Министерство образования и науки Российской Федерации

Государственное образовательное учреждение
высшего профессионального образования
«Оренбургский государственный университет»

Кафедра медико-биологической техники

А.Д. Стрекаловская, А.В. Рачинских, Т.А. Санеева

ТЕОРИЯ МАССОВОГО ОБСЛУЖИВАНИЯ

Методические указания
к лабораторной работе

Рекомендовано к изданию Редакционно-издательским советом Государственного образовательного учреждения высшего профессионального образования «Оренбургский государственный университет»

Оренбург
ИПК ГОУ ОГУ
2011

УДК 615.47:621.3(07)
ББК 53я7+31.23я7
С84

Рецензент – заведующий отделением медицинской техники ФГУ МНТК «Микрохирургия глаза» имени академика С.Н. Федорова В.И. Канюков

Стрекаловская, А.Д.

С84 Теория массового обслуживания: методические указания к лабораторной работе /А.Д. Стрекаловская, А.В. Рачинских, Т.А. Санеева; Оренбургский гос. ун-т. – Оренбург : ОГУ, 2011. – 28 с.

Методические указания устанавливают объем и содержание лабораторной работы, содержат основные элементы теории массового обслуживания.

Методические указания предназначены для студентов, обучающихся по программе высшего профессионального образования по специальности «Инженерное дело в медико-биологической практике» при изучении дисциплины «Эксплуатация и техническое обслуживание изделий медицинской техники».

УДК 615.47:621.3(07)
ББК 53я7+31.23я7

Содержание

Введение

В настоящее время происходит насыщение лечебно-профилактических учреждений изделиями медицинской техники (ИМТ), позволяющими использовать эффективные методики лечения и диагностики больных. Медицинская промышленность предоставляет врачам большой выбор изделий медицинской техники. Однако более или менее интенсивная эксплуатация ИМТ может привести к выходу их из строя. Для того, чтобы избежать дорогостоящего ремонта этих изделий после выхода их из строя, необходимо регулярно проводить их техническое обслуживание, что намного дешевле.

Только с помощью полноценного технического обслуживания ИМТ возможна минимизация затрат на их нормальную эксплуатацию в течение всего срока службы.

1 Принципы сервисного обслуживания

Карьера в сфере сервисного обслуживания электрических и электронных устройств может быть финансово привлекательной и приносить подлинное удовлетворение от работы. Эксперт обладает уникальным набором знаний в области электронной теории, техники решения проблем и квалификации в выполнении работ. Большинство электронных изделий и приборов содержат такие сходные элементы, как резисторы, конденсаторы, диоды, транзисторы, выводы, разъемы, провода. Понимание причин стандартных поломок этих элементов и способов их тестирования является необходимой предпосылкой для специалиста. В этой главе вы научитесь основам анализа решения проблем, узнаете распространенные неполадки и основные процедуры проверки работоспособности наиболее часто встречающихся электрических и электронных компонентов.

2 Анализ решения проблем

Прежде чем пытаться обслуживать прибор, вы должны сначала разработать концепцию решения проблем и применить ее к поиску неисправностей и ремонту. Первоначальный план действий таков:

1) анализ ситуации;

2) определение причин возникновения проблемы;

3) принятие решения.

Вы должны поступать именно в таком логическом порядке, в противном случае могут возникнуть ошибки, несчастные случаи, потери времени и лишние расходы. Например, многие специалисты по ремонту, обнаружив сгоревший предохранитель, просто заменяют его, вместо того, чтобы сначала определить причину возникновения проблемы. В результате может сгореть и следующий предохранитель.

Поэтому первым шагом в обслуживании устройства является анализ ситуации. Он предполагает критический обзор и всестороннее исследование возникшей проблемы, что позволяет специалисту понять причины, которые не позволяют прибору правильно работать. Это определяется простым осмотром общего состояния устройства.

Начните этот этап, задав вопросы заказчику и проведя наблюдения по следующим пунктам:

1) обсудите дефект с владельцем или пользователем;
2) сравните проблему с другими из вашего прошлого опыта;
3) может быть, неисправности и нет, а имеет место ошибка пользователя;
4) определите различия между текущим состоянием устройства и тем, которое должно быть при правильной работе;
5) оцените ситуацию в целом, отметив симптомы и необходимые изменения.

Определение причин возникновения проблемы вступает в силу, когда наблюдается отклонение от стандартного или желаемого состояния устройства. Примером является неправильно функционирующее или неработающее устройство. Поиск неисправностей представляет собой процесс определения причин проблемы. Первым шагом является организация работы. Начните с подготовки соответствующих схем, спецификаций производителя и руководств по техническому обслуживанию, инструментов и оборудования. Не старайтесь сократить этот этап, бросаясь сразу работать и тратя много времени на исправление устройства, в то время как простое чтение руководства по техническому обслуживанию может способствовать скорейшему решению проблемы. Другими словами, кто провалил этап планирования, тот гарантировал провал на пути устранения неполадок. Когда вы подготовились, выполните следующие операции:

1) опишите проблему;

2) сравните ситуацию с условиями работы устройства до возникновения неисправности;

3) опишите такие различия, как симптомы, шумы, запахи, которые были замечены при возникновении дефекта;

4) сравните: что есть и чего нет. Какие компоненты в порядке, а какие нет, и до какой степени они дефектны;

5) проанализируйте разницу с помощью тестирования, обращая особое внимание на неочевидные и непрямые связи. Например, небольшие изменения допусков элементов или цвета могут указывать на причину неисправности.

Когда вы определили истинную причину возникновения проблемы, то готовы перейти к заключительной фазе, которая называется «принятие решений».

На этом этапе специалист рассматривает различные варианты решения проблемы и выбор наилучшего. Например, если выяснено, что причиной неполадок стал электродвигатель, может быть несколько способов исправления. В зависимости от условий работы всей системы в целом можно починить двигатель или поставить новый той же модели. Третий вариант: выбрать более современную версию двигателя. Принимая решение, вы должны обратить внимание на преимущества и недостатки каждого способа. Планирование действий при аварийной ситуации учитывает будущие изменения всей системы: ожидаемый срок службы, условия работы и внесенные изменения. Например, может быть не совсем разумно ставить новый двигатель, если вся система в скором времени морально устареет и, в любом случае, будет заменена.

Помните о необходимости всегда выполнять все три фазы: ситуационный анализ, определение причин возникновения проблемы (поиск неисправностей) и принятие решения (ремонт). Для того чтобы стать умелым экспертом необходимо понимать важность этой последовательности и не изменять ей.

3 Неисправности схем

Большинство людей хотели бы, чтобы электрические и электронные изделия были гарантированно предохранены от неисправностей, но, к несчастью, это невозможно. Вероятно, большинство поломок - прямо или косвенно - возникают в результате неправильного использования или неудовлетворительного технического обслуживание.

Электрические или электронные неисправности можно классифицировать по основным причинам их возникновения следующим образом:

- тепло;
- влага;
- грязь и загрязнения;
- ненормальное или излишнее перемещение;
- неправильная установка;
- производственные дефекты;
- животные и грызуны.

Когда электронные приборы подвергаются слишком сильному тепловому воздействию, возникают проблемы. Тепло увеличивает сопротивление некоторых элементов схем, что в свою очередь приводит к возрастанию тока. Высокая температура заставляет материалы расширяться, высыхать, трескаться, вздуваться и изнашиваться гораздо быстрее, и, рано или поздно, устройство выйдет из строя.

Влага вызывает больший ток в цепях и может привести к поломке элементов. Вода и другие жидкости вызывает расширение, деформацию, ускоренный износ материалов и аномальный ток (короткие замыкания). Грязь, дым, испарения, абразивные материалы, сажа, жир, масла приводят к тому, что электронные устройства засоряются и покрываются липким налетом, начинают работать в ненормальном режиме и затем выходят из строя.

4 Теория массового обслуживания

Теория массового обслуживания (ТМО) представляет собой прикладную математическую дисциплину, занимающуюся исследованием показателей производительности технических устройств или систем массового обслуживания (СМО), предназначенных для обработки поступающих в них заявок на обслуживания заявок.

Для того чтобы понять необходимость ТМО и те последствия, к которым приводит игнорирование случайностей при расчете показателей обслуживания СМО, рассмотрим простейший пример. Пусть на некоторое обслуживающее устройство или обслуживающий прибор поступает поток заявок. Допустим, путем длительных наблюдений мы установили, что среднее число поступающих на прибор заявок постоянно и равно 6 в час. Спрашивается, какую производительность должен иметь прибор, чтобы успешно справляться с поступающим на него потоком заявок? Сам собой напрашивается ответ: прибор должен обслуживать в среднем 6 заявок в час или каждую заявку за 10 минут. Конечно, осторожный проектировщик всегда сделает небольшой запас, скажем, в 10 % на всякие непредвиденные обстоятельства и предложит производительность прибора, соответствующую обслуживания одной заявки за 9 минут. Дальнейшее увеличение производительности прибора вряд ли целесообразно, поскольку тогда он будет большую долю времени простаивать. Итак, ответ готов: прибор должен обслуживать заявку в среднем за 9 минут. При этом заявки перед прибором не должны накапливаться, а сам прибор в среднем 6 минут каждый час будет простаивать.

Однако на практике весьма быстро было подмечено следующее обстоятельство. Да, прибор действительно был свободен 10 %времени. Но в очень многих случаях перед прибором возникала весьма значительные очереди. В частности, при пуассоновском входящем потоке и экспоненциальном обслуживании при таких исходных данных в среднем перед обслуживающим прибором скапливается очередь из 8 заявок. Поиски причин этого явления выявили и виновника: им оказался именно элемент случайности в поступлении и обслуживании заявок.

Дальнейший ход событий предсказать не трудно. Раз виноваты случайные явления, а случайными явлениями занимается теория вероятности, то необходимо для анализа СМО применять методы этой дисциплины. Таким образом, сформировался еще один раздел теории веброятности - теория массового обслуживания. Родоначальником ТМО считается сотрудник Копенгагенской телефонной компании известный датский ученный А. К. Эрланг, который первым предположил для описания процессов, происходящих в СМО, использовать марковские процессы с дискретным (конечным или счетным) множеством состояний. Это нетрудно понять, если учесть, что основным практическим потребителем результатов ТМО были телефонные сети, а к настоящему времени добавились сети передачи данных, информационно-вычислительные сети и т.д.

Пик своего развития ТМО достигла в 50-70-е годы. Затем интерес к ТМО несколько ослабел. Это было связано с несколькими причинами, например, математической. Здесь нужно отметить, что, с одной стороны, характерной особенностью задач ТМО является необходимость почти для каждой СМО искать собственные методы исследования, а с другой - большой интерес исследователей к ТМО привел к тому, что задачи, допускающие простые решения, особенно в вычислительном плане, уже были решены. Кроме того, у аналитических методов появился серьезный конкурент - имитационное моделирование.

Однако в последнее время снова возродился интерес к задачам ТМО, обусловленный не только новыми проблемами, возникшими в практической жизни и особенно в областях, связанных с разработкой и применением вычислительной техники, но и новыми математическими подходами к их решению. Одним из таких подходов является алгоритмический подход, возникший в связи с широким применением вычислительной техники, в частности, персональных компьютеров в научных исследованиях, и предлагающий получение решений задач ТМО в виде тех или иных вычислительных алгоритмов. Алгоритмический подход, проигрывая традиционным аналитическим методам в наглядности полученных результатов, возможности их использования в задачах оптимизации и т. п., тем не менее, обладает и несомненным преимуществом, которое заключается в его ориентации на создание, в конечном итоге, комплексов и пакетов прикладных программ и таблиц, что в практической жизни часто оценивается гораздо выше даже очень "красивых" формул.

5 Основные элементы ТМО

Многие понятия теории массового обслуживания можно проиллюстрировать на одном важном примере: взлет и посадка самолетов в крупном аэропорту - операция, представляющая интерес для многих людей, пользующихся этим видом транспорта.

Допустим, что аэропорт имеет несколько взлетно-посадочных (*параллельных каналов*). Эти полосы ведут к большему или меньшему числу дорожек, оканчивающихся на аэровокзале (*последовательные каналы*). После того как самолет, прибывший в соответствии с определенным *распределением входящего потока*, приземляется, он присоединяется к очереди самолетов, ожидающих обслуживания (продвижение по дорожке к месту выгрузки). Таким образом, *выходящий поток* одной очереди становится *входящим потоком* для другой. Очередь существует как на земле (взлет самолетов), так и в воздухе (посадка самолетов). Обе эти очереди имеют свое распределение входящего потока. Приземляющиеся самолеты могут прибывать *группами*, при этом члены каждой группы должны кружить над аэропортом и приземлятся по порядку. (Если полоса очень широкая, то нетрудно представить посадку самолетов группами.) *Длительность операций обслуживания* (время приземления или взлета) около минуты. В любом случае имеется некоторое *распределение времени обслуживания*. Если для различных типов самолетов отведены различные взлетно-посадочные полосы, которые могут быть длиннее, например, для реактивных самолетов, то распределение времени обслуживания может меняться от одной полосы к другой.

При выборе самолетов для посадки важно определить соответствующий *показатель эффективности*. Например, если желательно *минимизировать общее время ожидания* пассажиров, то вначале нужно производить посадку самолетов с большим количеством людей.

Здесь же часто производится обслуживание с *приоритетом*, когда разрешается посадка снижающемуся самолету раньше, чем взлет ожидающемуся. Эта система с приоритетом распространяется также на случай аварийной обстановки, когда вследствие крайней необходимости разрешается посадить первым самолет, прибывший позже. Нередко приоритет на посадку дается реактивным самолетам из-за ограниченного запаса топлива.

Иногда порядок обслуживания таков, что прибывающий самолет присоединяется к очереди эшелонированных самолетов, ожидающих посадки, а затем выбор самолета на посадку производится случайным образом (одна из форм обслуживания с приоритетом). Так, например, если самолет находится ближе других к точке, в которой он может выйти из зоны ожидания, то ему будет дана команда на посадку. В промежутке времени между получением приоритета на посадку и командой "посадку разрешаю" самолет выходит из эшелона и направляется к аэродрому. Это время известно как захода на посадку. Время приземления затрачивается на операцию посадки и продолжается до того момента, когда самолет сворачивает с взлетно-посадочной полосы.

Самолет, ожидающий посадки, может, находится в положении, близком к критическому (в это время другие самолеты будут действительно в критическом положении), он может принять решение *присоединится к более* *короткой очереди* в ближайшем аэропорту и приземлится там. Прибывающий самолет может не выстраиваться в эшелон, а уходить в другой аэропорт (*отказ становится в очередь*). В этом случае говорят, что аэропорт "потерял" этот самолет. Случается, что самолет отправляется в соседний аэропорт после того, как, присоединившись к очереди, он прождал больше, чем предполагалось (*по кидание очереди до начала обслуживания*). Можно рассматривать приземляющийся самолет участвующим в *цикле*, если он присоединяется к очереди самолетов, ожидающих взлета, и снова включается к очереди самолетов, ожидающих взлета, и снова включается во входной поток системы. Если приземляющийся самолет имеет информацию о размерах очереди эшелонированных самолетов, ожидающих посадки в соседнем аэропорту, то он может присоединится к этой очереди. Если у него есть информация еще об одном аэропорте, то он может отправится и туда (редкий случай). Это движение туда и обратно при наличии нескольких очередей называется переходом из одной очереди в другую (*возможность выбора очереди*).

Аэропорт может временно закрываться, и прибывший самолет будет вынужден отправится в другой аэропорт, если число эшелонированных самолетов, ожидающих посадки, достигнет заданной величины. Операция обслуживания может быть ускорена путем оборудования специальных гасителей скорости, которые позволяют самолетам приземлятся на главной полосе с большой скоростью.

Основной проблемой пи управлении аэропортом является связь. Если входящий поток как на земле, так и в воздухе велик, то аэропорт должен быстро связываться с самолетами и получать ответ. При организации связи важной проблемой является определение числа операторов и *каналов связи*, необходимых для регулирования различных состояний перегруженности, которые могут возникнуть. В данном случае необходимо выбрать *оптимальное число каналов* для обслуживания требований, поступающих в соответствии с данным распределением. Можно произвести сравнение *стоимость дополнительного канала* со стоимостью возросшего объема обслуживания существующими каналами.

Важной проблемой является наличие соответствующего места для ожидания в очереди. Например, при проектировании аэропорта существенным моментом является наличие наземной рулежной дорожки для самолетов, готовых к влету.

Во многих задачах ТМО для определения необходимого показателя эффективности достаточно знать распределение входящего потока, дисциплину очереди (например, случайный выбор, обслуживание в порядке поступления или с приоритетом) и распределение времени обслуживания. В других задачах нужно иметь дополнительную информацию. Например, в случае отказов в обслуживании нужно определить вероятность того, что поступившее требование получит отказ сразу после прибытия или через некоторое время, т.е. покинет очередь до или после присоединения к ней.

С теоретической точки зрения очередь можно рассматривать как потоки, походящие через систему пунктов обслуживания, соединенных последовательно или параллельно. На поток оказывают влияние различные факторы; они могут замедлять его, приводить к насыщению и т.д.

6 Виды систем массового обслуживания

6.1 Системы массового обслуживания при наличии входного и выходного потоков

В данном разделе рассматриваются СМО, в которых имеется как входной поток, так и поток обслуженных клиентов. Исследуются такие структуры, в которых *параллельно* функционируют c узлов (приборов), так что одновременно могут обслуживаться сразу c клиентов. При этом все обслуживающие приборы с точки зрения быстродействия предполагаются эквивалентными. Схематически такая обслуживающая система изображена на рисунке 1. Заметим, что в любой (произвольно выбранный момент) времени всех находящихся в *системе* клиентов следует разделить на тех, кто находится в *очереди* и, следовательно, ждет, когда его начнут обслуживать, и тех, кто уже *обслуживается.*

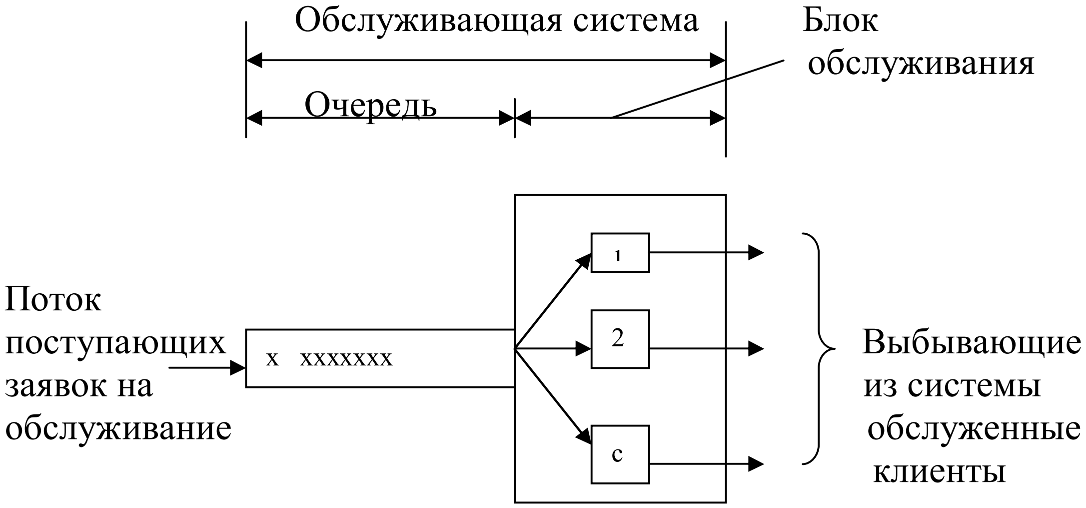

Рисунок 1

Обозначения, которые представляют наиболее подходящими для СМО с параллельно "включенными" приборами, давно уже унифицированы и имеют следующую структуру

$$(a/b/c): (d/e/f), \tag{1}$$

где символы a, b, c, d, e и f ассоциированы с конкретными наиболее существенными элементами модельного представления процессов массового обслуживания и интерпретируются следующим образом:

a - распределение моментов поступлений заявок на обслуживание;

b - распределение времени обслуживания (или выбытий обслуженных клиентов);

c - число параллельно функционирующих узлов обслуживания (c=1, 2, …, ∞);

d - дисциплина очереди;

e - максимальное число допускаемых в *систему* требований (число требований в очереди+число требований, принятых на обслуживание);

f - емкость источника, генерирующего заявки на обслуживание.

Для конкретизации a и b приняты следующие стандартные обозначения:

M - пуассоновское распределение моментов поступления заявок на обслуживание или выбытый из системы обслуживанных клиентов (или экспоненциальное распределение интервалов времени между моментами последовательных поступлений или продолжительностей обслуживания клиентов);

D - фиксированный (детерминированный) интервал времени между моментами последовательных поступлений в систему заявок на обслуживание или детерминированная (фиксированная) продолжительность обслуживания;

E_k - распределение Эрланга или гамма-распределение интервалов времени между моментами последовательных поступлений требований в обслуживающую систему или продолжительностей обслуживания (при этом под k понимается параметр распределения);

GI - распределение произвольного вида моментов поступления в систему заявок на обслуживание (или интервалов времени между последовательными поступлениями требований);

G - распределение произвольного вида моментов выбытия из системы обслуженных клиентов (или продолжительностей обслуживания).

Для иллюстрации рассмотрим структуру (M/D/10)/(GD/N/∞). В соответствии с принятыми обозначениями здесь речь идет о СМО с пуассоновским входным потоком, фиксированным временем обслуживания и десятью параллельно функционирующими узлами обслуживания. Дисциплина очереди не регламентирована, что подчеркивается парой символов GD. Кроме того, независимо от того, сколько требований поступает на вход обслуживающей системы, данная система (очередь+обслуживаемые клиенты) не может вместить более N требований (клиентов), т.е. клиенты, не попавшие в блок ожидания, вынуждены обслуживаться в другом месте.

Наконец, источник, порождающий заявки на обслуживание, имеет неограниченную (бесконечно большую) емкость.

Конечная цель анализа систем и процессов массового обслуживания заключается в разработке критериев (или показателей) эффективности функционирования СМО. В этой связи важно сразу же подчеркнуть одно важное обстоятельство: поскольку процесс массового обслуживания протекает во времени, то нас будет интересовать только стационарный процесс.

При выполнении условий стационарности нас будут интересовать следующие операционные характеристики СМО:

P_n - вероятность того, что в *системе* находится n клиентов (заявок на обслуживание);

L_s - среднее число находящихся в *системе* клиентов (заявок на обслуживание);

L_q - среднее число клиентов *очереди* на обслуживание;

W_s - средняя продолжительность пребывания клиента (заявки на обслуживание) в системе;

W_q - средняя продолжительность пребывания клиента (заявки на обслуживание) в *очереди*.

$$L_s = \sum_{n=0}^{\infty} np_n, \ L_q = \sum_{n=c}^{\infty} (n-c)p_n \qquad (2)$$

по определению.

Между L_s и W_s (как и между L_q и W_q) существует строгая взаимосвязь, так что, зная числовые значения одной из этих величин, можно легко найти значение другой величины. В частности, если частота поступлений в систему заявок на обслуживание равняется λ (интенсивность поступления требований), то мы имеем

$$L_S = \lambda W_S, \ L_q = \lambda W_q. \qquad (3)$$

Приведенные выше соотношения справедливы и при гораздо менее жестких предположениях, не налагающих никаких специальных ограничений ни на распределение моментов последовательных поступлений требований, ни на распределение продолжительностей обслуживания. Однако в тех случаях, когда частота поступлений заявок на обслуживание равняется λ, но не все заявки имеют возможность попасть в обслуживающую систему (например, из-за недостаточно большой вместимости блока ожидания), соотношения (1) необходимо видоизменить путем такого нового определения параметра λ, которое позволило бы учесть только действительно "допускаемые" в систему требования. Тогда, вводя в рассмотрение

$$\lambda_{ЭФФ} = \left(\begin{array}{l} \text{Эффективная частота поступлений,} \\ \text{т.е. количество требований, действи-} \\ \text{тельно допущенных в блок ожидания} \\ \text{обслуживающей системы, в единицу} \\ \text{времени} \end{array} \right)$$

$$L_S = \lambda_{ЭФФ} \cdot W_S$$

$$L_q = \lambda_{ЭФФ} \cdot W_q \qquad (4)$$

будем иметь

$$\lambda_{ЭФФ} = \beta\lambda \;, 0 < \beta < 1 . \qquad (5)$$

В общем случае это означает, что только часть поступающих заявок на обслуживание действительно "проникает" в систему. Но в любом случае можно установить зависимость $\lambda_{ЭФФ}$ от $L_S L_q$ следующим образом. По определению

$$\left\{\begin{array}{c}\text{Средняя продол-}\\ \text{жительность пре-}\\ \text{бываний в } \textit{системе}\end{array}\right\} = \left\{\begin{array}{c}\text{Средняя продолжи-}\\ \text{тельность пребы-}\\ \text{вания требований}\\ \text{в } \textit{очереди}\end{array}\right\} + \left\{\begin{array}{c}\text{Среднее время}\\ \textit{обслуживания}\end{array}\right\}$$

$$W_S = W_q + 1/\mu . \qquad (6)$$

Если средняя скорость обслуживания равняется μ и, следовательно, средняя продолжительность обслуживания равняется 1/μ, то справедливо следующее соотношение:

$$L_S = L_q + \lambda/\mu. \qquad (7)$$

Умножая левую и правую части этого соотношения на λ, получаем

$$\lambda_{ЭФФ} = \mu\left(L_S - L_q\right) \qquad (8)$$

Последнее соотношение остается справедливым и в том случае, если λ заменить на $\lambda_{ЭФФ}$. При этом для $\lambda_{ЭФФ}$ можно записать

$$p_n \to L_S = \sum_{n-0}^{\infty} np_n \to W_S = L_S/\lambda \to W_q = W_S - 1/\mu \to L_q = \lambda W_q . \qquad (9)$$

При анализе всех рассматриваемых ниже моделей основное внимание будет сосредоточено на получении формул для p_n, поскольку, зная p_n, нетрудно определить значение всех основных операционных характеристик интересующего нас процесса массового обслуживания в указанном ниже порядке.

Отметим, что в большинстве случаев при вычислении значений p_n в рамках соответствующей математической модели особые трудности не встречаются. Что же касается распределений продолжительностей ожидания, то их численная оценка может оказаться далеко не простой. Таким образом, в большинстве случаев удобнее вычислять W_S и W_q через L_S и L_q.

Пример - Рассмотрим СМО с одним обслуживающим прибором. Пусть среднее количество требований, поступающих в систему в течение часа, равняется трем(λ), а скорость обслуживания составляет 8 (μ) требований в час. Вероятность p_n того, что в системе окажется n требований, определяется на основе данных, полученных в результате наблюдений за функционированием системы. Допустим, что мы имеем следующие статистические оценки:

n	0	1	2	3	4	5	6	7	8
p_n	0.625	0.234	0.088	0.033	0.012	0.005	0.002	0.001	0

Как мы видим ниже, значения p_n вычисляются с помощью формул, которые приходится специально выводить для каждого конкретного типа моделей массового обслуживания.

На основе приведенных выше исходных данных можно вычислить L_S, W_S, W_q и L_q. Начнем с определения среднего числа требований, находящихся в обслуживающей системе требования:

$$L_S = \sum_{n=0}^{\infty} n \cdot p = 0 \cdot 0.625 + 1 \cdot 0.234 + 2 \cdot 0.088 + 3 \cdot 0.033 + 4 \cdot 0.012 + 5 \cdot 0.005 + 6 \cdot 0.002 + 7 \cdot 0.001 = 0.6$$

Поскольку λ=3, для средней продолжительности пребывания требования в системе имеем

$$W_S = L_S / \lambda = 0.6 / 3 \approx 0.2ч, \tag{10}$$

$$W_q = W_S - 1/\mu = 0.2 - 1/8 = 0.075ч. \tag{11}$$

Учитывая, что μ=8, получаем оценку средней продолжительности пребывания в очереди откуда следует, что среднее количество находящихся в очереди "клиентов" равняется

$$L_q = \lambda \cdot W_q = 3 \cdot 0.075 = 0.225. \tag{12}$$

Используя в качестве исходных данных, приведенные в предыдущем примере, вычислим:

1) среднее количество находящихся в очереди требований, используя при этом непосредственно известные значения p_n.

По определению

$$L_q = \sum_{n=c}^{\infty} (n-c) p_n \tag{13}$$

Подставляем соответствующие значения

$$L_q = \sum_{n=2}^{\infty} (2-1) \cdot 0.088 + (3-1) \cdot 0.033 + (4-1) \cdot 0.012 + (5-1) \cdot 0.005 + (6-1) \cdot 0.002 + (7-1) \cdot 0.001 = 0.225 \tag{14}$$

2) среднее количество клиентов, которые обслуживаются системой.

По определению среднее количество клиентов, которые обслуживаются системой равно L_S-Lq. Из приведенных выше формулах находим

$$L_S - L_q = \lambda W_S - \lambda W_q = \lambda (W_S - W_q) = \lambda (W_q + 1/\mu - W_q) = \lambda / \mu = 3/8 = 0.375 \tag{15}$$

При увеличении параметра λ будет увеличиваться L_S и L_q, а при увеличении параметра μ будет уменьшаться W_S и W_q.

7 Оптимальная скорость обслуживания

Рассмотрим одноканальную модель массового обслуживания со средней частотой поступления требований, равной λ, и со средней скоростью обслуживания, равной μ. Предполагается. Что скорость обслуживания поддается регулированию; требуется определить ее оптимальное значение на основе надлежащим образом построенной стоимостной модели. Введем следующие обозначения:

C_1 - выражения в стоимостной форме выигрыш за счет увеличения на единицу значения μ в течение единичного интервала времени;

C_2 - "цена" ожидания (т.е. обусловленные вынужденным ожиданием экономические потери) в единицу времени и в расчете на одно требование;

$TC(\mu)$ - стоимостный показатель, определяемый формулой

$$TC\,(\mu)=C_1\mu+ C_2L_S. \tag{16}$$

Следует отметить, что затраты на обслуживание, отнесенные к единице времени, прямо пропорциональны μ, а затем в единицу времени, обусловленные пребыванием заявок на обслуживание в режиме ожидания, равняются среднему значению числа требований, находящихся в СМО, умноженному на "цену" ожидания, определенную в расчете на одно требование и отнесенную к единице времени.

Поскольку μ является величиной непрерывной, ее оптимальное значение может быть получено путем приравнивания к нулю первой производной $TC\,(\mu)$ поμ. Например, для частного случая (M/M/1):(GD/∞/∞) - процесса

$$TC(\mu) = C_1\mu + C_2\lambda/(\mu-\lambda), \tag{17}$$

и, следовательно, для оптимального значения μ имеем

$$\mu = \lambda + \sqrt{C_2 \lambda / C_1}. \tag{18}$$

В ситуации, когда в блоке ожидания обслуживающей системы может находиться не более N клиентов, т.е. если имеет место (M/M/1):(GD/N/∞) процесс, стоимостную модель можно видоизменить, с тем чтобы за счет увеличения значения N уменьшить число клиентов, которых СМО может потерять. В данном случае величина N рассматривается как управляющая переменная, оптимальное значение которой (вместе с μ) определяется путем минимизации

$$TC(\mu, N) = C_1 \mu + C_2 L_S + C_3 N + C_4 \lambda p_N, \tag{19}$$

где C_3 - "стоимость" увеличения (на единицу времени) вместимости блока ожидания обслуживающей системы;

C_4 - экономические потери, связанные с невозможностью включить в блок ожидания системы еще одного нуждающегося в обслуживании клиента;

λp_N - есть число клиентов, потерянных системой в единицу времени.

Пример - Вычислительный центр коллективного пользования располагает большой электронно-вычислительной машиной (ЭВМ) с разветвленной системой считывающих и быстропечатающих устройств. Один из пользователей хочет определить оптимальную скорость (число перфокарт в минуту) считывающего устройства. Потребности в работе с ЭВМ возникают у пользователей случайно, так что входной поток заявок на обслуживание электронно-вычислительными средствами характеризуется пуассоновским законом распределения вероятностей; отметим, что средняя пропускная способность вычислительного центра составляет 50 программ в течение 8-ми часового рабочего дня. Пусть средний размер программы таков, что она уменьшается на одной тысяче перфокарт. Опыт показывает, что распределение продолжительностей считывания записанных на перфокартах программ является экспоненциальным. По оценкам клиентуры просрочка в выполнении заявленной потребностей в работе с ЭВМ на один день обходится в 10 долл. Вычислительный центр при планировании своей месячной пропускной способности исходит из того, что плата за увеличение пропускной способности считывающего устройства на сотню перфокарт в минуту составляет 100 долл.

Нетрудно убедиться, что скорость считывания, равная 100 перфокартам в минуту, эквивалентна 100·8·60=48000 перфокарт в день, или 48000:1000=48 программно-вычислительных реализаций в день. Полагая, что количество рабочих дней в одном месяце равняется 22, устанавливаемое на месяц указанное выше повышение быстродействия считывающего устройства обходится клиентуре 100 долл. : 22=4,55 долл. в день. Поскольку именно такое повышение быстродействия считывающего устройства позволяет ежедневно реализовать на ЭВМ на 48 процедур больше. Плата за реализацию одной дополнительной процедуры равняется 4,55 долл. : 48=0,0948 долл. Используя введенные выше обозначения, имеем C_1 =0,0948 долл. Поскольку C_2=10 долл. За одну процедуру в день, а λ=50 процедур день, получаем оптимальное значение μ: $\mu = 50 + \sqrt{(10 \cdot 50) \div 0{,}0948} = 123$ процедуры в день.

Переводя этот показатель в количество перфокарт в 1 мин, находим, что оптимальное значение пропускной способности считывающего устройства равняется (123·1000) : (8·60)=256 перфокарт в 1 мин.

8 Оптимальное число обслуживающих приборов

Рассмотрим мультиканальную модель. Стоимостная модель массового обслуживания в данном случае должна быть ориентирована на определение оптимального числа обслуживающих приборов, которое мы обозначили выше через с. предполагается, что значения λ и μ фиксированы. Интегральная стоимость показателей задается формулой:

$$TC(c) = cC_i + C_2 L_S(c), \tag{20}$$

где C_1 - отнесены к единице времени затраты на обеспечение функционирования одного дополнительного обслуживающего прибора;

$L_S(c)$- среднее число находящихся в обслуживающей системе требований. Оптимальное значение с находим из условий

$$TC(c-1) \geq TC(c) u TC(c+1) \geq TC(c), \tag{21}$$

что эквивалентно неравенству

$$L_S(c) - L_S(c+1) \leq C_1 / C_2 \leq L_S(c-1) - L_S(c). \tag{22}$$

Величина C_1/C_2 теперь является указателем того, где должен начинаться поиск оптимального значения с.

Пример - На складе запасных частей производится замена вышедших из строя механических узлов новыми. Заявки на замену поступают в соответствии с пуассоновским распределением вероятностей со средним значением количества заявок, равным λ=17,5. Каждый работник данного склада способен удовлетворить в среднем μ=10 заявок в час. Затраты, ассоциированные с добавлением к штату обслуживающих одного человека, оцениваются в 6 долл. в час. Произведенные потери из-за простоя станка в период замены тех или иных вышедших из строя узлов и (или) деталей составляет 30 долл. в час. Сколько человек должен включать штат обслуживающего персонала на складе запасных частей?

Определение оптимального значения с приведены в таблице 1.

Таблица 1 – Оптимальные значения

C	$L_S(c)$	$L_S(c-1)$- $L_S(c)$
1	2	3
1	∞	-
2	7.467	∞
3	2.217	5.25

Продолжение таблицы 1

1	2	3
4	1.842	0.375
		$\leftarrow C_1/C_2=0.2$
5	1.769	0.973
6	1.754	0.015
7	1.75	0.004

Поскольку $C_1/C_2=6/30=0{,}2$ имеем

$$L_S(4)-L_S(5)=0.073<0.2<0.375=L_S(3)-L_S(4). \tag{23}$$

Следовательно, оптимальное количество работников равняется 4.

Если $C_1=10$ долл., и $C_2=20$ долл., то оптимальное количество работников равно:

$$L_S(3)-L_S(4)=0.375<0.5<5.25=L_S(2)-L_S(3) \tag{24}$$

9 Моделирование с учетом предпочтительности уровня обслуживания

К моделям, в которых осуществляется учет предпочтительного уровня обслуживания, переходят из-за трудностей получения числовых значений стоимостных показателей (параметров) процесса массового обслуживания; при этом весь анализ производится на основе более примитивных оценок операционных характеристик, исследуемых СМО. При использовании таких моделей в ходе поиска "оптимальных" значений основных параметров проектируемой системы обращаются непосредственно к ее операционным характеристикам. При этом "оптимальность" связывают с возможностью обслуживающей системы удовлетворить некоторый желательный с точки зрения, принимающего решение, уровень активности системы. Эти желательные уровни определяются путем оценок верхних предельных значений тех конкурирующих экономических показателей, между которыми лицо, принимающее управляющее решение, хочет установить "баланс".

В мультиканальной модели задача заключается в определении оптимального значения числа обслуживающих приборов с с учетом того, что "конкурирующими" являются следующие показатели: средняя продолжительность ожидания W_S и доля времени (X), в течение которого обслуживающий прибор вынужденно бездействует. Эти показатели и определяют потенциальный характер процесса массового обслуживания. Обозначим верхние предельные значения W_S и X через α и β соответственно. Определим число обслуживающих приборов так, чтобы $W_S \le \alpha u X \le \beta.$ Выражение для X имеет вид

$$X = 100/c\sum_{n=0}^{c}(c-n)p_n = 100(1-\rho/c). \tag{25}$$

Решение задачи может быть найдено элементарным способом, если построить графики функций W_S=W_S(C) и X=X(C). Указав графически уровни α и β, можно сразу же выявить приемлемый диапазон значений C, т.е. такой диапазон значений C, для которого выполняются оба указанных выше условия.

Пример - Вернувшись к условию предыдущего примера, предположим, что цель заключается в определении такого числа работников на складе запасных частей, при котором среднее время ожидания от момента подачи заявки до момента получения требуемой запасной части не превышало бы 20 минут, одновременно потребуем, чтобы доля времени, в течение которого обслуживающий персонал на складе вынужден "простаивать", не превышала бы 15 %.

В таблице 2 приведены значения W_S и X для различных C. С увеличением C величина W_S уменьшается.

Таблица 2

C	1	2	3	4	5	6	7	8
W_S, мин	∞	25,6	7,6	6,3	6,1	6,0	6,0	6,0
X, %	0	12,5	41,7	56,3	65,0	70,8	75,0	78,0

Из таблицы 2 видно, что при переходе от C=2 к C=3 происходит резкое уменьшение W_S. Любое последующее увеличение значения C приводит к незначительному изменению W_S. Обратившись теперь к X, мы видим, что переход от C=2 к C=3 более чем в три раза увеличивает долю времени, в течение которого работники вынуждены простаивать. Следовательно, выбор между вариантом, когда C=2 и вариантом, когда C=3, должен осуществляться с учетом того, насколько значимо снижение простоя каждого станка из-за необходимости заменить в нем определенные детали с 25, 6 до 7,6 мин; при этом следует, естественно, помнить, что уменьшение W_S с 26,5 до 7,6 мин. влечет за собой увеличение средней доли вынужденных простоев обслуживающего персонала на складе запасных частей с 12,5 % до 41,7 %.

Список использованных источников

1 Неразрушающий контроль и диагностика: справочник / под ред. В.В.Клюева. – М. : Машиностроение, 2005. – 490 с.

2 Неразрушающий контроль металлов и изделий: справочник / под ред. Г.С. Самойловича. – М. : Машиностроение, 2004. – 456 с.

3 Политехнический словарь / гл. ред. И.И. Артоболевский. – М. : Советская энциклопедия, 2008. – 607 с.

4 РД 03-606–03. Инструкция по визуальному и измерительному контролю. – М. : Из-во государственного унитарного предприятия «Научно-технический центр по безопасности в промышленности Госгортехнадзора России», 2003. – 101 с.

5 Воронкин, Ю.Н. Методы профилактики и ремонта промышленного оборудования / Ю.Н. Воронкин, Н.В. Поздняков. – М. : Академия, 2002. – 240 с.

6 Кормильцин, Г.С. Основы монтажа и ремонта технологического оборудования / Г.С. Кормильцин, О.О. Иванов. – Тамбов: Изд-во ТГТУ, 2001. – 87 с.