ПРИМЕР выполнения и оформления заданий по СМО

СМО представляет собой техническое устройство, состоящее из двух узлов, которые могут независимо друг от друга выходить из строя. Граф системы массового обслуживания представлен на рис. 10

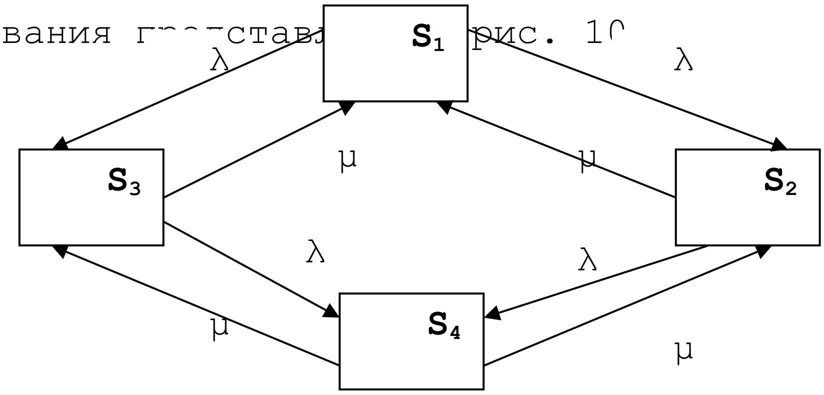


Рис. 10.

СМО может находиться в следующих состояниях:

S_1 – оба узла исправны, техническое устройство выполняет свои функции;

S_2 – первый узел исправен и работает, второй – неисправен;

S_3 – первый узел неисправен, второй узел исправен;

S_4 – оба узла неисправны, техническое устройство ремонтируется.

λ_1 – поток неисправностей первого узла;

μ_1 – поток ремонтов первого узла;

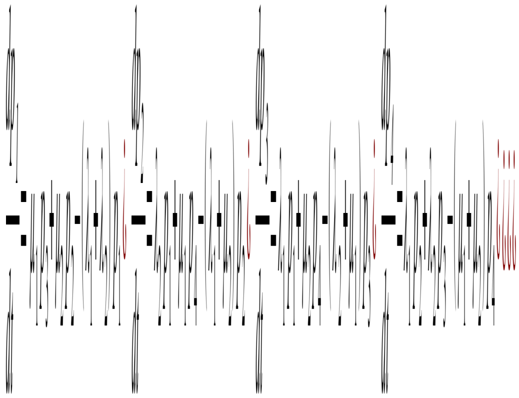
λ_2 – поток неисправностей второго узла;

μ_2 – поток ремонтов второго узла.

Каждому из состояний можно поставить в соответствие вероятность нахождения СМО в данном состоянии:

- S_1 соответствует вероятность p_1;
- S_2 соответствует вероятность p_2;
- S_3 соответствует вероятность p_3;
- S_4 соответствует вероятность p_4.

По графу системы массового обслуживания составляем систему уравнений Колмогорова (1), по которой можно определить вероятности состояний системы. Для этого следует исключить одно из уравнений системы, затем решить полученную неоднородную систему линейных уравнений любым методом (Крамера, Гаусса).



(1)

Рассмотрим пример имитационной модели, позволяющей исследовать описанную выше СМО. Имитационная модель разработана средствами VBA for Excel и табличного процессора MS Excel. Интерфейс имитационной модели представлен на рис. 1.

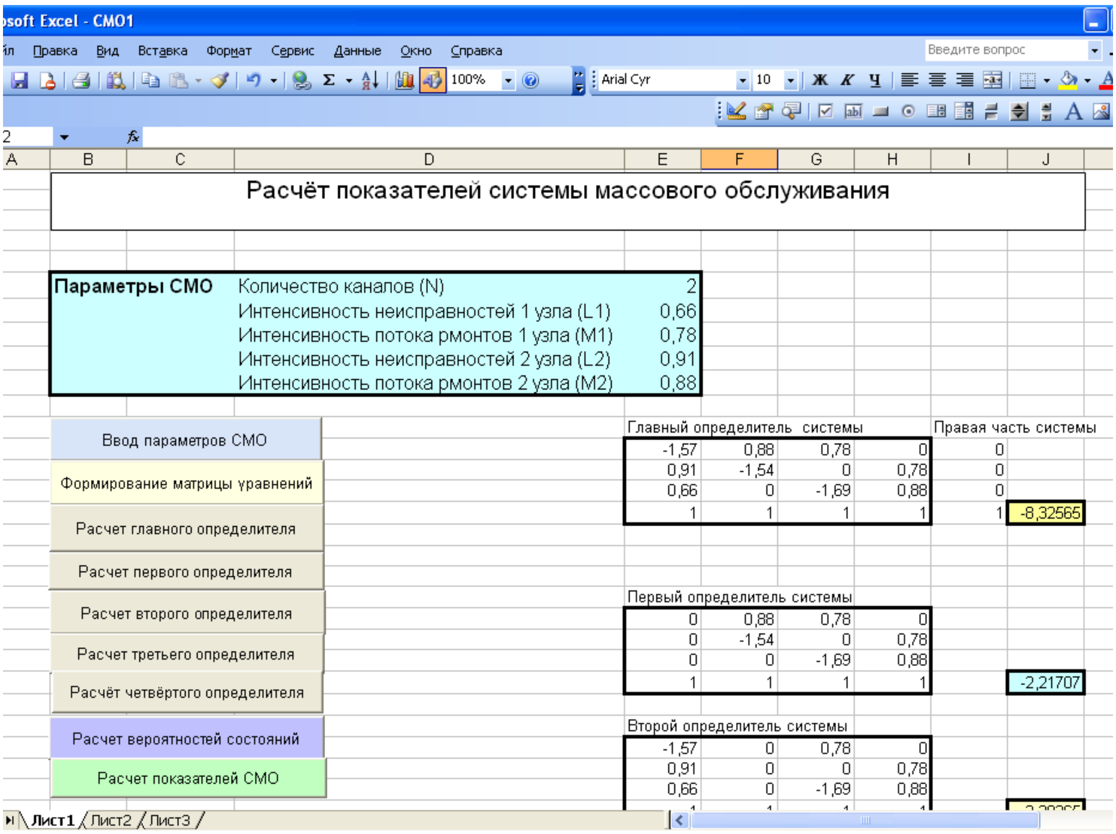


Рис. 1. Интерфейс имитационной модели

Основными элементами управления имитационной модели исследования параметров СМО являются кнопки, с которыми связаны соответствующие макросы. Основные функции и операции, реализованные в имитационной модели СМО:

- ввод показателей СМО (количества каналов обслуживания, интенсивности потоков неисправностей узлов технического устройства, интенсивности потоков ремонтов узлов технического устройства);
- формирование системы уравнений Колмогорова (см. систему уравнений 1);
- расчёт главного определителя системы уравнений Колмогорова;
- расчёт первого определителя системы уравнений Колмогорова;
- расчёт второго определителя системы уравнений Колмогорова;
- расчёт третьего определителя системы уравнений Колмогорова;
- расчёт четвёртого определителя системы уравнений Колмогорова;
- расчёт вероятностей состояний СМО: p_1, p_2, p_3, p_4;
- расчёт относительной пропускной способности СМО (q).

Для ввода параметров СМО используется объект User-Form1, который вызывается по нажатию одноимённой кнопки (см. рис. 2).

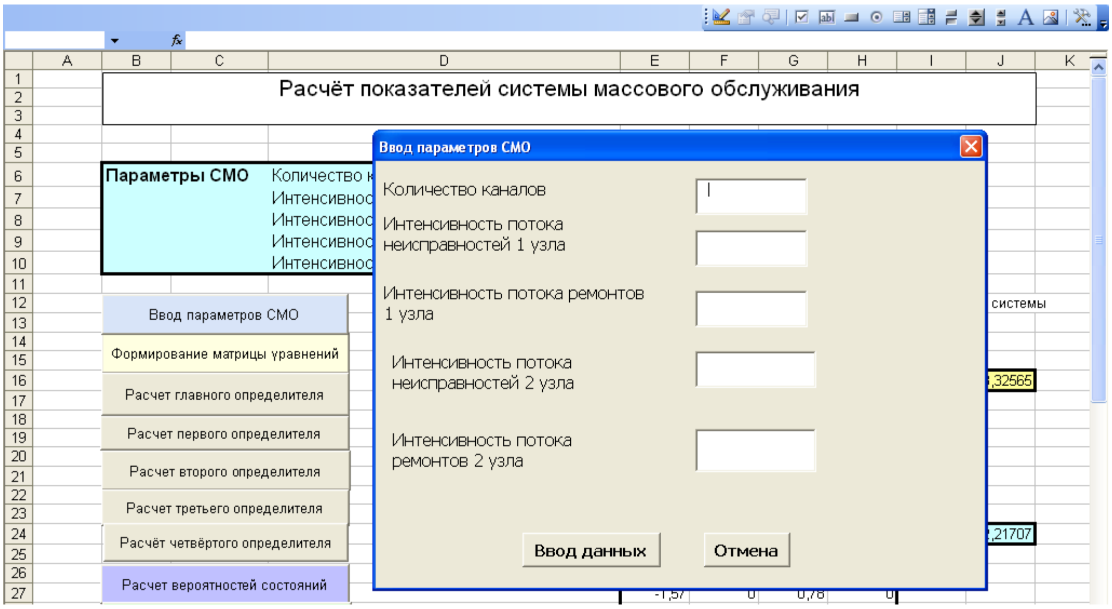


Рис. 2. Форма ввода параметров СМО

Для выполнения все указанных выше функций имитационной модели записываются макросы в автоматическом режиме. Для этого выбирают пункт меню «Сервис», закладку «Макрос» - «Начать запись». Именуют макрос, затем выполняют требуемую для расчётов последовательность действий, после чего нажимают кнопку или закладку «Макрос» - «Остановить запись». Для расчёта определителей используется математическая функция МОПРЕД(), которая в режиме макроса вычисляет нужный определитель.

Статические эксперименты в построенной моделью двухканальной СМО с отказами можно выполнять, варьируя параметры СМО (интенсивности потоков неисправностей узлов технического устройства и потоков их ремонтов). Можно пронаблюдать как изменение параметров СМО влияет на результаты её функционирования – на вероятности состояний, относительную пропускную способность.

В повседневной жизни к системам массового обслуживания относятся телефонные и автозаправочные станции, билетные кассы, торговые предприятия, парикмахерские, мастерские и т.п. В таких системах два основных потока: входной — поток заявок и выходной поток обслуживания. *Поток заявок* образуют клиенты (покупатели), желающие приобрести какой-либо товар. *Выходной поток* образуют продавцы, обслуживающие покупателей. Если интенсивность обслуживания мала, то образуется очередь. Последнюю можно ликвидировать или быстро сократить, используя несколько каналов обслуживания (несколько телефонных аппаратов, билетных касс, торговых точек и т.д.).

Для того, чтобы понять, как решаются такие задачи массового обслуживания, рассмотрим сначала основные понятия и определения.

Теория систем массового обслуживания (СМО) впервые была разработана датским математиком А. К. Эрлангом применительно к запросам, поступающим на телефонную станцию. Поэтому основные понятия и определения сохраняются из практики обеспечения телефонной связи независимо от фактического назначения конкретной СМО.

Системы массового обслуживания предназначены для обслуживания потока заявок или требований, поступающих на вход в случайные моменты времени. Каждая СМО состоит из некоторого числа каналов обслуживания, в качестве которых в зависимости от вида системы могут выступать: линии связи, приемные пункты, рабочие точки, подъездные пути, испытательные стенды, технологические агрегаты, ремонтные бригады и т.д. Выполнение поступившей заявки, т.е ее обслуживание, продолжается некоторое время (тоже случайное), после чего канал освобождается и готов принять следующую заявку.

Поступающие на вход системы массового обслуживания требования-заявки следуют одно за другим и образуют непрерывный поток событий. Конечно, невозможно заранее предсказать, например, когда какому-то абоненту вздумается позвонить по телефону своему партнеру, но если рассматривать всех абонентов телефонной станции, то несмотря на случайный характер каждого отдельного события, за 1 час (60 мин) было, например, 30 телефонных вызовов, то в среднем одна заявка приходится на интервал в 2 мин. Следовательно, среднее число событий в единицу времени — интенсивность потока λ — будет равна 0.5.

В простейшем потоке интенсивность является постоянной величиной, т.е $\lambda=const$ во времени. Такие простейшие потоки называются стационарными.

Системы массового обслуживания могут быть двух типов: СМО *с отказами,* в которых заявка, поступившая в тот момент, когда все каналы заняты, получает отказ и не обслуживается; СМО *с ожиданием,* в которых каждая заявка, прибывшая в систему, когда в ней нет свободных каналов, остается и ожидает, пока освободится какой-нибудь канал и ее возьмут на обслуживание. По аналогии с системами обслуживания населения ожидающие заявки называют *очередью*.

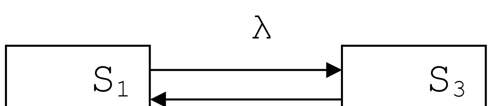
Основы теории СМО рассмотрим на примере. Пусть производственная система состоит из двух устройств, каждое из которых производит одну и ту же продукцию. Устройства в ходе работы могут выйти из строя (отказать). Отказавшее устройство немедленно начинают ремонтировать. Рассматриваемая система имеет четыре состояния:

S_1 — оба устройства работают;

S_2 — первое устройство ремонтируется (после отказа), второе работает;

S_3 — второе ремонтируется, первое работает; S_4 — оба ремонтируются.

Граф состояний такой производственной системы будет



следующим (рис. 3):

μ

Рис. 3. Граф состояний производственной системы

Переходы S_1 - S_2; S_2 – S_4; S_1 - S_3; S_3 – S_4 совершаются в результате происходящих в системе отказов. Обратные переходы являются следствием ремонтных работ. Отказы и окончания — являются случайными величинами.

Пусть λ_1 — интенсивность потока отказов первого устройства; λ_2— интенсивность потока отказов второго устройства; μ_1 — интенсивность потока окончаний ремонтов первого устройства; μ_2 — интенсивность потока окончаний ремонтов второго устройства.

Рассмотрим конкретное состояние, например S_1. Из этого состояния возможны переходы в состояния S_2 и S_3 — с суммарной вероятностью $\lambda_1 + \lambda_2$, отнесенной к единице времени. В стационарном режиме интенсивность потока событий равна вероятности за конечный промежуток времени, деленной на этот промежуток времени. Таким образом, число уходов из состояния S_1, в единицу времени в рассматриваемом коллективе систем равно: $N\,p_i(\lambda_1 + \lambda_2)$.

Здесь видно общее правило: совершаемое в единицу времени число переходов S_i в S_j равно произведению числа систем в состоянии S_i (в исходном состоянии) на вероятность перехода, отнесенную к единице времени. Мы рассмотрели уходы из состояния S_1.

Приходы в это состояние совершаются из S_2 и S_3. Поскольку рассматривается стационарный режим, то числа уходов и приходов для каждого состояния должны быть сбалансированы. Следовательно: $Np_1(\lambda_1 + \lambda_2) = Np_2\,\mu_1 + Np_3\,\mu_2.$

Рассматривая баланс уходов и приходов для каждого из четырех состояний и сокращая в уравнениях общий множитель N, получаем следующие уравнения относительно вероятностей p_1, p_2, p_3, p_4:

$$S_1\colon (\lambda_1 + \lambda_2)\, p_1 = \mu_1 p_2 + \mu_2 p_3.$$
$$S_2\colon (\lambda_2 + \mu_1)\, p_2 = \lambda_1 p_1 + \mu_2 p_4.$$
$$S_3\colon (\lambda_1 + \mu_2)\, p_3 = \lambda_2 p_1 + \mu_1 p_4.$$
$$S_4\colon (\mu_1 + \mu_2)\, p_4 = \lambda_2 p_2 + \lambda_1 p_3.$$

Нетрудно убедиться, что четвертое уравнение может быть получено сложением первых трех. Вместо этого уравнения воспользуемся уравнением: $p_1 + p_2 + p_3 + p_4 = 1$, которое означает, что система с достоверностью находится в каком-либо из четырех состояний. Таким образом, приходим к системе уравнений:

$$\left.\begin{aligned} (\lambda_1 + \lambda_2)\, p_1 &= \mu_1 p_2 + \mu_2 p_3; \\ (\lambda_2 + \mu_1)\, p_2 &= \lambda_1 p_1 + \mu_2 p_4; \\ (\lambda_1 + \mu_2)\, p_3 &= \lambda_2 p_1 + \mu_1 p_4; \\ p_1 + p_2 + p_3 + p_4 &= 1. \end{aligned}\right\}$$

Это уравнения Колмогорова, записанные для системы, граф состояний которой показан на рис. 13. Рассуждая аналогичным образом, можно составить уравнения Колмогорова и для других СМО.

Примеры задач с СМО

Задача № 1. Имеется производственная система, производящая некоторую продукцию. Граф состояний такой системы показан на рис. 13. Предположим, что второе устройство в данной системе более современное и имеет производительность вдвое более высокую, чем первое устройство. Первое устройство приносит в единицу

$$3p_1 = 2p_2 + 3p_3,$$

времени доход, равный 5 условным единицам, а второе — 10 единицам. Отказы второго устройства происходят в среднем вдвое чаше, чем первого; поэтому положим, что $\lambda_1 = 1$, $\lambda_2 = 2$. Интенсивности потоков окончаний ремонтов примем равными $\mu_1 = 2$, $\mu_2 = 3$. Используя заданные интенсивности отказов и потоков окончаний ремонтов, перепишем уравнения Колмогорова в виде:

Решая эту систему уравнений, находим: $p_1 = 0.4$; $p_2 = 0.2$; $p_3 = 0.27$; $p_4 = 0.13$, Это означает, что в среднем 40% времени оба устройства работают одновременно (состояние S_1 на рис. 13); 20% времени работает только первое устройство, а второе при этом ремонтируется (состояние S_2); 21% времени работает только второе устройство, а первое при этом ремонтируется (состояние S_3); 13% времени оба устройства одновременно находятся в состоянии ремонта (состояние S_4). Нетрудно подсчитать доход, которой дает система из двух рассматриваемых устройств в единицу времени:

$$(5 + 10) \times 0.4 + 5 * 0.2 + 10 \times 0.27 = 9.7 \text{ усл. ед.}$$

Предположим, что предлагается некоторая рационализация, позволяющая вдвое сократить время ремонта либо первого, либо второго устройства. По ряду причин рационализацию можно применить только к одному из устройств. Спрашивается, какое устройство следует выбрать, первое или второе? Это конкретный пример практической ситуации, когда пользуясь теорией массового обслуживания, надо обосновать принятие решения.

Допустим, что выбирается первое устройство. В результате рационализации интенсивность потока окончаний ремонтов этого устройства увеличивается вдвое, так что теперь $\mu_1 = 4$, а остальные интенсивности остаются прежними: $\lambda_1=1$, $\lambda_2= 2$, $\mu_2= 3$. Уравнения Колмогорова принимают теперь следующий вид:

$$\left.\begin{aligned} 3p_1 &= 4p_2 + 3p_3 , \\ 6p_2 &= p_1 + 3p_4 , \\ 4p_3 &= 2p_1 + 4p_4 , \\ p_1 + p_2 + p_3 + p_4 &= 1 . \end{aligned}\right\}$$

Решая эту систему, находим: p_1 = 0.48; p_2= 0.12; p_3= 0.32; p_4 = 0.08. С учетом полученных вероятностей определим доход, который теперь будет давать рассматриваемая система:

(5 + 10) x 0.48 + 5 x 0.12 + 10 x 0.32 = 11 усл. ед.

Если же мы выберем второе устройство, то в результате рационализации удвоится интенсивность μ_2 . В этом случае: $\lambda_1 = 1$, $\lambda_2 = 2$, $\mu_1 = 2$, $\mu_2 = 6$. Уравнения Колмогорова примут вид:

$$\left.\begin{aligned} 3p_1 &= 2p_2 + 6p_3 , \\ 4p_2 &= p_1 + 6p_4 , \\ 7p_3 &= 2p_1 + 2p_4 , \\ p_1 + p_2 + p_3 + p_4 &= 1 . \end{aligned}\right\}$$

Решая эту систему уравнений, находим: p_1 = 0.5; p_2 = 0.25; p_3 = 0.17; p_4 = =0.08. Подсчитываем доход: (5+ 10) x 0.5 + 5 x 0.25+ 10 x 0.17= 10.45 усл. ед.

Таким образом, мы видим, что выгоднее применить, рационализацию к первому устройству.

Теперь рассмотрим систему массового обслуживания с отказами. Самый простой пример СМО с отказами — это автоматическая телефонная станция. Если вызываемый абонент занят, то даются короткие гудки и ожидать бесполезно. В зависимости от степени необходимости в обслуживании заявки либо покидают систему, либо обращаются повторно. *Одноканальная система массового* *обслуживания* — это самая простая СМО, на которой можно рассмотреть основные закономерности ее работы.

На вход системы поступает поток заявок с интенсивностью λ. Заявка, поступившая в момент, когда система свободна, сразу же берется на обслуживание. Следующая заявка, прибывшая в момент, когда канал обслуживания занят, получает отказ. Время обслуживания заявки имеет случайную продолжительность, но имеется какое-то среднее значение, в результате чего на выходе образуется поток обслуживания с интенсивностью μ. Наглядно поток обслуживания можно представить таким образом, что если бы канал обслуживания был непрерывно загружен, то из него выходил бы поток обслуженных заявок (рис. 4).

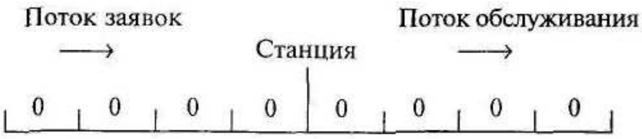


Рис. 4. Поток заявок и поток обслуживания в системе массового обслуживания

Если среднее время обслуживания одной заявки (в примере с телефонной станцией это средняя продолжительность одного разговора) составляет 0.5 мин, то интенсивность потока обслуживания μ = 1/0.5 = 2. Одноканальная СМО может находиться только в одном из двух состояний: S_o — свободна, S_1, — занята. Граф состояний, показывающий возможные переходы из одного состояния в другое, изображен на рис. 5.

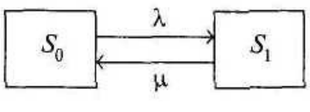


Рис. 5. Граф одноканальной СМО с отказами

Возможность нахождения СМО в свободном состоянии S_0 определяется какой-то, пока нам неизвестной вероятностью

p_0. Соответственно p_1 — это вероятность того, что система находится в занятом состоянии S_1. Так как система может находиться только в одном из двух состояний, то в каком-то из них она всегда находится, поэтому сумма вероятностей равна единице: $p_0 + p_1 = 1$.

Чтобы система могла пребывать в этих двух состояниях, воздействия, выводящие ее из состояния S_0, должны уравновешиваться воздействиями, возвращающими систему обратно в это состояние. Величина каждого воздействия определяется произведением интенсивности потока на соответствующую вероятность, т.е. $\lambda p_0 = \mu p_1$. Из этого выражения определяем: $p_1 = \frac{\lambda}{\mu} p_0$.

Учитывая, что сумма вероятностей всегда равна

$$p_0 = \frac{1}{1 + \lambda/\mu} = \frac{\mu}{\lambda + \mu}.$$

$$p_1 = \frac{\lambda}{\mu}\ p_0 = \frac{\lambda}{\lambda + \mu}.$$

единице, получим:

$$p_0 + \frac{\lambda}{\mu} p_0 = 1.$$

Основные параметры СМО с отказами: относительная пропускная способность и абсолютная пропускная способность, а также вероятность получения отказа. *Относительная пропускная способность* q определяется вероятностью того, что в момент заявки канал свободен и она будет обслужена, т.е. для одноканальной системы $q = p_0$. *В* пределе, когда процесс уже установился значение относительной пропускной способности СМО будет равно

$$q = \frac{\mu}{\lambda + \mu}.$$

Абсолютная пропускная способность A определяется произведением относительной пропускной способности на интенсивность потока требований: $A = q\,\lambda$. В пределе она становится равной $A = \lambda\mu / (\lambda + \mu)$.

Вероятность того, что заявка будет обслужена, определяется p_0, а *вероятность отказа* — p_1,. Таким образом, вероятность того, что канал будет занят:

$$p_{отк.} = p_1 = \lambda / (\lambda + \mu) .$$

Другой пример. Сборочный участок производит в один час 90 блоков, т.е. интенсивность потока λ = 1.5 блоков в 1 мин. На этом участке работает контролер, который выборочно проверяет изготовленные блоки аппаратуры, средняя продолжительность контрольных операций *s* = 1.25 мин. Если в момент прибытия очередного блока контролер занят, то этот блок сразу же перелается на дальнейшие операции без промежуточного контроля. Производство непрерывное и продолжается до обнаружения дефекта в одном из блоков, в этом случае технологический процесс останавливается и выясняются причины неисправности.

Необходимо определить, какая часть выпускаемой продукции в таких условиях подвергается контролю и какая часть пропускается на дальнейшие операции без контроля (т.е. какая часть получает отказ от прохождения контрольных операций).

Определим параметр μ потока обслуживания μ = 1/1.25 = 0.8. Относительная пропускная способность *q* = 0.8 : (1.5 + 0.8) = 0.348. Таким образом, контрольным операциям будет подвергаться менее 35% продукции участка. Абсолютная пропускная способность *A* = 1.5 * 0.348 =

0.52. Вероятность отказа в обслуживании, т.е. пропуска на дальнейшую обработку без контроля, равна (1 - 0.35) = 0.65.

Интересно, что если увеличить производительность труда контролера и таким образом снизить продолжительность контрольных операций, то пропускная способность системы, конечно, повысится, однако далеко не до такой степени, как может показаться на первый взгляд. Допустим, что с оснащением контроля новым, более производительным оборудованием s снизилось в 2 раза и соответственно в 2 раза увеличилась интенсивность потока обслуживания, т.е. $\mu = 1.6$. Тогда при той же интенсивности потока заявок получим

$q = 1.6/(1.5 + 1.6) = 0.516$, т.е. контролироваться будет около 52% всех изделий, а не 70%, как можно было бы ожидать.

Рассмотрим теперь *многоканальные системы массового обслуживания.*

Для повышения пропускной способности СМО надо увеличить число каналов обслуживания, т.е. число линий связи в телефонной системе, количество контролеров на производстве и т.д. Для потребителя это будет удобно, но общая эффективность системы при этом может снизиться, так как каждый новый канал требует дополнительных затрат на установку и обслуживание.

Граф двухканальной системы массового обслуживания с отказами будет иметь вид, показанный на рис. 6.

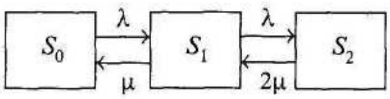


Рис. 6. Граф двухканальной системы массового обслуживания

Состояние S_1 — это состояние, когда в СМО имеется одна заявка и один канал занят, а второй свободен. Из состояния S_0 в состояние S_1 систему переводит поток заявок с интенсивностью λ. Как только приходит первая заявка, один канал становится занятым, тот же поток переводит СМО из первого состояния во второе, когда заняты оба канала и следующим заявкам будет даваться отказ.

Если в системе занят один канал, то этот канал производит μ обслуживании в единицу времени. Теперь пусть система находится в состоянии S_2, т.е. в ней работают два канала. В состояние S_1 система будет переходить, если обслуживание закончил либо первый, либо второй канал. Таким образом, суммарная интенсивность потока обслуживания будет равна 2.
Для состояния S_0 баланс воздействий будет, $\lambda\ p_0 = \mu\ p_1$ откуда получим:

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

Воздействия, выводящие из состояния S_1 (стрелки, направленные из S_1), будут равны $\lambda\ p_1 + \mu\ p_1$. Они компенсируются воздействиями, приводящими в это состояние (стрелки, направленные внутрь S_1): $\lambda p_0 + 2\mu p_2$.

Баланс воздействий будет равен: $2\mu p_2 + \lambda p_0 = \mu p_1 + \lambda p_1$.

С учетом того, что $\lambda\ p_0 = \mu\ p_1$ получим: $2\mu p_2 = \lambda p_1$
или иначе

$$p_2 = \frac{\lambda}{2\mu} p_1 = \frac{\lambda}{2\mu}\frac{\lambda}{\mu} p_0 = \frac{\lambda^2}{2\mu^2} p_0 .$$

Так как сумма всех вероятностей по-прежнему должна равняться единице, получаем:

$$p_0 + \frac{\lambda}{\mu} p_0 + \frac{\lambda^2}{2\,\mu^2} p_0 = 1.$$

Откуда следует:

$$p_0 = \frac{1}{1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2}}; \quad p_2 = \frac{\lambda^2}{2\,\mu^2} p_0 .$$

Произведя по полученным формулам соответствующие расчеты из предыдущего примера, получим q = 62%. Таким образом, производительность двух контролеров больше, чем одного, работающего в 2 раза быстрее.

Граф трехканальной системы массового обслуживания с отказами имеет вид, как на рис. 7.

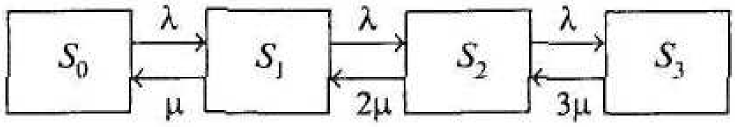


Рис. 7. Граф трехканальной СМО с отказами

Повторяя рассуждения, аналогичные предыдущим, можно получить:

$$p_0 = \frac{1}{1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \frac{\lambda^3}{3\mu^3}}; \quad p_3 = \frac{\lambda^3}{2 \times 3\,\mu^3} p_0 .$$

Проделав расчеты с данными для предыдущего примера, в случае трехканальной системы получим q = 81%. Для многоканальных СМО вводится еще один параметр — среднее число занятых каналов $k_{ср} = А/\mu = \lambda\, q/\mu.$

Для системы контроля с тремя контролерами получим $k_{ср}$ = 1.52. Таким образом, работы не хватает для загрузки даже двух контролеров, но все три не обеспечивают 100%-ную проверку всей выпускаемой продукции. Причина такого положения заключается в случайном характере поступления изделий на контроль.

Можно проверить, что получится, если увеличить число контролеров. Хотя, наверное, уже очевидно, что подобный подход явно нельзя назвать эффективным.

На рис. 8 изображен граф n-канальной системы массового обслуживания с отказами.

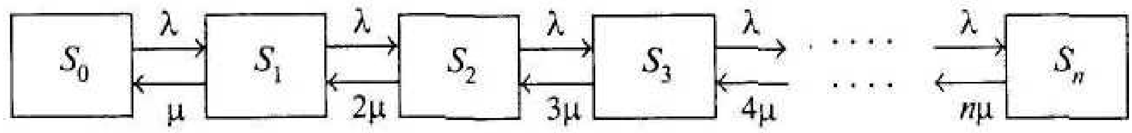


Рис. 8. Потоки в многоканальной системе массового обслуживания

Такой же процедурой, которая применялась для 2- и 3-канальных СМО, можно получить:

$$p_0 = \frac{1}{1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \ldots \frac{\lambda^n}{n!\,\mu^n}}; \quad p_n = \frac{\lambda^n}{n!\,\mu^n}\, p_0 \,.$$

где $n! = 1 \times 2 \times 3 \ldots n$.

Вероятность отказа равна p_n, а относительная пропускная способность: $q = 1 - P_n$.

В производственной системе с четырьмя контролерами и при тех же интенсивностях потоков, которые указаны в этом примере, получим q = 92%, а среднее число занятых каналов k — 1.75.

Теперь должно быть ясно, что 100% - ной проверки всей продукции таким путем не добиться. Следовательно, необходимо изменить систему обслуживания и перейти к СМО с ожиданием.

Системы массового обслуживания СМО с ожиданием

Рассмотрим СМО с одним каналом, на вход которого требования поступают с интенсивностью λ. Заявка, поступившая в момент, когда канал занят, не покидает систему, а становится в очередь и ожидает. Граф состояний такой системы показан на рис. 9.

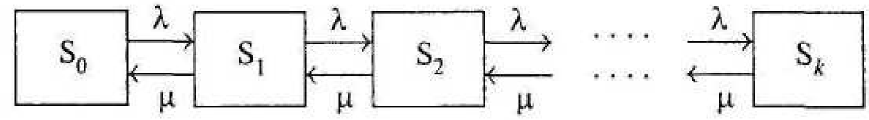


Рис. 9. Граф системы массового обслуживания с ожиданием

Состояние S_0 соответствует свободному каналу; S_1 означает, что канал занят, но очереди нет; S_2 - канал занят и одна заявка стоит в очереди; S_3 - в очереди две заявки и т.д. В состоянии S_k например, канал занят и $(k - 1)$ заявок ожидают обслуживания. По стрелкам слева направо систему из одного состояния в другое переводит поток заявок **с** интенсивностью λ, а по стрелкам справа налево переводит поток обслуживании, имеющий интенсивность μ. Всякий раз при переходе из одного состояния в другое очередь изменяется на единицу.

Для получения вероятности начального состояния можно использовать уравнение $\lambda\ p_0 = \mu\ p_1$, откуда $p_1 = (\lambda/\mu)\ p_0$. Величину λ/μ называют интенсивностью нагрузки СМО. в дальнейшем будем обозначать ее ρ. Для устойчивой работы СМО с ожиданием необходимо, чтобы средняя интенсивность потока обслуживания была больше интенсивности потока заявок, т.е $\mu > \lambda$ и, следовательно, $\rho < 1$. Если же $\lambda > \mu$, то система не справится **с** обслуживанием и очередь будет расти до бесконечности.

Используя введенные обозначения, вероятность состояния S_1 можно записать в виде: $p_1 = \rho\ p_0$. Чтобы получить вероятности p_2 и p_3 можно использовать полученные ранее выражения: $p = \rho^2\ p_0$, $p_3 = \rho^3\ p_0$. Аналогично можно получить выражение для произвольного члена: $p_k = \rho^k\ p_0$.

Для определения p_0 напишем выражение для суммы вероятностей:

$$p_0 + \rho p_0 + \rho^2 p_0 + \ldots + \rho^k p_0 = 1.$$

Величина $1 + \rho + \rho^2 + \ldots + \rho^k$ представляет собой сумму членов геометрической прогрессии, она равна $1/(1 - \rho)$. Поэтому $p_0 = 1 - \rho$, откуда получаем $p_к = \rho^k (1 - \rho)$.

Используя это выражение, можно определить характеристики системы массового обслуживания с ожиданием, существенные для ее функционирования: среднюю длину очереди, среднее число заявок в системе, среднее время пребывания в системе и вероятность образования очереди.

С вероятностью p_2 в очереди стоит одна заявка, с вероятностью p_3 — две заявки и с вероятностью p_k в очереди находится $(k - 1)$ заявок.

Следовательно,

$$L_{ср} = 1p_2 + 2p_3 + \ldots + (k-1)\, p_k =$$
$$= \rho^2 (1-\rho)\,(1 + 2\rho + 3\rho^2 + \ldots + k\,\rho^{k-1})\,.$$

Сумма геометрической прогрессии $1 + 2\rho + 3\rho^2 + \ldots$ равна $1/(1 - \rho)^2$, поэтому

$$L_{ср} = \frac{\rho^2}{1-\rho} = \frac{\lambda}{\mu\,(\mu-\lambda)}.$$

Среднее число заявок, находящихся в системе обслуживания, состоит из среднего числа находящихся в очереди и среднего числа находящихся на обслуживании, включая интервалы, когда очереди не было. Эта величина ей принимает значение 0, если канал свободен. Вероятность такого состояния равна $p_0 = 1 - \rho$. Если канал занят, значит заявки обслуживаются, и ω принимает значение 1. Вероятность этого равна $1 - p_0 = \rho$. Следовательно,

$$\omega = L_{ср} + o = \frac{\rho^2}{1-\rho} + \rho = \frac{\rho}{1-\rho}.$$

Среднее время ожидания в очереди равна

$$T_{ож} = \frac{1}{\mu}\,\frac{\rho}{1-\rho}.$$

$$\rho^2$$

среднему числу заявок в очереди, деленному на интенсивность потока обслуживания в одной или в другой форме:

Вероятность образования очереди равна вероятности того, что в системе будет более одного требования, т.е.

$$p_k = 1 - p_0 - p_1 = 1 - (1 - \rho) - \rho\ (1 - \rho) = \rho^2.$$

Рассмотрим такую же систему контроля продукции, которая была в СМО с отказами, но теперь установим такой порядок, при котором контролер проверяет всю продукцию. Если контролер будет занят, блоки ожидают, пока он освободится. Интенсивность нагрузки в первом случае будет: $\rho = \lambda/\mu = 1.5/0.8 = 1.87 > 1.$

При указанных условиях данный режим контроля невозможен, поскольку будет непрерывно возрастать. Во втором случае, т.е после модернизации контрольного оборудования: $\rho = 1.5/1.6 = 0.9375 < 1.$

В системе будут проверяться все 100% изделий, поэтому прежние параметры (относительная и абсолютная пропускная способность) теперь теряют смысл. Интерес представляет средняя длина очереди, т.е. среднее число изделий, ожидающих, пока контролер освободится и возьмет их на проверку. Для ее определения используем формулу для $L_{ср}$ = 0.8789/(1 - 0.9375) = 14.06. Среднее число изделий, находящихся в системе, рассчитывается по формуле для $\omega_{ср}$ = =0.9375/0.0625 = 15. Среднее округленное время ожидания в системе контроля определяется по формуле для $T_{ож}$: $T_{ож} = 0.9375^2/(1.5 \times 0.0625) = 10$ мин.

Время ожидания находится в допустимых пределах, и систему технического контроля с ожиданием можно считать вполне приемлемым вариантом системы технического контроля, обеспечивающей 100%-ную проверку всех блоков. Вероятность образования очереди при заданных выше интенсивностях потока изделий и производительности контроля p_k= 0,88.

Задача № 2.

Фирма организует у себя телефонную связь. Аналитически известны интенсивность потока заявок λ. и интенсивность потока обслуживании μ. Необходимо обосновать оптимальное количество каналов обслуживания. Очевидно, что чем больше количество каналов, тем вероятность обслуживания (вероятность связи) выше, но при этом может снизиться эффективность работы станции из-за простоев в этих каналах и лишних затрат на обслуживание.

Решение.

Данную задачу можно описать n-канальной системой с отказами.

Граф состояний такой системы показан на рис. 10.

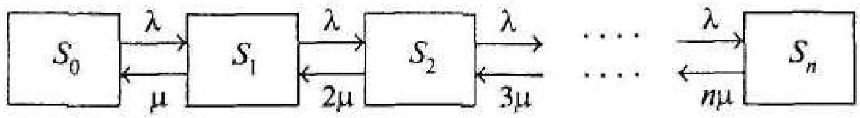


Рис. 10.

Состояния системы:

S_o — все каналы свободны;

S_1 , — занят один канал, остальные свободны;

S_2 — заняты два канала, остальные свободны;

S_n — заняты все *n* каналов.

Уравнения Колмогорова для такой системы:

$$\left.\begin{aligned}
&\lambda\, p_0 = \mu\, p_1;\\
&(\lambda + \mu)\, p_1 = \lambda\, p_0 + 2\mu\, p_2;\\
&(\lambda + 2\mu)\, p_2 = \lambda\, p_1 + 3\mu\, p_3;\\
&\dots \quad \dots \quad \dots\\
&(\lambda + k\mu)\, p_k = \lambda\, p_{k-1} + (k+1)\mu\, p_{k+1};\\
&\dots \quad \dots \quad \dots\\
&(\lambda + (n-1)\mu)\, p_{n-1} = \lambda\, p_{n-2} + n\,\mu\, p_n;\\
&p_0 + p_1 + p_2 + \dots + p_n = 1.
\end{aligned}\right\}$$

Решая эту систему уравнений, легко можно получить значения p_0, p_1, p_2 и т.д.

Предположим, что на телефонную станцию поступает в среднем 1.5 заявки в минуту, а поток обслуживании имеет интенсивность, равную 0.5 заявки в минуту. Следовательно,

$\lambda / \mu = 3$. Вероятность обслуживания поступившей заявки для

$$Q = 1 - p_n = 1 - \frac{(\lambda/\mu)^n}{n!}\, p_0$$

$$p_0 = \left(1 + \lambda/\mu + \frac{(\lambda/\mu)^2}{2!} + \frac{(\lambda/\mu)^3}{3!} + \dots + \frac{(\lambda/\mu)^n}{n!}\right)^{-}$$

где

$$p_k = \frac{(\lambda/\mu)^k}{k!}\, p_0\,, \quad (k = 1, 2, 3, \dots, n).$$

Среднее число занятых каналов:

$$N = (\lambda/\mu)\,(1 - (\lambda/\mu)^n/n!\; p_0)$$

Для трех каналов ($n = 3$) получим следующие результаты. Вероятность обслуживания заявки Q = 0.65, что составляет 65%. При этом среднее число занятых каналов $N = 1.96$, что составляет 65% от всех трех каналов. Соответственно 35% поступающих в систему заявок получают отказ.

Увеличим число каналов обслуживания до 4. Получим вероятность обслуживания заявки $Q = 0.79$, что составляет

79%. Вероятность отказа уменьшается до 21%. Вместе с тем число занятых каналов становится равным 2.38, что составляет 60% от всего числа каналов. Мы видим, что при сравнительно небольшом снижении процента занятых каналов (с 65% до 60%) происходит существенное увеличение вероятности обслуживания — с 65 до 79%.

В случае 5 каналов Q — 89%, процент занятых каналов — 53%.

В случае 6 каналов Q = 94%, процент занятых каналов — 47%.

Подведем итоги.

При увеличении каналов с 3 до 4:

— количество занятых каналов снижается на 5 %;

— вероятность обслуживания возрастает на 14 %.

При увеличении каналов с 4 до 5:

— количество занятых каналов снижается на 7%;

— вероятность обслуживания возрастает на 10%.

При увеличении каналов с 5 до 6:

— количество занятых каналов снижается на 6%;

— вероятность обслуживания возрастает на 5%.

Таким образом, в динамике мы видим, что увеличение каналов с 3 до 4 является оптимальным, так как при минимальном снижении числа занятых каналов наблюдается максимальный прирост вероятности обслуживания. Дальнейшее увеличение каналов невыгодно из-за простоев в них.

Задача № 3.

На автозаправочной станции имеется одна колонка и площадка, на которой могут находится одновременно не более *т* автомашин. Если все места на площадке заняты, то очередная машина, прибывшая к станции, не останавливается, а проезжает мимо. Аналитически было выявлено, что на автозаправочную станцию в среднем в минуту прибывает поток машин с интенсивностью λ_1, а поток обслуживания с интенсивностью μ определяется длительностью заправки.

Менеджеров интересуют вероятность отказа в обслуживании и среднее время ожидания в очереди в зависимости от мест в очереди *т.*

Решение.

Данную задачу можно представить в виде одноканальной системы с ограниченной очередью. Число мест в очереди *т.* Если все места заняты, то очередная заявка, поступающая в систему, получает отказ. Граф состояний такой системы показан на рис. 11.

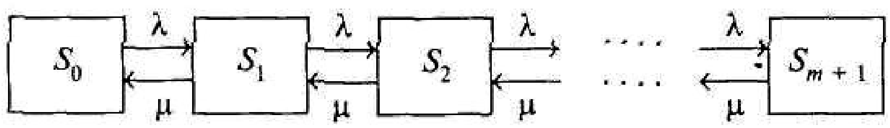


Рис. 11

Состояния системы;

S_0, — канал свободен;

S_1 — канат занят, идет обслуживание, но очереди нет;

S_2 — канал занят, одна заявка стоит в очереди;

S_3 — канал занят, в очереди стоят две заявки;

S_{m+1} — канал занят, в очереди стоят *т* заявок. Уравнения Колмогорова для такой системы:

$$\left.\begin{aligned}
&\lambda p_0 = \mu p_1;\\
&(\lambda + \mu)\, p_1 = \lambda p_0 + \mu p_2;\\
&\dots \quad \dots\\
&(\lambda + \mu)\, p_m = \lambda p_{m-1} + \mu p_{m+1};\\
&p_0 + p_1 + p_2 + \dots + p_m + p_{m+1} = 1.
\end{aligned}\right\}$$

Решая эту систему и вводя = λ/μ, получаем: вероятность свободного канала

$$p_0 = \frac{1}{1 + \rho + \rho^2 + \rho^3 + \dots + \rho^{m+1}} = \frac{1 - \rho}{1 - \rho^{m+2}};$$

$$p_k = \rho^k p_0,$$

где $k = (1, 2, 3, \dots m+1)$

Вероятность отказа p_{m+1}.

Среднее число заявок в очереди:

$$r = \sum_{k=1}^{m} k\, p_{k+1},$$

где p_{k+1} — вероятность того, что в очереди стоят k заявок.

Среднее время ожидания в очереди: r/λ.

Предположим, что на автозаправочную станцию прибывает в минуту в среднем одна машина. Следовательно, λ = 1.

Предположим, что длительность заправки составляет в среднем 2 мин. Следовательно, μ = 1/2. Таким образом, ρ = λ/μ = 2.

Если число мест в очереди $т$ = 3, то вероятность отказа $p_{т+1}$, = 51,6%, а среднее время ожидания в очереди равно 2,1 мин.

Если число мест в очереди $т$ = 6, то вероятность отказа $p_{т+1}$, = 50,2%. а среднее время ожидания в очереди равно 5 мин.

Видно, что если $\rho > 1$, то при больших τ вероятность отказа стабилизируется, становясь равной $(\rho - 1)/\rho$. Чтобы существенно снизить вероятность отказа, необходимо (если нельзя уменьшить ρ) переходить к многоканальным системам.

Задача № 4.

В порту с одним причалом выгружаются прибывающие суда. Аналитически известны интенсивность потока заявок λ и интенсивность потока обслуживании (разгрузка судов) μ. При этом может образоваться очередь.

Менеджеров, организующих работу порта, интересуют вероятности очередей размером k и вероятность отсутствия очереди.

Решение.

Данную задачу можно представить в виде одноканальной системы с неограниченной очередью. Граф состояний такой системы показан ниже.

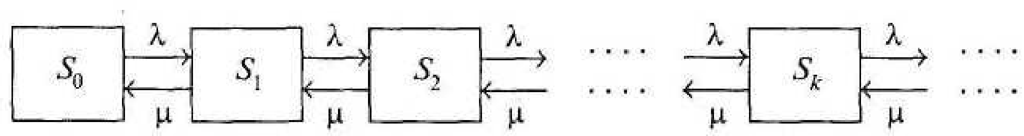


Рис. 12.

Состояния системы:

S_0 — канал свободен (очереди нет);

S_1 — канал занят (идет выгрузка одного судна), но очереди нет;

S_2 — канал занят, в очереди стоит одна заявка;

S_3 — канал занят, в очереди стоят две заявки;

…

S_k — канал занят, в очереди стоят $(k - 1)$ заявок. Эта система характеризуется бесконечным числом дискретных состояний.

Вероятность обслуживания без очереди (состояние S_0):

$$p_0 = 1 - \rho$$

Вероятность очереди из $(k - 1)$ заявок:

Если условие $\rho < 1$ не выполняется, то стационарный режим в рассматриваемой системе не устанавливается: очередь при $t \to \infty$ растет неограниченно.

Задача № 5.

Имеется инструментальный склад, обслуживающий несколько цехов фирмы. Аналитически известны интенсивность потока требований на инструмент λ и интенсивность потока обслуживании μ за смену. Известны также потери в единицу времени: от простоя в очереди — n усл. ед., на содержание кладовщика — m усл. ед.

Менеджеров, организующих производственный процесс, интересует среднее время ожидания обслуживания и среднее время обслуживания при разном количестве кладовщиков s инструментального склада. Также важно найти оптимальное количество кладовщиков с учетом затрат в единицу времени на простой в очереди и на содержание кладовщика.

Решение.

При работе одного кладовщика данную задачу можно представить в виде одноканальной системы обслуживания с неограниченной очередью: $\rho = \lambda/\mu.$

При $\rho > 1$ очередь растет неограниченно.

При $\rho < 1$ имеем следующие показатели.

Вероятность отсутствия очереди:

$$p_0 = 1 - \rho.$$

Вероятность очереди из $(k - 1)$ заявок:

$$p_k = \rho^k (1 - \rho) \text{ или } p_k = \rho^k p_0.$$

Среднее время ожидания $T_c = 1/\mu\ (1/(1-\rho)).$ в системе

$$T_c = T_{ож} + T_{обс}.$$

Среднее время ожидания обслуживания:

$$T_{ож} = 1/\mu\ (1/(1-\rho)).$$

Среднее время обслуживания:

$$T_{обс} = 1/\mu.$$

При работе s кладовщиков задачу можно описать как многоканальную систему с неограниченной очередью.

Если $\rho/s < 1$, то существуют финальные вероятности.

Если $\rho/s \geq 1$, то очередь растет до бесконечности.

При этом ρ может быть больше 1.

Предположим, что условие $(\rho/s) < 1$ выполнено. Тогда вероятность

$$p_0 = \left(1 + \rho/1! + \rho^2/2! + \ldots + \rho^s/s! + \frac{\rho^{s+1}}{s!\,(s-\rho)}\right)^{-1}.$$

Среднее число заявок в очереди:

$$L_{оч} = \frac{\rho^{s+1}\, p_0}{s \times s!\,(1-\rho/s)^2}.$$

Среднее число заявок в системе (с учетом уже обслуживающихся заявок):

$$L_c = L_{оч} + \rho.$$

Среднее время пребывания заявки в очереди:

$$T_{оч} = (1/\lambda) \times L_{оч}.$$

Среднее время пребывания заявки в системе:

$$T_c = (1/\lambda) \times L_c.$$

Предположим, что затраты в единицу времени на простой составляют 7 усл. ед., а на содержание одного кладовщика 5 усл. ед. Тогда получим следующие результаты при разном количестве кладовщиков (полагаем, что λ = 1.6, μ = 0.9, ρ = =1.77).

При s = 2: T_c = 5.11, общие затраты 7x 5.11 + 5x2 = 45.77 усл. ед.

При s = 3: T_c = 1.42, общие затраты 7x 1.42 + 5x3 = 24.94 усл. ед.

При s = 4: T_c = 1.17, общие затраты 7x 1.17 + 5 x 4 = 28.19 усл. ед.

Видно, что с экономической точки зрения выгодно держать на складе трех кладовщиков.