lim_{s \rightarrow \infty} \frac{\alpha}{n + (s+1)\beta} = 0.$$

Из сходимости этого ряда следует также, что величина

$$\frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)},$$

а значит, и вероятности \mathbf{p}_{n+s} , при неограниченном увеличении \mathbf{s} становятся сколь угодно малыми.



Таким образом:

в системе $M / M / n$ с ограничением на время ожидания при любых α и β (т. е. при любых значениях параметров λ , μ и ν) существуют предельные вероятности p_i (а значит, и стационарный режим), которые могут быть найдены по формулам (3.12) и (3.13).



Основные характеристики функционирования СМО в стационарном режиме.

- Среднее число заявок, находящихся в очереди.

$$\bar{N}_q = M(s) = \sum_{s=1}^{\infty} s \cdot p_{n+s} = \frac{\frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{s \alpha^s}{\prod_{l=1}^s (n + l\beta)}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)}}.$$



- Вероятность потери заявки

можно оценить как отношение среднего числа заявок, уходящих из очереди в единицу времени не обслуженными, к среднему числу заявок, поступающих в систему в единицу времени.

Поэтому

$$\begin{aligned} P_{отказа} &= \bar{N}_q \cdot \frac{\nu}{\lambda} = \bar{N}_q \cdot \frac{\mu}{\lambda} = \bar{N}_q \cdot \frac{\beta}{\alpha} = \\ &= \frac{\beta}{\alpha} \cdot \frac{\frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{s \alpha^s}{\prod_{l=1}^s (n + l\beta)}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)}}. \end{aligned}$$



- Относительная пропускная способность системы

$$q = 1 - P_{отказа}.$$

- Среднее число заявок в системе

$$\begin{aligned}\bar{N} &= \sum_{j=0}^{\infty} j \cdot p_j = \sum_{k=0}^n k \cdot p_k + \sum_{s=1}^{\infty} s \cdot p_{n+s} = \\ &= \frac{\sum_{k=1}^n \frac{\alpha^k}{(k-1)!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{s \alpha^s}{\prod_{l=1}^s (n+l\beta)}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n+l\beta)}}.\end{aligned}$$



- Среднее число занятых каналов

$$\bar{N}_s = \bar{N} - \bar{N}_q .$$

- Вероятность того, что поступившая в СМО заявка сразу будет обслужена (без ожидания)

$$P_{обсл} = \sum_{k=0}^{n-1} p_k = \frac{\sum_{k=0}^{n-1} \frac{\alpha^k}{k!}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{l=1}^s (n + l\beta)}} .$$

- Вероятность того, что поступившая в СМО заявка будет некоторое время ожидать обслуживания

$$P_{ож} = 1 - P_{обсл} .$$



Как уже отмечалось, при $\nu \rightarrow \infty$ (т. е. $\beta \rightarrow \infty$) рассматриваемая СМО смешанного типа превращается в систему с отказами:

вероятности p_{n+s} в формулах (3.13) обращаются в нуль, а формулы (3.12) превращаются в формулы Эрланга (3.7).

Далее рассматривается другой крайний случай:

$$\nu \rightarrow 0 \quad (\text{т. е. } \beta \rightarrow 0) \quad -$$

чистая система с ожиданием.



СМО $M / M / n / \infty$

Это n -канальная СМО марковского типа с ожиданием.

В такой системе

$$P_{отказа} = 0.$$

Пусть

- входной поток требований (заявок) – простейший с интенсивностью λ ;
- поток обслуживаний – простейший с интенсивностью μ .



При $\beta \rightarrow 0$ формула (3.11) преобразуется в

$$p_0 = \frac{1}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{n^s}}.$$

Ряд

$$\sum_{s=1}^{\infty} \frac{\alpha^s}{n^s} = \sum_{s=1}^{\infty} \left(\frac{\alpha}{n} \right)^s$$

будет сходящимся при $\alpha < n$,
расходящимся при $\alpha \geq n$.

Условие существования
стационарного режима



При условии $\alpha < n$

$$\sum_{s=1}^{\infty} \frac{\alpha^s}{n^s} = \sum_{s=1}^{\infty} \left(\frac{\alpha}{n} \right)^s = \frac{\frac{\alpha}{n}}{1 - \frac{\alpha}{n}} = \frac{\alpha}{n - \alpha},$$

поэтому

$$p_0 = \frac{1}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n - \alpha)}}. \quad (3.14)$$



С учетом этого формулы (3.12) и (3.13) преобразуются к виду

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n-\alpha)}}, \quad 0 \leq k \leq n, \quad (3.15)$$

$$p_{n+s} = \frac{\frac{\alpha^{n+s}}{n!n^s}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n-\alpha)}}, \quad s \geq 1. \quad (3.16)$$



Таким образом:

в системе $M / M / n / \infty$ стационарный режим существует только при условии $\alpha < n$

(среднее число входящих заявок, приходящееся на среднее время обслуживания одной заявки не выходит за пределы возможностей n -канальной системы);

при этом условии предельные вероятности p_i могут быть найдены по формулам (3.15), (3.16).



Характеристики функционирования СМО в стационарном режиме.

- Среднее число заявок, находящихся в очереди

$$\bar{N}_q = \frac{\frac{\alpha^{n+1}}{n \cdot n! \left(1 - \frac{\alpha}{n}\right)^2}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n - \alpha)}}$$

(получается из аналогичной формулы для системы с ограничением на время ожидания при $\beta \rightarrow 0$).



Для системы с ожиданием можно использовать формулы (3.1) – (3.3).

- Среднее время ожидания в очереди.

По формуле (3.2)

$$W_q = \frac{\bar{N}_q}{\lambda}.$$

- Среднее число занятых каналов.

По формуле (3.3)

$$\bar{N}_s = \lambda \cdot W_s = \lambda \cdot \frac{1}{\mu} = \alpha.$$

- Среднее число заявок в системе.

$$\bar{N} = \bar{N}_q + \bar{N}_s = \bar{N}_q + \alpha.$$



- Среднее время пребывания заявки в системе

$$T = W_q + W_s = W_q + \frac{1}{\mu}.$$

- Вероятность того, что поступившая в СМО заявка сразу будет обслужена (без ожидания)

$$P_{обсл} = \sum_{k=0}^{n-1} p_k = \frac{\sum_{k=0}^{n-1} \frac{\alpha^k}{k!}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^{n+1}}{n!(n-\alpha)}}.$$

- Вероятность того, что поступившая в СМО заявка будет некоторое время ожидать обслуживания

$$P_{ож} = 1 - P_{обсл}.$$



СМО $M / M / n / K$

Это n -канальная СМО марковского типа с ограничением на длину очереди.

Пусть

- входной поток требований (заявок) – простейший с интенсивностью λ ;
- поток обслуживаний – простейший с интенсивностью μ ;
- заявка, заставшая все каналы занятыми, становится в очередь, только если в очереди находится менее чем K заявок; в противном случае поступившая заявка покидает систему не обслуженной.



Число возможных состояний системы конечно:

\mathbf{z}_0 — ни один канал не занят (очереди нет),

\mathbf{z}_1 — занят ровно один канал (очереди нет),

.....

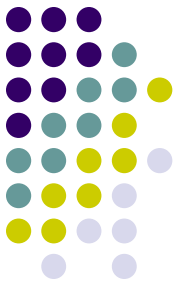
\mathbf{z}_n — заняты все \mathbf{n} каналов (очереди нет),

\mathbf{z}_{n+1} — заняты все \mathbf{n} каналов, одна заявка стоит в очереди,

.....

\mathbf{z}_{n+K} — заняты все \mathbf{n} каналов, \mathbf{K} заявок стоят в очереди.

Система дифференциальных уравнений, связывающая вероятности $\mathbf{p}_i(\mathbf{t})$, включает $\mathbf{n}+\mathbf{K}+1$ уравнение и имеет вид:



$$\left\{ \begin{array}{l}
 \frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t), \\
 \dots\dots\dots \\
 \frac{dp_k(t)}{dt} = \lambda p_{k-1}(t) - (\lambda + k\mu)p_k(t) + (k+1)\mu p_{k+1}(t) \\
 \quad (1 \leq k \leq n-1), \\
 \dots\dots\dots \\
 \frac{dp_n(t)}{dt} = \lambda p_{n-1}(t) - (\lambda + n\mu)p_n(t) + n\mu p_{n+1}(t), \\
 \dots\dots\dots \\
 \frac{dp_{n+s}(t)}{dt} = \lambda p_{n+s-1}(t) - (\lambda + n\mu)p_{n+s}(t) + n\mu p_{n+s+1}(t), \\
 \quad (1 \leq s < K), \\
 \dots\dots\dots \\
 \frac{dp_{n+K}(t)}{dt} = \lambda p_{n+K-1}(t) - n\mu p_{n+K}(t).
 \end{array} \right. \quad (3.17)$$



В данном случае СУР имеет вид:

$$\left\{ \begin{array}{l} -\lambda p_0 + \mu p_1 = 0, \\ \dots\dots\dots \\ \lambda p_{k-1} - (\lambda + k\mu)p_k + (k+1)\mu p_{k+1} = 0 \\ (1 \leq k \leq n-1), \\ \dots\dots\dots \\ \lambda p_{n-1} - (\lambda + n\mu)p_n + n\mu p_{n+1} = 0, \\ \dots\dots\dots \\ \lambda p_{n+s-1} - (\lambda + n\mu)p_{n+s} + n\mu p_{n+s+1} = 0, \\ (1 \leq s < K), \\ \dots\dots\dots \\ \lambda p_{n+K-1} - n\mu p_{n+K} = 0. \end{array} \right. \quad (3.18)$$



Предельные вероятности $p_0, p_1, \dots, p_n, \dots$ должны удовлетворять системе (3.18) и условию

$$\sum_{k=0}^{n+K} p_k = 1.$$

Нахождение этих вероятностей выполняется аналогично рассмотренным ранее случаям.



Итог:

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^K \left(\frac{\alpha}{n}\right)^s}, \quad k = 0, 1, \dots, n,$$

$$p_{n+s} = \frac{\frac{\alpha^n}{n!} \left(\frac{\alpha}{n}\right)^s}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^K \left(\frac{\alpha}{n}\right)^s}, \quad s = 1, 2, \dots, K,$$

где $\alpha = \frac{\lambda}{\mu}$.

Стационарный режим существует при любых λ и μ .



Характеристики функционирования СМО в стационарном режиме.

- Вероятность потери заявки

$$P_{отказа} = P_{n+K} = \frac{\frac{\alpha^n}{n!} \left(\frac{\alpha}{n}\right)^K}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^K \left(\frac{\alpha}{n}\right)^s}.$$

- Среднее число потерянных (не обслуженных) заявок в единицу времени

$$\lambda_n = \lambda \cdot p_{n+K}.$$



- Относительная пропускная способность системы

$$q = 1 - P_{\text{отказа}};$$

абсолютная пропускная способность системы

$$Q = \lambda q.$$

- Средняя доля времени простоя системы (относительное время простоя)

$$p_0 = \frac{1}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^K \left(\frac{\alpha}{n} \right)^s}.$$



- Среднее число заявок, находящихся в очереди

$$\bar{N}_q = \sum_{s=1}^K s \cdot p_{n+s} = \frac{\frac{\alpha^n}{n!} \sum_{s=1}^K s \cdot \left(\frac{\alpha}{n}\right)^s}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^K \left(\frac{\alpha}{n}\right)^s}.$$

- Среднее число заявок в системе

$$\bar{N} = \sum_{k=0}^n k \cdot p_k + \sum_{s=1}^K s \cdot p_{n+s} = \frac{\sum_{k=1}^n \frac{\alpha^k}{(k-1)!}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^K \left(\frac{\alpha}{n}\right)^s} + \bar{N}_q.$$