

ЛЕКЦИЯ №6

Курс: Моделирование систем

**Тема: Характеристики вычислительных систем,
представленных в виде моделей СМО**

1. Модель размножения и гибели

**2. Модель вычислительной системы в виде одноканальной
СМО с очередью**

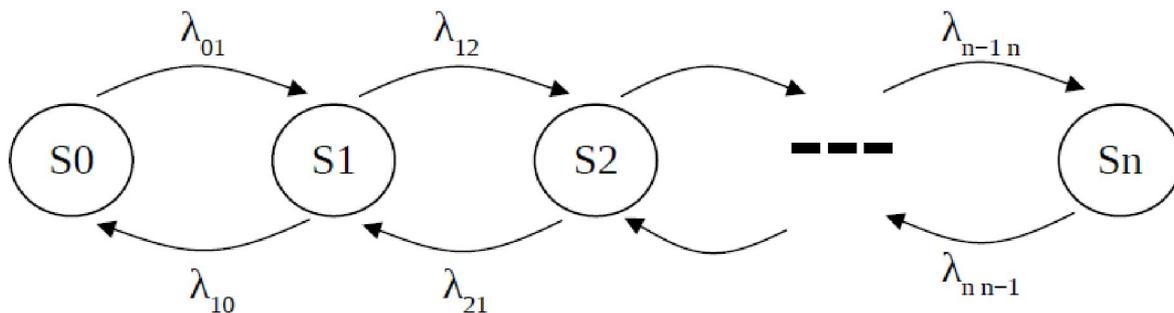
3. Примеры решения задач

1. Модель размножения и гибели

1.1 Граф модели размножения и гибели

Имея в распоряжении размеченный граф состояний, можно легко написать уравнения Колмогорова для вероятностей состояний, а также написать и решить уравнения для финальных вероятностей. Для некоторых случаев удастся получить решение этих уравнений в аналитическом виде.

Разновидностью марковской модели с дискретным числом состояний и непрерывным временем является модель размножения и гибели. Граф состояний этой модели имеет вид цепи. Интенсивности переходов из одного состояния в другое обозначены как λ_{ij} , а времена переходов распределены по показательному закону, т.е. все потоки, переводящие систему по стрелкам графа – простейшие.



1. Модель размножения и гибели

1.1 Граф модели размножения и гибели

Особенность этого графа в том, что все состояния системы (S_0, S_1, \dots, S_n) можно вытянуть в одну цепочку, в которой каждое из соседних состояний связано прямой и обратной стрелкой с каждым из соседних состояний – правым и левым, а крайние состояния S_0 и S_n только с одним соседним состоянием

Такая схема часто встречается в теории массового обслуживания.

Рассмотрим возможные состояния системы (S_0, S_1, \dots, S_n) и их вероятности:

$$p_0(t), p_1(t), \dots, p_n(t).$$

Очевидно, что для любого момента времени (условие нормировки):

$$\sum_{k=0}^n p_k(t) = 1$$

1. Модель размножения и гибели

1.2 Дифференциальные уравнения Колмогорова

Составим дифференциальные уравнения Колмогорова для всех вероятностей, используя приведенный граф состояний исходной модели.

Для этого зафиксируем момент времени t и найдем вероятность $p_0(t+\Delta t)$ того, что в момент времени $t+\Delta t$ система будет находиться в состоянии S_0 .

Исходя из представленной схемы графа состояний, это может произойти двумя способами:

- во – первых (событие A) - в момент t система находилась в состоянии S_0 , а за время Δt не перешла из него в состояние S_1 ;
- во – вторых (событие B) – в момент t система находилась в состоянии S_1 , а за время Δt перешла в состояние S_1 .

Складывая эти две возможности (по теореме сложения вероятностей) получим:

$$p_0(t + \Delta t) = p(A) + p(B)$$

1. Модель размножения и гибели

1.2 Дифференциальные уравнения Колмогорова

Вероятность события А найдем по теореме умножения вероятностей.

Вероятность того, что в момент t система находилась в состоянии S_0 , равна $p_0(t)$. Вероятность того, что за время Δt не придет ни одной заявки, равна $\exp(-\lambda_{01}\Delta t)$. Учитывая, что значение $\lambda_{01}\Delta t$ – мало, можно записать:

$$e^{-\lambda_{01}\Delta t} = 1 - \lambda_{01} \cdot \Delta t$$

Окончательно, для вероятности события А имеем:

$$p(A) = p_0(t) \cdot (1 - \lambda \cdot \Delta t)$$

1. Модель размножения и гибели

1.2 Дифференциальные уравнения Колмогорова

Найдем вероятность события $p(B)$. Вероятность того, что в момент t система была в состоянии S_1 , равна $p_1(t)$.

Вероятность того, что за время Δt система перейдет в состояние S_0 , равна – $1 - \exp(-\lambda_{10}\Delta t)$ или

$$1 - e^{-\lambda_{10}\Delta t} = \lambda_{10} \cdot \Delta t$$

Окончательно, для вероятности события B получим:

$$p(B) = p_1(t) \cdot \lambda_{10} \cdot \Delta t$$

1. Модель размножения и гибели

1.2 Дифференциальные уравнения Колмогорова

Суммируя вероятности событий А и В, окончательно получим,

$$p_0(t + \Delta t) = p_0(t) \cdot (1 - \lambda_{01} \cdot \Delta t) + p_1(t) \cdot \lambda_{10} \cdot \Delta t$$

Переносим в левую часть $p_0(t)$, делим на Δt и переходя к пределу, получим дифференциальное уравнение для $p_0(t)$:

$$\frac{dp_0(t)}{dt} = -\lambda_{01} \cdot p_0(t) + \lambda_{10} \cdot p_1(t)$$

1. Модель размножения и гибели

1.2 Дифференциальные уравнения Колмогорова

Аналогично, могут быть получены дифференциальные уравнения и для всех остальных вероятностей состояний:

$$\left. \begin{aligned}
 \frac{dp_0(t)}{dt} &= -\lambda_{01} \cdot p_0(t) + \lambda_{10} \cdot p_1(t) \\
 \frac{dp_1(t)}{dt} &= -(\lambda_{12} + \lambda_{10}) \cdot p_1(t) + \lambda_{01} \cdot p_0(t) + \lambda_{21} \cdot p_2(t) \\
 &\dots \\
 &\dots \\
 \frac{dp_n(t)}{dt} &= -\lambda_{n,n-1} \cdot p_n(t) + \lambda_{n-1,n} \cdot p_{n-1}(t)
 \end{aligned} \right\}$$

1. Модель размножения и гибели

1.3 Финальные стационарные уравнения Колмогорова

Вначале, после включения рассматриваемой системы в работу, протекающий в ней процесс не будет стационарным. Этот начальный процесс называется переходным – нестационарным. Однако, спустя некоторое время этот переходный процесс затухает и система перейдет в установившейся – стационарный режим работы.

Для стационарного режима вероятностные характеристики не зависят от времени. В этом стационарном режиме работы все вероятности $p_0(t)$, $p_1(t)$, ..., $p_n(t)$ стремятся к постоянным пределам p_0 , p_1 , ..., p_n , а все их производные стремятся к нулю.

1. Модель размножения и гибели

1.3 Финальные стационарные уравнения Колмогорова

Заменим в системе обыкновенных дифференциальных уравнений все вероятности их пределами, а все производные положим равными нулю. В этом случае получим систему уже не дифференциальных, а алгебраических уравнений:

$$\left. \begin{aligned} 0 &= -\lambda_{01} \cdot p_0 + \lambda_{10} \cdot p_1 \\ 0 &= -(\lambda_{12} + \lambda_{10}) \cdot p_1 + \lambda_{01} \cdot p_0 + \lambda_{21} \cdot p_2 \\ &\quad \dots \\ &\quad \dots \\ 0 &= -\lambda_{n,n-1} \cdot p_n + \lambda_{n-1,n} \cdot p_{n-1} \end{aligned} \right\}$$

К этим уравнениям необходимо добавить условие нормировки:

$$p_0 + p_1 + p_2 + \dots + p_n = 1$$

1. Модель размножения и гибели

1.3 Финальные стационарные уравнения Колмогорова

Разрешим полученную систему относительно неизвестных p_0, p_1, \dots, p_n .

Для первого уравнения имеем : $\lambda_{01} \cdot p_0 = \lambda_{10} \cdot p_1$

Для второго уравнения имеем : $(\lambda_{12} + \lambda_{10}) \cdot p_1 = \lambda_{01} \cdot p_0 + \lambda_{21} \cdot p_2$

Подставляя первое уравнение во второе, получим :

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

Аналогичное соотношение получим и для состояния S2

$$\lambda_{23} \cdot p_2 = \lambda_{32} \cdot p_3$$

Обобщая полученные соотношения можно записать и для произвольного состояния S_k

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

где $k=0,1,\dots,n$

1. Модель размножения и гибели

1.3 Финальные стационарные уравнения Колмогорова

Итоговая система преобразованных уравнений имеет следующий вид :

$$\left. \begin{aligned} \lambda_{01} \cdot p_0 &= \lambda_{10} \cdot p_1 \\ \lambda_{12} \cdot p_1 &= \lambda_{21} \cdot p_2 \\ &\dots \\ \lambda_{k-1,k} \cdot p_{k-1} &= \lambda_{k,k-1} \cdot p_k \\ \lambda_{n-1,n} \cdot p_{n-1} &= \lambda_{n,n-1} \cdot p_n \end{aligned} \right\}$$

Из первого уравнения можно получить

$$p_1 = \frac{\lambda_{01}}{\lambda_{10}} \cdot p_0$$

Из второго уравнения можно получить

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

Обобщая полученные соотношения для произвольного состояния S_k имеем:

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

1. Модель размножения и гибели

1.3 Финальные стационарные уравнения Колмогорова

В числителе стоит произведение всех интенсивностей, стоящих у стрелок, ведущих слева – направо (с начала и до данного состояния S_k), а в знаменателе – произведение всех интенсивностей, стоящих у стрелок, ведущих справа налево (с начала и до S_k).

$$P_k = \frac{\lambda_{k-1,k}}{\lambda_{k,k-1}} \dots \frac{\lambda_{12}}{\lambda_{21}} \cdot \frac{\lambda_{01}}{\lambda_{10}} \cdot P_0$$

Заметим, что все вероятности состояний выражены через одну вероятность – вероятность нахождения системы в начальном состоянии - p_0 .

Подставим, эти выражения в нормировочное условие получим:

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

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

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Характеристики вычислительной системы

Рассмотрим систему с одним каналом и неограниченной очередью, в которой отсутствуют ограничения по длине очереди и по времени ожидания.

Интенсивность входного потока заявок λ , среднее время обслуживания $t_{обс}$.

Необходимо найти финальные вероятности состояний, а также:

$L_{сист}$ – среднее число заявок в системе;

$W_{сист}$ – среднее время пребывания заявки в системе;

$L_{очер}$ – среднее число заявок в очереди;

$W_{очер}$ – среднее время пребывания заявки в очереди;

$P_{зан}$ – вероятность того, что канал занят.

Состояния системы:

S_0 – канал свободен, очереди нет;

S_1 – канал занят, очереди нет;

S_2 – канал занят, 1 заявка в очереди;

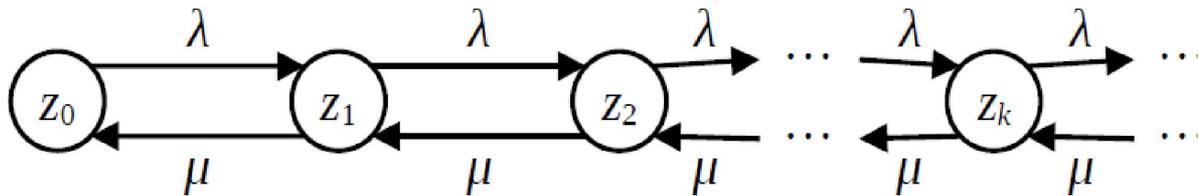
....

S_n – канал занят, $(n-1)$ заявка в очереди.

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Граф состояний вычислительной системы

Граф состояний одноканальной системы с бесконечной очередью приведен ниже. Приведенный граф описывает уже рассмотренную ранее модель «размножения и гибели», с бесконечным количеством состояний. Поэтому для описания работы данной системы используем уже полученные ранее формулы



$$P_0 = \frac{1}{\left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12}}{\lambda_{21}} \cdot \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{n-1,n}}{\lambda_{n,n-1}} \dots \frac{\lambda_{12}}{\lambda_{21}} \cdot \frac{\lambda_{01}}{\lambda_{10}} \right)}$$

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Коэффициент загрузки

Заметим, что среднее время обслуживания $t_{обс}$ и интенсивность обслуживания μ связаны следующей зависимостью:

$$\mu = 1/t_{обс}.$$

Заметим также, что абсолютную пропускную способность вычислять не надо – рано или поздно заявка будет обслужена, так как очередь неограничена; следовательно, $A = \lambda$. Относительная пропускная способность – вероятность того, что заявка будет обслужена, – $Q = 1$.

Предельная вероятность состояния p_0 может быть найдена из следующей формулы:

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

Если обозначить $\lambda/\mu = \rho$, то получим:

$$p_0 = \left(1 + \rho + \rho^2 + \dots + \rho^k + \dots \right)^{-1}$$

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Коэффициент загрузки

Из математики известно, что ряд в приведенной формуле сходится при $\rho < 1$.

При этом сумма ряда равна $1/(1-\rho)$. Тогда для финальной вероятности p_0 получим следующую формулу:

$$p_0 = 1 - \rho$$

Полученная вероятность p_0 означает, что канал обработки (сервер) свободен и очереди нет. Дополнительная величина к p_0 , равная $1 - p_0 = 1 - 1 + \rho = \rho$ означает, что канал занят обслуживанием, т.е. отношение $\rho = \lambda/\mu$ является мерой загрузки канала (сервера) и называется **коэффициентом загрузки**.

Таким образом, для такой системы финальные вероятности существуют, если величина $\rho = \lambda/\mu < 1$. Если $\rho \geq 1$, то очередь при $t \rightarrow \infty$ растет неограниченно.

Заметим, что СМО справляется с потоком заявок при $\rho = 1$, только если этот поток – регулярный, и время обслуживания тоже неслучайно, равно интервалу между заявками. В этом идеальном случае очередь в СМО вообще отсутствует, канал будет непрерывно занят и регулярно выпускать обслуженные заявки.

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Коэффициент загрузки

Финальные вероятности, как уже было сказано, образуют геометрическую прогрессию со знаменателем ρ .

Как это ни странно, максимальной из них оказывается p_0 – вероятность того, что канал будет вообще свободен. Как бы ни была нагружена система с очередью, если только она вообще справляется с потоком заявок ($\rho < 1$), самое вероятное число заявок в системе будет 0.

Далее вычислим среднее число заявок в системе.

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее число заявок в системе

Обозначим среднее число заявок в системе - L_{sys} .

Случайная величина Z – число заявок в системе – имеет возможные значения $0, 1, 2, 3, \dots, k$, с вероятностями $p_0, p_1, p_2, \dots, p_k$. Тогда будем иметь:

$$L_{sys} = 0p_0 + 1p_1 + 2p_2 + \dots + kp_k + \dots = \sum_{k=0}^{\infty} kp_k \quad (19.1)$$

Учитывая, что $p_k = \rho^k p_0 = \rho^k (1 - \rho)$ и подставляя в формулу (19.1) получим

$$L_{sys} = \sum_{k=0}^{\infty} kp_k = \sum_{k=1}^{\infty} k\rho^k (1 - \rho) = \rho(1 - \rho) \sum_{k=1}^{\infty} k\rho^{k-1}$$

Но величина $k\rho^{k-1}$ это производная по ρ от величины ρ^k

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее число заявок в системе

Учитывая последнее, можно записать формулу (19.1) в следующем виде:

$$L_{sys} = \rho(1-\rho) \sum_{k=1}^{\infty} k \rho^{k-1} = \rho(1-\rho) \sum_{k=1}^{\infty} \frac{d\rho^k}{d\rho} \quad (20.1)$$

Меняя местами операции дифференцирования и суммирования получим следующее выражение

$$L_{sys} = \rho(1-\rho) \sum_{k=1}^{\infty} \frac{d\rho^k}{d\rho} = \rho(1-\rho) \frac{d}{d\rho} \sum_{k=1}^{\infty} \rho^k \quad (20.2)$$

Учитывая, что сумма бесконечной убывающей прогрессии равна $1/(1-\rho)$, ее производная равна $1/(1-\rho)^2$, окончательно получим:

$$L_{sys} = \frac{\rho}{(1-\rho)} \quad (20.3)$$

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Длина очереди

Средняя длина очереди равна среднему числу заявок в системе минус среднее число обслуживаемых заявок.

Ранее было показано, что среднее число обслуживаемых заявок равно ρ .

Учитывая это, можно записать, что средняя длина очереди в СМО равна:

$$L = L_{sys} - \rho = \frac{\rho}{(1 - \rho)} - \rho = \frac{\rho^2}{(1 - \rho)}$$

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее время пребывания заявки в системе и в очереди

Выведем важную формулу, связывающую для предельного стационарного режима среднее число заявок L_{sys} , находящихся в системе массового обслуживания, и среднее время пребывания заявки в системе W_{sys} .

Пусть дана любая система массового обслуживания и связанные с ней два потока событий: поток заявок, прибывающих в систему, и поток заявок, покидающих СМО. Если в системе установился предельный стационарный режим, то среднее число заявок, прибывающих в систему за единицу времени, равно среднему числу заявок, покидающих ее: оба потока имеют одну и ту же интенсивность λ .

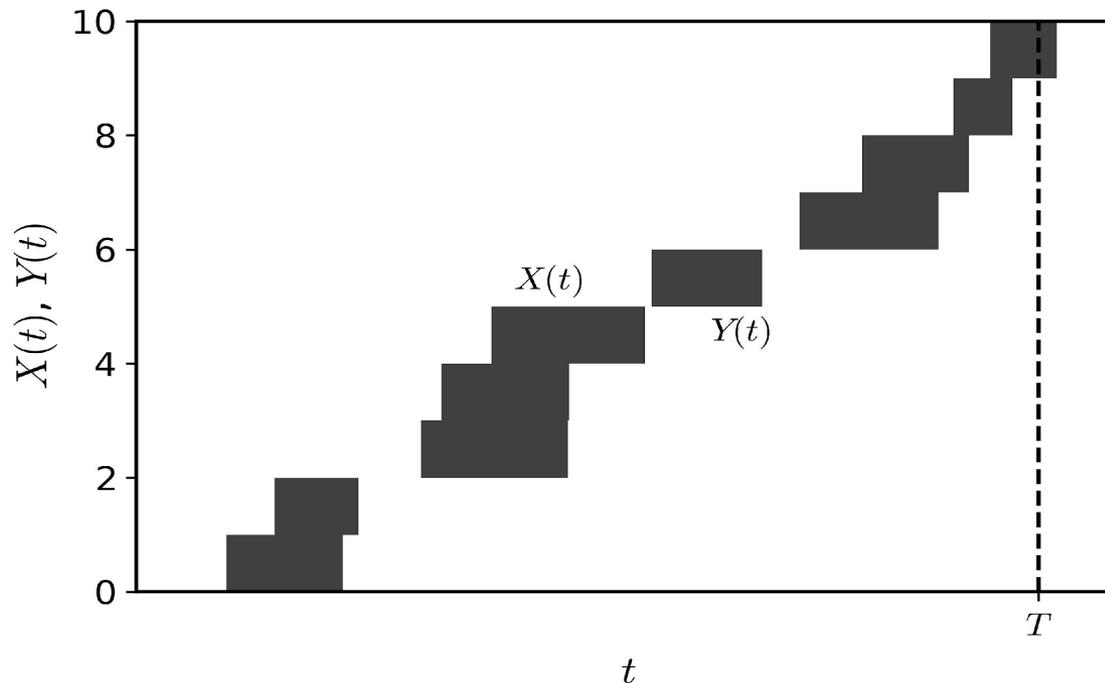
Пусть $X(t)$ – число заявок, прибывших в СМО до момента t , $Y(t)$ – число заявок, покинувших систему до момента t . И та, и другая функция являются случайными и меняются скачком: в моменты прихода заявок $X(t)$ увеличивается на 1, $Y(t)$ уменьшается на 1 в моменты ухода заявки.

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее время пребывания заявки в системе и в очереди

На временной диаграмме процесса поступления и ухода заявок изображены функции X , Y . Обе линии – ступенчатые; верхняя – $X(t)$, нижняя – $Y(t)$.

Очевидно, что для любого момента времени их разность $Z(t) = X(t) - Y(t)$ – есть не что иное, как число заявок, находящихся в СМО.



2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее время пребывания заявки в системе и в очереди

Рассмотрим очень большой промежуток времени T и вычислим для него среднее число заявок, находящихся в СМО. Оно будет равно интегралу от функции $Z(t)$ на этом промежутке, деленному на длину интервала T :

$$L_{sys} = \frac{1}{T} \cdot \int_0^T Z(t) dt$$

Но этот интеграл есть не что иное, как площадь заштрихованной фигуры, приведенной на слайде № 23.

Фигура состоит из прямоугольников, высота которых равна 1, а основание – равно времени пребывания соответствующей заявки в системе. Обозначим эти времена $t_1, t_2, t_3 \dots t_n$

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее время пребывания заявки в системе и в очереди

Таким образом, можно считать, что,

$$\int_0^T Z(t) dt = \sum_{i=0}^n t_i$$

где сумма распространяется на все заявки, пришедшие за время T . Разделим правую и левую часть данного выражения на T получим:

$$\frac{\int_0^T Z(t) dt}{T} = \frac{\sum_{i=0}^n t_i}{T}$$

В левой части уравнения есть среднее число заявок, находящихся в СМО :

$$L_{sys} = \frac{1}{T} \cdot \int_0^T Z(t) dt \quad \text{Отсюда получим}$$

$$L_{sys} = \frac{\int_0^T Z(t) dt}{T} = \frac{\sum_{i=0}^n t_i}{T}$$

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее время пребывания заявки в системе и в очереди

Разделим и умножим правую часть полученного ранее выражения на λ , тогда получим:

$$L_{sys} = \frac{\int_0^T Z(t) dt}{T} = \frac{\sum_{i=0}^n t_i \cdot \lambda}{T \cdot \lambda} = W_{sys} \cdot \lambda$$

Но $T\lambda$ -- есть не что иное, как среднее число заявок, пришедшее за время T . Если мы разделим сумму всех времен t_i на среднее число заявок, то получим среднее время пребывания заявки в системе W_{sys} . Итак, среднее время пребывания заявки в системе составит:

$$W_{sys} = \frac{L_{sys}}{\lambda} = \frac{\rho}{\lambda(1-\rho)}$$

А среднее время пребывания заявки в очереди составит:

$$W_L = \frac{L}{\lambda} = \frac{\rho^2}{\lambda(1-\rho)}$$

2. Модель вычислительной системы в виде одноканальной СМО с очередью

Среднее время пребывания заявки в системе и в очереди

Это и есть известная формула Литтла:

для системы массового обслуживания, при любом характере потока заявок, при любом распределении времени обслуживания, при любой дисциплине обслуживания среднее время пребывания заявки в системе равно среднему числу заявок в системе, деленному на интенсивность потока заявок.

$$W_{sys} = \frac{L_{sys}}{\lambda} = \frac{\rho}{\lambda(1-\rho)}$$

А среднее время пребывания заявки в очереди составит:

$$W_L = \frac{L}{\lambda} = \frac{\rho^2}{\lambda(1-\rho)}$$

3. Примеры решения задач

1. Найти абсолютную пропускную способность одноканальной ВС с очередью и среднее время нахождения заявки в очереди, если известно, что среднее число программ в очереди составило $L=10$, а интенсивность обработки программ сервером составила $\mu=1,0$ [1/сек].

Решение.

Используем следующую формулу для очереди L одноканальной ВС:

$$L = L_{sys} - \rho = \frac{\rho}{(1-\rho)} - \rho = \frac{\rho^2}{(1-\rho)}$$

Разрешим приведенное уравнение относительно ρ , получим следующее значение $\rho=0,916$;

Учитывая, что $\rho=\lambda/\mu$, найдем значение $\lambda= \rho*\mu=0,916*1=0,916$.

Принимая во внимание, что для одноканальной ВС с неограниченной очередью абсолютная пропускная способность равна интенсивности входного потока, т.е. все поступающие заявки будут обслужены, то тогда $A= \lambda=0,916$.

Используя формулу Литтла, найдем среднее время нахождения заявки в очереди:

$$W=L/ \lambda=10/0,916= 10,916$$