ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ КОНТРОЛЯ КАЧЕСТВА ФУНКЦИОНИРОВАНИЯ СИСТЕМ ОБСЛУЖИВАНИЯ В ТОРГОВЛЕ

Баканов М. И., д.э.н., профессор,
Степанов В. Г., экономист

Московский государственный университет коммерции

В условиях рыночных отношений, которые сложились в нашей стране за последние годы, процесс торгового обслуживания покупателей становится важнейшим объектом контроля и экономического анализа на предприятиях торговли.

Одним из основных показателей, характеризующих процесс обслуживания покупателей, является уровень качества торгового обслуживания. Данный показатель является интегральным, включающим ряд частных показателей, таких как культура обслуживания покупателей, скорость торгового обслуживания, стабильность товарного ассортимента, спектр услуг, предоставляемых покупателям и т. д.

Фактически уровень качества торгового обслуживания является показателем качества функционирования системы обслуживания на предприятии торговли. Под *системой обслуживания в торговле (СОТ)* будем понимать совокупность реализованных на предприятии торговли методов и средств различной природы, обеспечивающих удовлетворение потребностей покупателей в товарах и услугах приемлемого качества и за приемлемое время.

Существующая на предприятии торговли СОТ определяется организацией торгово-технологического процесса, системами управления и маркетинга. Поэтому уровень качества торгового обслуживания будет зависеть от применяемой на предприятии технологии продажи товаров, их качества, широты и глубины товарного ассортимента, стабильности поставок товаров, профессиональной подготовки торгового персонала и уровня его мотивации, наличия и размещения торгово-технологического оборудования в торговом зале и складских помещениях, а также от стратегических целей и тактических задач, решаемых руководством предприятия на рынке.

Очевидно, что уровень качества торгового обслуживания является важным фактором конкурентоспособности предприятия торговли в условиях рыночных отношений. Сегодня потребитель при выборе места покупок ориентируется не только на цены предлагаемых ему товаров, но и на качество его обслуживания. В случае же жесткой ценовой конкуренции на рынке высокий уровень качества торгового обслуживания часто становится главным аргументом в пользу конкретного предприятия. Поэтому в последние годы руководители предприятий торговли стали уделять контролю качества функционирования систем обслуживания повышенное внимание.

Однако контроль и анализ качества торгового обслуживания на практике связан с рядом сложностей. Само понятие "качество обслуживания" носит достаточно субъективный характер. Каждый руководитель определяет приемлемый для него уровень, учитывая ряд таких факторов, как местоположение предприятия, товарный ассортимент, уровень спроса на предлагаемые товары, контингент покупателей и степень их требовательности к качеству обслуживания, имеющиеся организационные и финансовые возможности. Большое значение имеют уровень культуры и ценностные установки самого руководителя, а также соответствующие традиции, сложившиеся на предприятии и в обществе. Поэтому сложно выработать какие-то единые критерии для оценки уровня качества торгового обслуживания. Кроме того, показатель качества торгового обслуживания плохо поддается формализации, так как включает показатели, большинство из которых являются атрибутивными, и поэтому их количественная оценка в значительной степени затруднена.

С учетом вышесказанного можно предложить эффективный метод контроля качества функционирования СОТ и объективной оценки уровня качества обслуживания покупателей на основе такого его важнейшего показателя, как *скорость торгового обслуживания*. Скорость торгового обслуживания определяется средним числом обслуженных покупателей в единицу времени в данной торговой точке или на данном рабочем месте.

Скорость обслуживания имеет существенное социальное значение, оказывая самое непосредственное влияние как на покупателей, так и на персонал и руководство предприятия торговли. Данный показатель влияет и на эффективность эксплуатации и обслуживания технических средств, используемых в торгово-технологическом процессе. Таким образом, сам по себе он является существенным фактором конкурентоспособности предприятия торговли (см. рис. 1).

Скорость торгового обслуживания определяется не только численностью персонала, занимающегося обслуживанием покупателей и интенсивностью его работы, но и организацией торгово-технологического процесса, организацией труда и мотивацией работников предприятия торговли. Она тесно взаимосвязана с другими показателями качества торгового обслуживания, такими как культура обслуживания, стабильность товарного ассортимента, спектр услуг, оказываемых покупателям и т.д. Поэтому в процессе исследования скорости обслуживания подвергаются анализу все факторы, воздействующие на процесс торгового обслуживания покупателей. Таким образом, осуществляется комплексный анализ качества функционирования СОТ, реализуется системный подход к решению задачи поддержания приемлемого уровня качества торгового обслуживания покупателей.

Важно и то, что данный показатель лучше других поддается формализации. Для анализа скорости торгового обслуживания может применяться значительный банк экономико-математических методов, которые позволяют оценить эффективность работы системы обслуживания на предприятии, а также рассчитать оптимальные показатели функционирования СОТ, обеспечивающие необходимую скорость и качество торгового обслуживания покупателей.

Все вышесказанное позволяет сделать вывод, что эффективный контроль качества работы СОТ со стороны руководства предприятия торговли может обеспечиваться путем анализа скорости торгового обслуживания покупателей. При этом анализ должен быть системным, т. е. осуществляться с учетом всех основных факторов, влияющих на процесс торгового обслуживания.

Анализ качества функционирования систем торгового обслуживания по своему содержанию относится к управленческому анализу хозяйственной деятельности предприятий торговли. По отношению к процессу управления он сочетает в себе характеристики перспективного, оперативного и ретроспективного анализа. Однако главной задачей анализа является постоянный контроль работы обслуживающей системы с целью обеспечения заданных показателей качества торгового обслуживания. Следовательно, необходим оперативный анализ качества функционирования СОТ на предприятии торговли. Если же в результате проведения оперативных исследований оказывается, что уровень качества торгового обслуживания ниже установленного по набору показателей, определенных на данном предприятии, то возникает необходимость в проведении более углубленного системного анализа работы системы обслуживания покупателей.

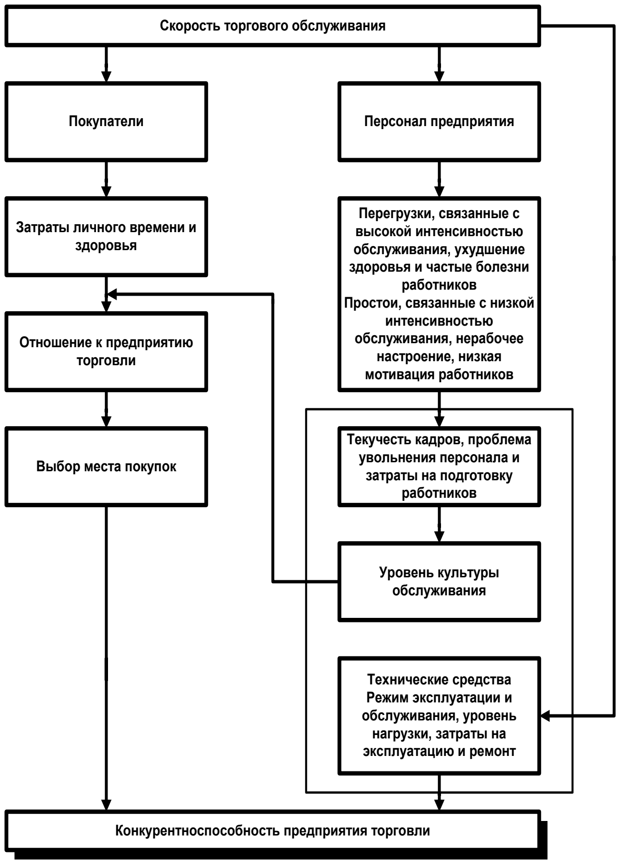

Рис. 1. Влияние скорости обслуживания на конкурентоспособность предприятия

Оперативность предварительного анализа - важнейшее требование, которое накладывается на все виды его обеспечения: организационно-экономическое, информационное, математическое, программное и техническое. В первую же очередь это требование касается информационного обеспечения анализа. Не секрет, что самые эффективные методы анализа часто не находят практического применения именно в связи с высокими требованиями, предъявляемыми к их информационному обеспечению. Оперативность анализа достигается, в первую очередь, выбором соответствующей информационной технологии, которая и определяет, в конечном итоге, математические, программные и технические средства для проведения анализа, а также необходимые меры организационного и экономического характера.

Источниками информации для проведения анализа качества работы СОТ являются данные наблюдений (сплошных и моментных) за работой СОТ, опросов торгового и обслуживающего персонала предприятия, а также первичные документы и данные бухгалтерского учета, которые в процессе анализа подвергаются соответствующей статистической обработке. Таким образом, информационное обеспечение анализа качества функционирования СОТ формируется на основе всех видов учетной экономической информации: оперативной, бухгалтерской и статистической.

Представим процедуру проведения оперативного анализа качества функционирования СОТ, которая, благодаря эффективной информационной технологии, может с успехом применяться на различных предприятиях торговли вне зависимости от их размеров, сложности организации торгово-технологического процесса и системы управления.

В первую очередь, представим математическую модель СОТ, на основе которой осуществляется экономико-математический анализ процесса торгового обслуживания.

СОТ строится из двух элементов: обслуживающей системы и входящего в нее потока требований на обслуживание. Под *обслуживанием* понимается совокупность определенных действий (операций), которые необходимо выполнить согласно поступающим требованиям. *Требованием (заявкой)* называется потребность в обслуживании, исходящая от объекта, поступающего в систему. Объекты могут быть любой природы, и с математической точки зрения для проведения анализа совершенно неважно, является ли объектом человек, техническое устройство или, например, электрический сигнал. Поэтому под требованием обычно понимают как исходящую от объекта потребность в обслуживании, так и сам объект. Последовательность объектов-требований, поступающих в систему с целью их обслуживания, образует *входящий поток требований*. Обслуживание требований в СОТ осуществляется средствами, которые так же, как и объекты, могут иметь различную природу и называются *каналами обслуживания*. Совокупность каналов обслуживания образует *обслуживающую систему*.

Основным показателем входящего потока требований является его *интенсивность* λ - среднее число требований, входящих в систему за единицу времени. Обратной величиной является *среднее время* $t_{тр}$ между последовательными моментами поступления требований в обслуживающую систему. Если требования поступают в систему через строго определенные промежутки времени, то поток называется *регулярным*. Однако в большинстве случаев поток требований является *нерегулярным*, т. е. время между последовательными моментами поступления требований в систему является случайной величиной. Следовательно, важной характеристикой входящего потока является закон распределения времени между моментами поступления требований в систему. Другой характеристикой потока, которая может учитываться в процессе анализа, является степень флуктуации интенсивности входящего потока требований в течение рассматриваемого периода времени. Количественно данная характеристика выражается с помощью коэффициента вариации.

Основным показателем работы каналов обслуживания является *интенсивность* μ или *среднее время обслуживания* $t_{обсл}$ ими поступающих требований. Так как время обслуживания в большинстве случаев является случайной величиной, то важной характеристикой работы канала обслуживания является также закон распределения времени обслуживания требований.

Обслуживающие системы различаются технологией обслуживания требований и, соответственно, делятся на типы.

Все необходимые операции по обслуживанию требований могут выполняться сразу или могут быть разбиты на последовательные этапы или фазы обслуживания. Соответственно, обслуживающие системы делятся на *однофазные* и *многофазные*. При этом на каждой фазе обслуживания может быть задействовано различное число каналов. Очевидно, что любую многофазную систему можно представить в виде совокупности однофазных.

Если однофазная система включает только один канал обслуживания, то она называется *одноканальной*, если же система включает несколько каналов, - то *многоканальной*. При этом возможны следующие варианты обслуживания: один канал может обслуживать сразу несколько требований; каждое требование обслуживается одновременно всеми каналами; каждое требование обслуживается отдельным каналом.

Если каналы обслуживания однофазной СОТ обладают одинаковыми характеристиками, то система называется *однородной*, если же каналы различаются по своим характеристикам, то система называется *неоднородной*.

В процессе обслуживания требований в системе может допускаться или не допускаться образование очереди. Соответственно, различают системы обслуживания, функционирующие в режиме *с очередями* и *с отказами (потерями требований)*. При работе системы в режиме с очередями могут существовать ограничения на длину очереди, время ожидания обслуживания в очереди, или же таких ограничений может не быть (системы с неограниченным ожиданием).

В зависимости от порядка обслуживания требований различают системы *без приоритета* и *с приоритетом*. В системах без приоритета обслуживание требований может осуществляться следующими способами: по принципу *FIFO ("первый пришел - первый обслужен")*; по принципу *LIFO ("последний пришел - первый обслужен")*; по принципу *случайного отбора*. В системах с приоритетами требования получают приоритеты согласно принятым в СОТ правилам. При этом обслуживание может осуществляться по принципу абсолютного или относительного приоритета. В первом случае обслуживание очередного требования прерывается, если в систему поступило требование с более высоким приоритетом. Во втором случае каждое требование обслуживается до конца.

Из вышесказанного следует, что обслуживающая система характеризуется следующими показателями: число фаз обслуживания, число каналов обслуживания на каждой фазе, индивидуальные характеристики каналов, режим и порядок обслуживания требований. Важной интегральной характеристикой работы фазы обслуживания системы является объем обслуживания $x(\Delta t)$, который определяется как суммарная интенсивность обслуживания требований всеми каналами данной фазы за период времени Δt (единицу времени).

Требования, обслуженные системой, образуют *выходящий поток*. Показатели выходящего потока, очевидно, являются производными и определяются характеристиками входящего потока требований и обслуживающей системы.

СОТ можно классифицировать и в зависимости от способа формирования потока требований. В случае, если после обслуживания требования покидают систему, СОТ называют *системами с разомкнутым потоком требований*. Если же уже обслуженные требования полностью или частично формируют входящий поток требований, то имеет место *система с замкнутым (циклическим) потоком требований*.

Характеристики СОТ в значительной степени зависят от природы объектов, входящих в систему для их обслуживания, а также каналов обслуживания. В общем случае на основе данного признака современные системы обслуживания можно разделить на технические и социально-технические. В технических системах массового обслуживания (СМО) и требования, и каналы обслуживания представляют собой устройства, механизмы, сигналы и т. д., т. е. имеют техническую природу. Такие СМО обычно используются в автоматических системах управления. Человек в таких системах не принимает участия в процессе обслуживания. В социально-технических системах ведущую роль в процессе обслуживания играет человек. Такие системы обслуживания используются в автоматизированных и традиционных (неавтоматизированных) системах управления.

Очевидно, что технические системы нехарактерны для сферы торговли. Хотя в современных условия на предприятиях торговли широко используются самые разнообразные технические средства, человек является главным объектом и субъектом торгового обслуживания. Поэтому СОТ являются социально-техническими системами. Однако при этом возможны случаи, когда обслуживающая система является чисто технической, например, торговый автомат или лифт в торговом зале.

Социальный характер процесса обслуживания оказывает существенное влияние на характеристики работы СОТ. В процессе анализа качества работы СОТ необходимо учитывать ряд факторов, характерных в основном для социально-технических систем обслуживания. Во-первых, это сильная флуктуация интенсивности входящего потока требований в течение небольших промежутков времени, которая объясняется колебаниями спроса на отдельные товары, изменениями товарного ассортимента, предпочтениями покупателей приобретать товары в определенные дни, часы и т.д. Во-вторых, в процессе торгового обслуживания, кроме массовых операций, присутствуют также мелкосерийные и единичные операции, которые оказывают существенное влияние на качество работы СОТ. В-третьих, в социально-технических системах показатели обслуживания в значительной степени зависят от показателей входящего потока требований. В частности, в процессе анализа необходимо учитывать, что интенсивность обслуживания покупателей продавцами обязательно будет зависеть от интенсивности потока покупателей, входящих в торговый зал. В этом смысле СОТ можно назвать *адаптивными* системами обслуживания. В-четвертых, необходимо помнить, что все мероприятия, которые будут осуществляться по результатам проведенного анализа работы СОТ, самым непосредственным образом затронут интересы людей, участвующих в процессе обслуживания.

Оперативный экономико-математический анализ качества функционирования СОТ основан на следующих предположениях:

1) С учетом специфики торговой деятельности в качестве основной модели выбирается разомкнутая СОТ, включающая *n* – канальную однофазную однородную обслуживающую систему, которая работает в режиме с очередями и ограничением на длину очереди, а интенсивность входящего в систему потока требований представляет собой случайную величину с произвольным законом распределения (см. рис. 2). На основе данной модели можно построить СОТ любой сложности и с любыми характеристиками: с ограничением на время ожидания обслуживания в очереди, с неограниченным ожиданием в очереди и с отказами, неоднородную, многофазную и замкнутую.

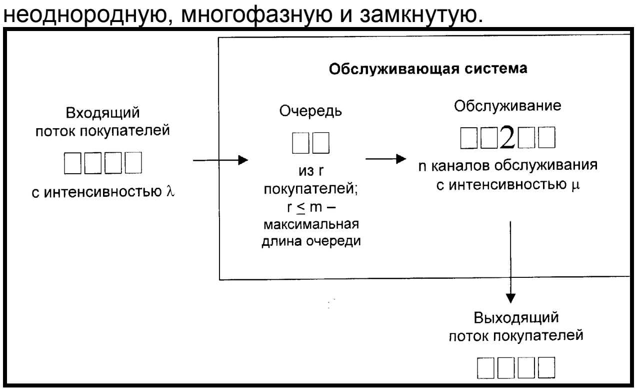

Рис. 2. Основная модель СОТ

2) Процесс обслуживания представляет собой марковскую цепь - случайный процесс с дискретным временем и дискретным конечным множеством состояний системы $S_0, S_1, ..., S_L$. Состояние системы S_k определяется числом требований $k = 0,\dots,L$ $(L = n + m)$, находящихся в системе в рассматриваемый момент времени. Переходы системы из одного состояния в другое происходят через равные промежутки времени Δt - шаги случайного процесса (см. рис. 3).

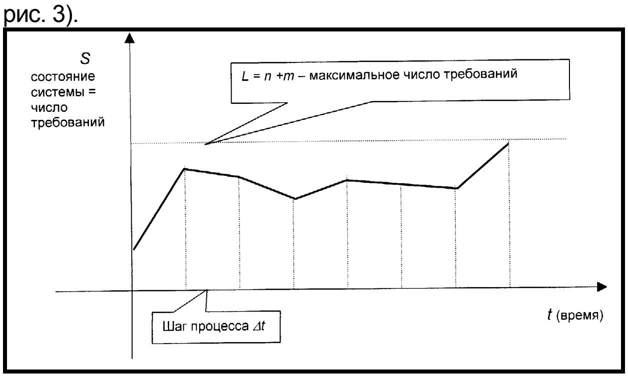

Рис. 3. Представление состояния системы с точки зрения числа находящихся в ней требований

3) Случайный процесс изменения состояния СОТ является однородным по времени, т. е. вероятность перехода системы из одного состояния в другое за шаг процесса Δt не зависит от времени (номера шага) и описывается с помощью матрицы перехода системы (матрицы переходных вероятностей):

$$P = \begin{pmatrix} p_{00}\, p_{01} \cdots p_{0L} \\ p_{10}\, p_{11} \cdots p_{1L} \\ \dots\dots\dots \quad \dots\dots \\ p_{L0}\, p_{L2} \cdots p_{LL} \end{pmatrix},$$

где p_{ij} - вероятность того, что система за время Δt перейдет из состояния S_i в состояние S_j. При этом должны выполняться следующие равенства:

$$\sum_{j=0}^{L} p_{ij} = 1, \qquad i = \overline{0, L}.$$

Учитывая вышесказанное, рассмотрим процесс проведения оперативного анализа, который включает ряд этапов.

1. ОПЕРАТИВНЫЙ АНАЛИЗ КАЧЕСТВА ФУНКЦИОНИРОВАНИЯ СОТ

1.1. Исследование качества работы системы путем опроса персонала

Опрос торгового и обслуживающего персонала является необходимым этапом в процессе анализа социально-технических систем. Опрос осуществляется в форме коллективной и (или) индивидуальной беседы с обязательным анкетированием работников с целью формализации получаемой информации о качестве работы СОТ. В ходе опроса выявляются "узкие места" в организации торгово-технологического процесса, в распределении функций между торговыми работниками и их мотивации, недостатки в управлении текущими задачами в процессе торгового обслуживания.

Эффективным способом получения данных о распределении нагрузки на торговый персонал и технические средства в течение различных периодов времени обслуживания покупателей является применение такого метода шкальных оценок, как *семантический дифференциал*. На рис. 4 приведен пример семантического дифференциала для оценки распределения нагрузки на работника в течение рабочей недели.

Фактически в результате применения данного метода шкальных оценок можно получить карту распределения нагрузки на персонал и (или) технические средства в течение различных периодов времени обслуживания. Для большей формализации получаемых результатов длину шкалы семантического дифференциала необходимо выбирать таким образом, чтобы ее легко можно было привязать к какой-либо шкале баллов. Например, для количественной оценки нагрузки можно воспользоваться шкалой баллов от 0 до 10.

Использование семантического дифференциала позволяет достаточно быстро отрегулировать режим работы торгового и обслуживающего персонала, технических средств, используемых в торгово-технологическом процессе. Кроме того, полученные результаты дают возможность определить периоды времени, в течение которых необходимо осуществлять наблюдения за работой СОТ. В результате опроса персонала можно установить примерные параметры функционирования СОТ: максимальную, минимальную и среднюю интенсивность потока покупателей, входящих в систему в различные периоды времени, среднюю длину очереди у рабочего места, среднее и максимальное время обслуживания покупателей и т.д. Однако более точные данные о качестве работы СОТ можно получить только в ходе наблюдений за функционированием системы на следующих этапах.

1.2. Определение времени и интенсивности обслуживания требований путем проведения сплошных наблюдений за работой системы

На втором этапе осуществляется наблюдение за работой каналов обслуживания с целью определения основных параметров обслуживания. В результате сплошных наблюдений за работой каналов строится ряд распределения времени обслуживания покупателей, определяется среднее время обслуживания требований и интенсивность обслуживания (см. табл. 1).

$$t_{обсл} = \frac{\sum_{j=1}^{R} \tau_j f_j}{\sum_{j=1}^{R} f_j};$$

$$\tau_j = \frac{t_j + t_{j-1}}{2}, \qquad j = 1,\dots, R;$$

$$\mu = \frac{1}{t_{обсл}}.$$

Периоды проведения сплошных наблюдений выбираются на основе данных опроса торгового и обслуживающего персонала, проведенного на предыдущем этапе анализа.

Число наблюдений определяется согласно статистической теории выборочного метода. Например, можно воспользоваться следующей формулой:

$$n = \frac{t^2 \sigma^2}{\Delta^2},$$

где

t - коэффициент доверия (аргумент функции Лапласа), который определяется по таблице значений функции Лапласа в соответствии с заданной доверительной вероятностью;

σ - среднеквадратическое отклонение;

Δ - точность оценки.

Таблица 1

Наблюдение за работой каналов обслуживания

№ интервала j	Интервал времени обслуживания, Δ_j	Частота f_j
1	$t_0 - t_1$	f_1
2	$t_1 - t_2$	f_2
3	$t_2 - t_3$	f_3
…	….	…
R	$t_{R-1} - t_R$	f_R

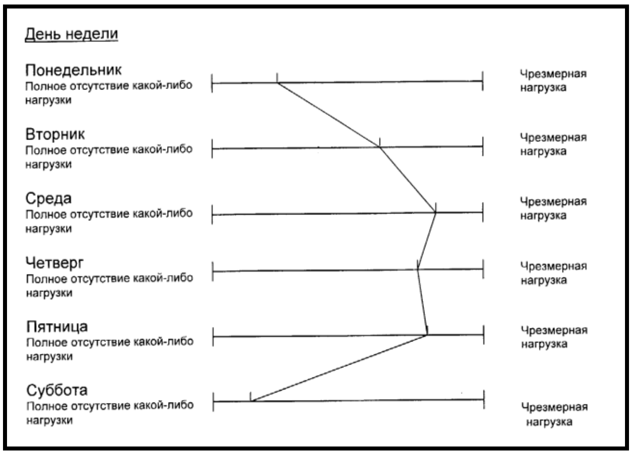

Рис. 4. Семантический дифференциал для оценки распределения нагрузки в течение рабочей недели

Однако в большинстве случаев получаемое по указанной формуле число наблюдений будет корректироваться с учетом ограничений по времени проведения наблюдений и факторов организационно-экономического характера.

Здесь необходимо отметить, что проведение сплошных наблюдений за работой СОТ является достаточно трудоемким и относительно дорогостоящим делом, требующим либо отвлечения собственных работников от их основной работы для наблюдения за процессом торгового обслуживания, либо привлечения сторонних наблюдателей. Однако в большинстве случаев показатели работы каналов обслуживания являются вполне стабильными в течение длительных периодов времени, и поэтому в сплошных наблюдениях за работой обслуживающей системы при каждом проведении оперативного анализа нет особой необходимости. Они необходимы при существенных изменениях интенсивности потока требований, поступающих в систему, а также изменениях в организации торгово-технологического процесса, распределении функций между торговыми работниками и т.д.

1.3. Моментные наблюдения за состоянием каналов и обслуживающей системы

Определение параметров работы каналов обслуживания на предыдущем этапе позволяет осуществить оперативный анализ состояния каналов обслуживания и обслуживающей системы в целом и дать оценку качеству ее функционирования. Оперативный экономико-математический анализ основан на проведении моментных наблюдений за состоянием каналов обслуживания. Получаемые в результате наблюдений данные используются для построения матрицы перехода системы за шаг процесса.

Данные моментных наблюдений регистрируются в специальном статистическом формуляре, который называется *Картой состояния каналов обслуживания* (см. табл. 2).

Важнейшими вопросами, которые необходимо решить на данном этапе, являются определение периода времени для проведения наблюдений, числа наблюдений и периода времени между последовательными наблюдениями, т. е. величины шага процесса.

Время проведения моментных наблюдений необходимо выбирать исходя из данных, полученных в ходе опроса персонала, в частности, на основании карты распределения нагрузки, полученной с помощью семантического дифференциала. Выбор числа моментных наблюдений осуществляется аналогично тому, как это делалось на предыдущем этапе при проведении сплошных наблюдений. В качестве шага процесса необходимо выбрать величину, примерно равную среднему времени обслуживания или кратную этому значению, т. е.

$$\Delta t = k\, t_{обсл},\ k \geq 1.$$

Таблица 2

Карта состояния каналов обслуживания

Момент времени наблюдения	Число требований в системе			
	Канал № 1	Канал № 2	Канал № 3	Всего
10:00	2	1	1	4
10:05	1	2	1	4
10:10	1	2	0	3
10:15	2	1	1	4
10:20	2	1	0	3
10:25	1	2	1	4
10:30	1	1	1	3
…	…	…	…	…

Тогда удается не только оценить состояние обслуживающей системы путем построения матрицы ее перехода, но и определить интенсивность входящего потока требований, непосредственно не наблюдая за ним.

При выборе шага процесса необходимо учитывать требование однородности каналов и в формулярах объединять данные только по каналам, однородным по своим характеристикам. Если же каналы неоднородные, то их состояние необходимо фиксировать через различные интервалы времени и использовать для этого различные формуляры.

1.4. Построение матрицы перехода системы и расчет вектора предельного распределения системы по состояниям

На основе полученных в результате моментных наблюдений данных о состоянии каналов и системы обслуживания строится матрица перехода системы за шаг процесса:

$$P = \begin{pmatrix} p_{00}\, p_{01} \dots p_{0L} \\ p_{10}\, p_{11} \dots p_{1L} \\ \dots\dots\dots\dots \quad \dots\dots \\ p_{L0}\, p_{L2} \dots p_{LL} \end{pmatrix}.$$

Например, используя данные *Карты состояния каналов обслуживания*, представленной в примере выше, можно построить матрицы перехода каждого из каналов и системы в целом за шаг процесса $\Delta t = 5$ *мин.* в период времени с 10:00 до 10:30. Для канала № 1 матрица перехода между состояниями S_1 и S_2 будет выглядеть следующим образом:

$$P_1 = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

Для обслуживающей системы переходы между состояниями S_3 и S_4 будут осуществляться с вероятностями, представленными следующей матрицей:

$$P_{сист} = \begin{bmatrix} 0 & 1 \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}.$$

Далее на основе матрицы перехода системы необходимо вычислить вектор предельного распределения системы по состояниям $p = (p(0), p(1),\dots,p(L))$, который характеризует работу СОТ в устойчивом (или, другими словами, стационарном) состоянии. Здесь $p(k)$ - вероятность того, что система за единицу времени Δt перейдет в состояние S_k, $k = 0,\dots,L$. Очевидно, что должно выполняться следующее равенство:

$$\sum_{k=0}^{L} p(k) = 1\ .$$

Расчет вектора p осуществляется путем решения системы линейных уравнений

$$A \times p = b,$$

где

$$A = \begin{bmatrix} 0 & 1 - p_{00} + p_{10} & \dots & 1 - p_{00} + p_{L0} \\ 1 - p_{11} + p_{01} & 0 & \dots & 1 - p_{11} + p_{L1} \\ \dots & \dots & \dots & \dots \\ 1 - p_{LL} + p_{0L} & 1 - p_{LL} + p_{1L} & \dots & 0 \end{bmatrix};$$

$$b = \begin{bmatrix} 1 - p_{00} \\ 1 - p_{11} \\ \dots \\ 1 - p_{LL} \end{bmatrix};\ p = \begin{bmatrix} p(0) \\ p(1) \\ \dots \\ p(L) \end{bmatrix}.$$

Замечание. Данный расчет нельзя осуществлять для цепей Маркова с так называемыми поглощающими состояниями, т. е. состояниями, из которых можно вернуться только обратно (p_{ii}=1). Однако такие процессы не характерны для систем обслуживания.

После построения матрицы перехода и определения вектора предельного распределения системы по состояниям можно переходить к следующему этапу оперативного анализа качества работы СОТ - расчету показателей, характеризующих качество функционирования СОТ.

1.5. Расчет показателей качества функционирования СОТ

Само по себе предельное распределение вероятностей, как, впрочем, и матрица перехода, содержит важную информацию о характеристиках работы СОТ, позволяя оценить вероятности нахождения системы в тех или иных состояниях и переходов между ними. Кроме того, имеется возможность рассчитать следующие показатели качества работы системы обслуживания:

1. Среднее время пребывания системы в состоянии S_k:
   $T_k = P(k)\ T,\ k = 0,\dots,L.$
2. Средняя периодичность возвращения системы

в состояние S_k, выраженная в шагах процесса Δt: $\omega_k = 1/P(k)$, и в единицах времени:

$$T\omega_k = \Delta t / P(k), k = 0,\dots,L.$$

3. Среднее число требований в системе:

$$\overline{L} = \sum_{k=0}^{L} kP(k).$$

4. Средняя длина очереди:

$$\overline{r} = \sum_{k=n+1}^{L}(k-n)P(k).$$

5. Среднее число занятых каналов:

$$\overline{z} = \overline{L} - \overline{r} = \sum_{k=0}^{L} \min\{k; n\}P(k).$$

6. Среднее число каналов, простаивающих в ожидании требований:

$$\overline{w} = n - \overline{z} = \sum_{k=0}^{n-1}(n-k)P(k).$$

7. Коэффициенты занятости и простоя каналов:

$$K_з = \frac{\overline{z}}{n};$$

$$K_{пр} = 1 - K_з = \frac{\overline{w}}{n}.$$

8. Суммарное время:

* эксплуатации каналов системы:

$$T_z = \overline{z}T;$$

* простоя каналов системы:

$$T_w = \overline{w}T;$$

* пребывания в очереди всех требований:

$$T_r = \overline{r}T.$$

9. Распределение времени, в течение которого:	%	В единицах времени
* в системе отсутствуют требования	$P_0 = P(0)$	$T P_0$
* в системе отсутствует очередь	$P_z = \sum_{k=0}^{n} P(k)$	$T P_z$
* в системе существует очередь	$P_w = \sum_{k=n+1}^{L} P(k)$	$T P_w$

10. Если известны расходы, связанные с работой системы:

- эксплуатационные расходы C_z в расчете на единицу времени работы одного канала;
- издержки C_w, связанные с простоем одного канала в течение единицы времени;
- издержки C_r, связанные с пребыванием в очереди одного требования в течение единицы времени, то можно вычислить общие расходы по эксплуатации СОТ: $C = C_z T_z + C_w T_w + C_r T_r$.

1.6. Определение характеристик входящего потока требований за шаг процесса

Выбор в качестве шага процесса Δt среднего времени обслуживания требований каналом обслуживающей системы $t_{обсл}$ позволяет, используя данные *Карты состояния каналов обслуживания*, рассчитать интенсивность входящего потока требований λ за шаг процесса.

Пусть шаг процесса равен среднему времени обслуживания требований каналами, т. е. $\Delta t = t_{обс.}$. Если обслуживающая система включает n каналов, то объем обслуживания за шаг процесса $x(\Delta t)$ будет примерно равен числу каналов n.

Рассмотрим изменение числа требований за шаг процесса с момента времени t_0 по t_1. Обозначим через $S(t_0)$ и $S(t_1)$ число требований в системе соответственно в моменты времени t_0 и t_1, а через λ - интенсивность потока требований, входящих в систему за шаг процесса. Тогда имеет место следующее соотношение для числа требований за шаг процесса (см. рис. 5):

$$S(t_0) + \lambda \approx x(\Delta t) + S(t_1)$$

Следовательно,

$$\lambda \approx x(\Delta t) + S(t_1) - S(t_0).$$

Объем обслуживания за шаг процесса будет зависеть от числа требований в начальный момент времени t_0. Таким образом, имеем:

$$x(\Delta t) = \min\{ n; S(t_0)\} \text{ и } \lambda \approx \min\{ n; S(t_0)\} + S(t_1) - S(t_0).$$

Если полученное значение будет отрицательным, то можно предположить, что соответствующее количество требований, находящихся в очереди в момент времени t_0, за шаг процесса покинуло очередь, не дождавшись обслуживания, т. е. получили отказ. Можно также предположить, что произошло повышение интенсивности обслуживания. Это вполне возможно, так как при расчетах мы ориентируемся на среднее время обслуживания. Однако и в первом, и во втором случаях интенсивность входящего в систему потока равна нулю.

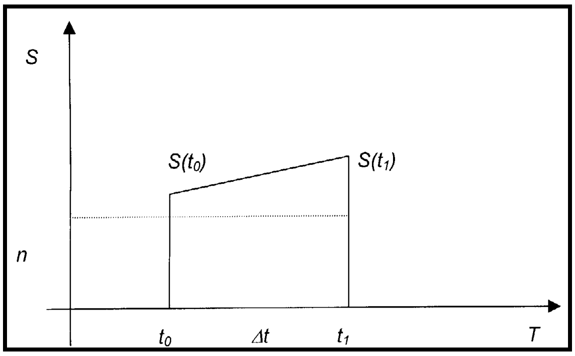

Рис. 5. Изменение числа требований за шаг процесса

Таблица 3

Интенсивности входящего в систему потока требований

Интенсивность λ	Частота f_k
λ_1	f_1
λ_2	f_2
λ_3	f_3
λ_4	f_4
λ_5	f_5
…	…
λ_{s-1}	f_{s-1}
λ_s	f_s

Следовательно, можно записать формулу для нахождения интенсивности потока за шаг процесса следующим образом:

$$\lambda \approx \max\{ \min\{n; S(t_0)\} + S(t_1) - S(t_0); 0\} \quad (1).$$

Рассчитав интенсивности входящего в систему потока требований на каждом из шагов процесса в различные периоды времени, можно построить вариационный ряд (ряд распределения) интенсивности потока и найти ее среднее значение (см. табл. 3).

$$\lambda = \frac{\sum_{k=1}^{s} \lambda_k f_k}{\sum_{k=1}^{s} f_k}.$$

Воспользуемся данными, приведенными в табл. 3 *Карты состояния каналов обслуживания*, и построим ряд распределения интенсивности потока требований, входящего в трехканальную обслуживающую систему. Используя формулу *(1)*, получаем:

λ [10:00;10:05] = 3; λ [10:05;10:10] = = 2; λ [10:10;10:15] = 4; λ [10:15;10:20]= =2; λ [10:20;10:25] = 4; λ [10:25;10:30] = 2.

Построим вариационный ряд интенсивности потока за шаг процесса (см. табл. 4).

Таблица 4

Вариационный ряд интенсивности потока

λ	*f*
2	3
3	1
4	2

Вычислим среднее значение интенсивности потока требований:

$$\lambda = \frac{2 \cdot 3 + 3 \cdot 1 + 4 \cdot 2}{3 + 1 + 2} = \frac{17}{6} \approx$$

$$\approx 2{,}83 \ \frac{чел.}{шаг \ процесса} \approx 0{,}57 \ \frac{чел.}{мин.}$$

1.7. Анализ полученных данных и принятие соответствующих управленческих решений

Полученные в ходе опроса персонала и наблюдений данные позволяют вполне объективно оценить качество функционирования системы обслуживания, выявить возможные недостатки в организации торгово-технологического процесса и управления персоналом предприятия торговли. Собственно говоря, на этом оперативный анализ качества функционирования СОТ завершается. Далее на основе оценки полученных руководством предприятия данных принимается решение о проведении более углубленного системного анализа качества работы системы обслуживания или же полагается, что качество функционирования СОТ удовлетворяет принятым на предприятии требованиям и, таким образом, анализ завершается.

Рис. 6. Этапы проведения системного анализа качества функционирования СОТ

Необходимо отметить, что периодичность проведения рассмотренных выше исследований в значительной степени будет зависеть от специфики работы конкретного предприятия торговли и реализации его СОТ. При планировании проведения оперативного или более углубленного системного анализа необходимо учитывать такие факторы, влияющие на показатели качества работы СОТ, как колебания спроса на основные товары, изменения в организации торгово-технологического процесса, системе управления, изменение численности торгового и обслуживающего персонала.

В случае если данные, полученные в ходе проведения оперативного анализа, свидетельствуют о неудовлетворительном качестве работы обслуживающей системы, необходимо переходить к более масштабному системному анализу работы СОТ. На рис. 6 показан порядок его проведения.

Как видно из приведенной схемы, на первоначальном этапе требуется исследовать основные факторы, оказывающие влияние на процесс торгового обслуживания. К таким факторам относятся организация торгово-технологического процесса, товарный ассортимент, а также организация труда и мотивация торгового персонала. Только после выявления и устранения соответствующих недостатков будут созданы условия для проведения анализа с применением экономико-математических методов.

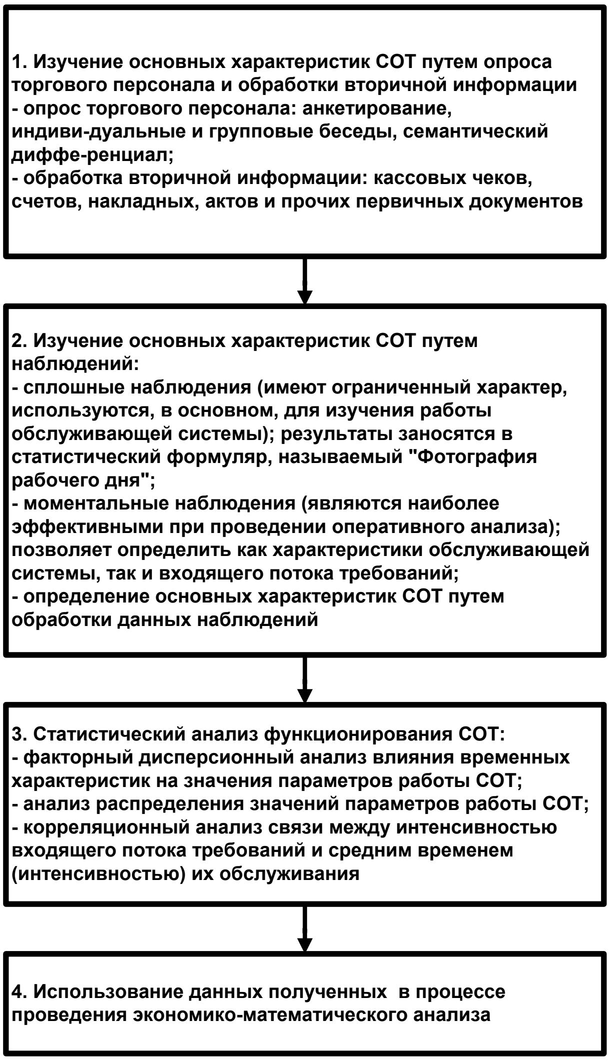

Рис. 7. Стадии формирования информационного обеспечения системного анализа качества функционирования СОТ

В математическом плане системный анализ представляет собой более углубленный анализ качества функционирования СОТ, который предполагает осуществление статистического анализа данных наблюдений и, при необходимости, сбор дополнительной информации о работе СОТ, уточнение математической модели СОТ, а также оптимизацию работы СОТ на основе выбранного критерия качества работы системы обслуживания. Стадии формирования информационного обеспечения экономико-математического анализа представлены на рис. 7.

Достаточный объем собранной информации о работе системы обслуживания позволяет осуществить уточнение математической модели СОТ по следующей схеме:

Марковская цепь → Пуассоновский процесс → Общая классическая модель СОТ (разомкнутая, однофазная, однородная, с очередями) → Частные модели СОТ (замкнутая, с отказами, с неограниченным ожиданием в очереди, с ограниченным временем ожидания).

Однако здесь необходимо отметить, что для применения большинства методов оптимизации уточнения модели СОТ не требуется.

Применение подхода, основанного на рассмотрении процесса функционирования СОТ как марковского случайного процесса, позволяет использовать значительный банк экономико-математических методов оптимизации ее работы на основе различных критериев качества (см. рис. 8).

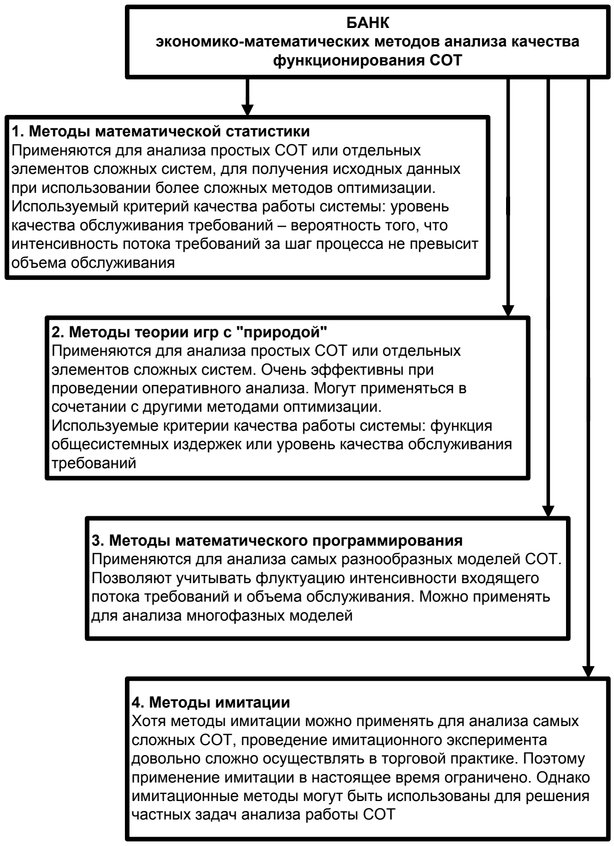

Рис. 8. Классификация экономико-математических методов оптимизации функционирования СОТ

Далее рассмотрим процедуру проведения углубленного экономико-математического анализа качества функционирования СОТ, который осуществляется в рамках системного анализа, и соответствующие информационные технологии, применяемые при этом. При этом будем предполагать, что на предприятии предварительно были осуществлены необходимые мероприятия по совершенствованию организации торгово-технологического процесса, системы управления персоналом, анализа и прогнозирования спроса, а также мероприятия по обеспечению стабилизации товарного ассортимента.

2. УГЛУБЛЕННЫЙ ЭКОНОМИКО-МАТЕМАТИЧЕСКИЙ АНАЛИЗ КАЧЕСТВА ФУНКЦИОНИРОВАНИЯ СОТ

Как уже отмечалось выше, целью углубленного анализа качества работы системы обслуживания является определение оптимальных, с точки зрения выработанных руководством предприятия критериев качества торгового обслуживания покупателей, значений показателей работы СОТ. К таким показателям относятся, в первую очередь, число каналов на каждой из фаз обслуживания, характеристики и режим работы каналов.

Возможно, что для проведения оптимизации потребуется организовать сбор дополнительной информации о характеристиках входящего потока требований и системы их обслуживания. Поэтому на данном этапе анализа могут быть осуществлены рассмотренные выше процедуры проведения опроса персонала и наблюдений за работой системы обслуживания. Однако для того чтобы максимально сократить объем требующейся информации о характеристиках СОТ, необходимо предварительно осуществить статистический анализ уже имеющихся данных о работе СОТ, а затем при необходимости выполнять такой анализ параллельно с процессом сбора дополнительной информации о системе. Поэтому первый этап углубленного экономико-математического анализа можно определить следующим образом.

2.1. Статистический анализ данных о качестве работы СОТ и сбор дополнительной информации

Статистический анализ данных, полученных в результате наблюдений за работой СОТ, включает:

1) Оценку точности значений показателей функционирования СОТ. Получаемые в результате наблюдений значения показателей функционирования СОТ нуждаются в соответствующей статистической оценке их точности. Особенно это важно в случае, когда время проведения наблюдений ограничено и, соответственно, объем выборки оказывается небольшим. Тогда точечная оценка показателей оказывается слишком грубой. Поэтому следует применять интервальную оценку.

Для интервальной оценки интенсивности λ входящего потока за шаг процесса, необходимо построить доверительный интервал. Пусть на основе выборки объема N значений случайной величины Λ - числа требований, поступающих в систему за шаг процесса, определена выборочная средняя

$$\lambda = \frac{1}{N}\sum_{i=1}^{N}\lambda_i .$$

Построим доверительный интервал $\Delta = (\lambda - \delta;$ $\lambda + \delta)$, который покрывает неизвестный параметр λ' с заданной надежностью $\alpha = P(\lambda - \delta < \lambda' < \lambda + \delta)$, где положительное число δ - точность оценки параметра. Чаще всего надежность α выбирается равной 0.90; 0.95 или 0.99.

Тогда, задав значение надежности оценки α и зная δ, можно определить верхнюю границу доверительного интервала Δ: $\lambda_{верх} = \lambda + \delta$. Следовательно, с надежностью α можно будет утверждать, что интенсивность потока не превысит данного значения.

Для малых выборок, когда объем $N < 30$, при определении δ используют распределение Стьюдента и полагают, что

$$\delta = t_\alpha \frac{s}{\sqrt{N}},$$

где t_α - значение случайной величины T, имеющей распределение Стьюдента с $k = N - 1$ степенями свободы. Значение t_α зависит от объема выборки N (степеней свободы k) и надежности α (в нашем случае - от уровня качества обслуживания p) и определяется по таблице значений $t_\alpha = t(\alpha, k)$; s - исправленное среднеквадратическое отклонение,

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\lambda_i - \lambda)^2} .$$

Для выборок объема $N \geq 30$ полагают

$$\delta = t\frac{\sigma}{\sqrt{N}},$$

где t - аргумент функции Лапласа - определяется из соотношения $\Phi(t) = \alpha / 2$ по таблице значений данной функции; σ - среднеквадратическое отклонение,

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\lambda_i - \lambda)^2} .$$

Аналогично осуществляется интервальная оценка интенсивности обслуживания μ.

2) Факторный анализ влияния времени на значения показателей работы СОТ. Важным элементом анализа функционирования СОТ является изучение влияния различных факторов на показатели работы системы. В первую очередь, вызывает интерес влияние на интенсивность входящего потока требований фактора времени. Другими словами, необходимо определить, существенно ли изменяется интенсивность потока в различные периоды времени. Если влияние этого фактора незначительно, то данные о потоке требований, собранные в различные периоды времени, можно объединить, тем самым увеличивая объем необходимой выборки для более точной оценки интенсивности потока требований. Если же влияние фактора времени существенно, то данные об интенсивности потока за разные периоды объединять нельзя, и анализ работы СОТ необходимо выполнять для каждого из периодов отдельно.

Для оценки влияния качественного фактора на изучаемую величину можно воспользоваться дисперсионным анализом. Основная идея дисперсионного анализа заключается в сравнении дисперсии, порождаемой изучаемым фактором (факторной дисперсии), и дисперсии, порождаемой другими случайными причинами (остаточной дисперсии). Если различие между дисперсиями значимо, то фактор оказывает существенное влияние на изучаемую величину.

Предположим наиболее общий случай, когда в различные периоды времени $T_1, T_2, \ldots, T_p$ производилось различное число наблюдений за интенсивностью входящего потока требований: соответственно $q_1, q_2, \ldots, q_p$. Результаты наблюдений можно представить в виде табл. 5.

Групповые средние интенсивностей на каждом уровне вычисляются по формуле:

$$\bar{\lambda}_j = \frac{1}{q_j}\sum_{i=1}^{q_j}\lambda_{ij}, \quad j = 1,\ldots, p.$$

Соответственно, общая средняя будет вычисляться по формуле:

$$\bar{\lambda} = \frac{1}{N}\sum_{j=1}^{p}\sum_{i=1}^{q_j}\lambda_{ij}, \qquad N = \sum_{j=1}^{p} q_j.$$

Таблица 5

Результаты наблюдений

Уровни фактора (периоды)	T_1	T_2	…	T_p
Значения интенсивности входящего потока на разных уровнях	λ_{11}	λ_{12}	…	λ_{1p}
	λ_{21}	λ_{22}	…	λ_{2p}
	…	…	…	…
	$\lambda_{q1\,1}$	$\lambda_{q2\,2}$	…	$\lambda_{qp\,p}$
Групповые средние	$\bar{\lambda}_1$	$\bar{\lambda}_2$	…	$\bar{\lambda}_p$

Рассмотрим величину

$$S_{общ} = \sum_{j=1}^{p}\sum_{i=1}^{q_j}(\lambda_{ij} - \bar{\lambda})^2.$$

Эта величина называется *общей суммой квадратов* и характеризует вариацию отдельных значений интенсивностей от средней под влиянием как фактора времени, так и других случайных факторов. Представим эту сумму в следующем виде:

$$S_{общ} = \sum_{j=1}^{p}\sum_{i=1}^{q_j}(\lambda_{ij} - \bar{\lambda})^2 = \sum_{j=1}^{p}\sum_{i=1}^{q_j}(\lambda_{ij} - \bar{\lambda}_j + \bar{\lambda}_j - \bar{\lambda})^2 =$$

$$= \sum_{j=1}^{p}\sum_{i=1}^{q_j}(\lambda_{ij} - \bar{\lambda}_j)^2 + 2\sum_{j=1}^{p}\sum_{i=1}^{q_j}(\lambda_{ij} - \bar{\lambda}_j)(\bar{\lambda}_j - \bar{\lambda}) + \sum_{j=1}^{p}\sum_{i=1}^{q_j}(\bar{\lambda}_j - \bar{\lambda})^2 =$$

$$= \sum_{j=1}^{p}\sum_{i=1}^{q_j}(\lambda_{ij} - \bar{\lambda}_j)^2 + \sum_{j=1}^{p} q_j(\bar{\lambda}_j - \bar{\lambda})^2 = S_{ост} + S_{факт}$$

Сумма $S_{факт}$ называется факторной суммой и характеризует влияние на случайную величину (интенсивность входящего потока требований) фактора времени. Сумма $S_{ост}$ называется остаточной суммой и характеризует влияние на интенсивность потока других случайных факторов.

Для оценки значимости влияния фактора времени на интенсивность потока требований воспользуемся *критерием Фишера – Снедекора*. Для этого необходимо вычислить следующее отношение:

$$F = \frac{S_{факт}/(p-1)}{S_{ост}/(N-p)},$$

где F - случайная величина, имеющая распределение Фишера с $(p-1)$ и $(N-p)$ степенями свободы. Назовем это значение наблюдаемым $F_{набл}$.

Критерий Фишера – Снедекора имеет следующий вид:

$$P\,[\,F_{набл} > F_{кр}(\alpha;\, p-1;\, N-p)] = \alpha,$$

где α - уровень значимости, который определяет вероятность того, что предположение о незначительном влиянии фактора времени на интенсивность потока будет ошибочным; $F_{кр}(\alpha;\, p-1;\, N-p)$ - критическая точка распределения Фишера, значение которой определяется по соответствующей таблице.

Тогда, если $F_{набл} > F_{кр}(\alpha;\, p-1;\, N-p)$, то предположение о незначительном влиянии фактора времени будет отвергнуто, и, следовательно, в различные периоды времени интенсивность входящего в СОТ потока требований существенно отличается. Если же будет выполняться соотношение

$F_{набл} \leq F_{кр}(\alpha;\, p-1;\, N-p)$,

то интенсивность потока не зависит от того или иного периода времени, на котором она изучается.

Если мы имеем дело с социально-технической системой обслуживания, то можно исследовать влияние фактора времени и на изменение среднего времени (интенсивности) обслуживания. Очевидно, что схема расчетов будет аналогичной указанной выше. В случае существенного влияния фактора времени на показатели обслуживания необходимо обрабатывать данные о работе каналов в различные периоды времени раздельно.

3) Анализ распределения значений показателей работы СОТ. Выше уже говорилось о том, что в рамках углубленного экономико-математического анализа можно осуществить уточнение исходной модели СОТ. Например, можно доказать, что процесс поступления требований в систему является пуассоновским, а время их обслуживания подчиняется показательному закону распределения. Это дает возможность воспользоваться классическими методами теории массового обслуживания для расчета показателей качества работы СОТ.

Таблица 6

Эмпирическое распределение случайной величины *X*

Варианты	Эмпирические частоты
x_1	f_1
x_2	f_2
…	…
x_s	f_s

Если необходимо проверить гипотезу о распределении случайной величины по определенному закону распределения, то необходимо для тех же вариант вычислить теоретические частоты $f'_1, f'_2, \ldots, f'_s$. Затем вычисляем случайную величину:

$$\chi^2 = \sum_{i=1}^{s}\frac{(f_i - f'_i)^2}{f'_i} \qquad (2)$$

Это значение называется наблюдаемым значением критерия Пирсона и обозначается как $\chi^2_{набл}$.

В качестве критерия проверки гипотезы о предполагаемом распределении случайной величины X используют следующее выражение:

$$P\,[\,\chi^2_{набл} > \chi^2_{кр}(\alpha;\, k)] = \alpha,$$

где

α - уровень значимости, который определяет вероятность того, что принятая гипотеза о совпадении теоретического и эмпирического распределений будет ошибочной. Обычно α выбирают равной *0.05; 0.01; 0.025*;

ν - число степеней свободы, которое определяется из соотношения: $\nu = s - 1 - u$, где u - число параметров предполагаемого распределения, которые оценены по данным выборки;

$\chi^2_{кр}$ $(\alpha;\ \nu)$ - критическая точка распределения χ^2, определяемая по таблице значений критических точек.

Тогда, если $\chi^2_{набл} \le \chi^2_{кр}(\alpha;\ \nu)$, то нет оснований отвергать гипотезу о предполагаемом законе распределения случайной величины X. Если же $\chi^2_{набл} > \chi^2_{кр}(\alpha;\ \nu)$, то гипотеза отвергается.

Теперь рассмотрим применение критерия Пирсона для подтверждения гипотез о предполагаемых распределениях показателей работы СОТ. Прежде всего, необходимо проверить правильность предположения о том, что входящий поток требований является пуассоновским. При этом будем учитывать важное свойство, которым обладают случайные величины, распределенные по закону Пуассона: математическое ожидание случайной величины равно ее дисперсии, т. е.

$$\lambda = \sigma^2, \quad \sigma^2 = \frac{\sum_{i=1}^{s} (\lambda_i - \lambda)^2 f_i}{\sum_{i=1}^{s} f_i}.$$

Если это условие выполняется, то в первом приближении можно предположить, что поток требований пуассоновский. Для более точной оценки выдвинутого предположения воспользуемся критерием Пуассона. Вначале вычисляем теоретические частоты f_i'. Формула Пуассона имеет вид:

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

Тогда теоретические частоты f'_i вычисляются следующим образом:

$$f_i' = NP_i(k) = N\frac{\lambda^k e^{-\lambda}}{k!}, \quad N = \sum_{i=1}^{s} f_i,\ k = \lambda_1, \ldots, \lambda s,\ i = 1, \ldots, s$$

Проверка гипотезы о предполагаемом законе неизвестного распределения производится при помощи специально подобранной случайной величины (критерия) - *критерия согласия*. Для проверки гипотез воспользуемся *критерием Пирсона* χ^2 (*"хи квадрат"*). Пусть получено эмпирическое распределение случайной величины X, приведенное в табл. 6 значение критерия Пирсона. Затем по заданному уровню значимости α и числу степеней свободы $\nu = s-2$ $(u = 1)$, так как распределение Пуассона имеет один параметр λ) определяется критическое значение критерия Пирсона $\chi^2_{кр}(\alpha;\ \nu)$ по таблице значений критических точек распределения χ^2.

Определив по таблице $\chi^2_{кр}(\alpha;\ \nu)$, сравниваем: если $\chi^2_{набл} \le \chi^2_{кр}(\alpha;\ \nu)$, то можно принять выдвинутое предположение о том, что входящий поток является пуассоновским.

Теперь рассмотрим применение критерия Пирсона для обоснования предположения о том, что время обслуживания требований является случайной величиной, распределенной по показательному закону. Проверим следующее условие, которое выполняется для случайных величин, распределенных по показательному закону:

$$M(t) = t_{обсл} = \frac{1}{\mu} = \sigma, \qquad где \quad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{R}(\overline{t_i} - t_{обсл})^2 f_i}$$

т. е. математическое ожидание $M(T)$ случайной величины T равно среднеквадратическому отклонению σ.

Таким образом, если вычислить значение σ и будет выполняться соотношение $\sigma \approx 1/\mu$ или $\sigma \approx t_{обсл}$, то можно предположить, что время обслуживания требований распределено по показательному закону. Точную оценку мы получим, если вновь воспользуемся критерием Пирсона.

Вычислим теоретические частоты f_i`$, i = I, \ldots, R$. Воспользуемся тем свойством, что вероятность попадания в интервал (a, b) случайной величины T, распределенной по показательному закону определяется формулой:

$$P(a < T < b) = e^{-\mu a} - e^{-\mu b}.$$

Тогда вероятность P_i попадания случайной величины T в интервал $\Delta_i = [t_{i-1};\ t_i]$ будет определяться таким образом:

$$P_i = P(t_{i-1} < T < t_i) = e^{-\mu t_{i-1}} - e^{-\mu t_i}, \quad i = 1, \ldots, R.$$

Поэтому теоретические частоты f_i' будут вычисляться по формуле:

$$f_i' = NP_i = N(e^{-\mu t_{i-1}} - e^{-\mu t_i}), \quad N = \sum_{i=1}^{R} f_i, \qquad i = 1, \ldots, R$$

Затем по формуле *(2)* вычисляем наблюдаемое значение критерия Пирсона. По заданному уровню значимости α и числу степеней свободы $\nu = s - 2$ ($u = 1$, так как показательное распределение имеет один параметр μ) определяется критическое значение критерия Пирсона $\chi^2_{кр}(\alpha;\ \nu)$ по таблице значений критических точек распределения χ^2. Определив по таблице значение $\chi^2_{кр}(\alpha;\ \nu)$, сравниваем:

если $\chi^2_{набл} \le \chi^2_{кр}(\alpha;\ \nu)$, то можно принять выдвинутое предположение о том, что время обслуживания имеет показательное распределение.

В общем случае показатели работы СОТ могут быть распределены по закону, отличному от рассмотренных выше. Например, интенсивность входящего потока требований может быть распределена равномерно или нормально. Критерий Пирсона позволяет проверить гипотезу о равномерном и нормальном распределении интенсивности входящего потока или времени обслуживания.

4) Корреляционный анализ связи между интенсивностью входящего потока требований и средним временем (интенсивностью) их обслуживания. Если СОТ является социально-технической системой, и наблюдаются существенные перегрузки персонала в отдельные периоды времени, то важное значение приобретает анализ существования возможной связи между интенсивностью потока требований λ и средней продолжительностью их обслуживания $t_{обсл}$.

Для этого на основе интервального ряда распределения времени обслуживания (см. табл. 1) необходимо построить таблицу распределения времени обслуживания в зависимости от значений интенсивности потока (см. табл. 7). Вычислим середины интервалов продолжительностей обслуживания:

$$\tau_i = \frac{t_i + t_{i-1}}{2}, \qquad i = 1, \ldots, R.$$

Таблица 7

Распределение времени обслуживания

Интервалы	Интенсивности потока

времени обслуживания	λ_1	λ_2	…	λ_s
τ_1	f_{11}	f_{12}	…	f_{1s}
τ_2	f_{21}	f_{22}	…	f_{2s}
…				
τ_R	f_{R1}	f_{R2}	…	f_{Rs}

В табл. 7 отражена *статистическая зависимость* $t_{обсл}$ от λ, так как при изменении интенсивности потока требований происходит изменение распределения продолжительности их обслуживания. Продолжим обработку полученных данных с целью выявления возможной корреляционной связи между изучаемыми показателями.

Корреляционная связь между случайными величинами означает, что изменение значений одной величины приводит к изменению среднего значения другой. Следовательно, необходимо вычислить средние значения продолжительностей обслуживания при каждом значении λ_j, $j = 1,\ldots,s$:

$$t_j = \frac{\sum_{i=1}^{R} \tau_i f_{ij}}{\sum_{i=1}^{R} f_{ij}}$$

Получаем следующую таблицу, отражающую зависимость $t_{обсл}$ от λ, приведенную в табл. 8.

Таблица 8

Зависимость $t_{обсл}$ от λ

Интенсивность потока	Время обслуживания
λ_1	t_1
λ_2	t_1
…	…
λ_s	t_1

Далее по табличным данным строится *эмпирическая линия регрессии $t_{обсл}$ на λ.*, а затем подбирается зависимость $t_{обсл} = t_{обсл}(\lambda)$ в явном виде. Эта зависимость называется *выборочным уравнением регрессии $t_{обсл}$ на λ*. Для выбора нужной зависимости из множества допустимых используют метод наименьших квадратов.

Согласно *методу наименьших квадратов* из множества допустимых зависимостей вида $t_{обсл} = t_{обсл}(\lambda)$ выбирается та, для которой выполняется следующее условие:

$$Z = \sum_{j=1}^{s} (t_j - t_{обсл}(\lambda_j))^2 \to \min,$$

где t_j и $t_{обсл}(\lambda_j)$ - соответственно табличные и аналитические вычисленные значения продолжительности обслуживания требований.

Затем, используя тот же критерий, вычисляем параметры выборочного уравнения регрессии.

Для проверки полученной зависимости $t_{обсл} = t_{обсл}(\lambda)$ на адекватность реальным данным можно воспользоваться величиной

$$\varepsilon = \frac{100\,\%}{s} \sum_{j=1}^{s} \left| \frac{t_{обсл}(\lambda_j) - t_j}{t_j} \right|,$$

называемой *средней относительной погрешностью прогнозирования*.

При ε **<10%** имеем высокую точность прогнозирования; при ε от *10%* до *20%* имеет место хорошая точность. Точность прогнозирования можно считать удовлетворительной, если ε приняло значение от *20* до *50 %* и в случае, если ε **>50%**,имеет место неудовлетворительная точность.

Для определения тесноты связи между изучаемыми признаками необходимо вычислить *корреляционное отношение* η, если зависимость $t_{обсл} = t_{обсл}(\lambda)$ нелинейная или *коэффициент корреляции* $r_{t\lambda}$ - в случае прямой линии регрессии:

$$\eta = \sqrt{\frac{\sum_{j=1}^{s} (t_{обсл}(\lambda_j) - \bar{t}_{обсл})^2}{\sum_{j=1}^{s} (t_{обсл\,j} - \bar{t}_{обсл})^2}},$$

$$r_{t\lambda} = \frac{s \sum_{j=1}^{s} t_j \lambda_j - \sum_{j=1}^{s} t_j \sum_{j=1}^{s} \lambda_j}{s \sqrt{\sum_{j=1}^{s} (t_j - \bar{t}_{обсл})^2 \sum_{j=1}^{s} (\lambda_j - \bar{\lambda})^2}}.$$

Здесь $\bar{t}_{обсл}$ - средняя продолжительность обслуживания требований, $\bar{\lambda}$ - средняя интенсивность потока требований.

Таблица 9

Качественная оценка тесноты связи

Значение коэффициента корреляции по абсолютной величине	Качественная характеристика связи
менее 0.1	связи нет
0.1 - 0.3	слабая связь
0.3 - 0.5	умеренная связь
0.5 - 0.7	заметная связь
0.7 - 0.9	высокая связь
более 0.9	весьма высокая связь

Величина η изменяется от *0* до *1*. Чем ближе η к *1*, тем теснее связь между временем обслуживания и интенсивностью потока требований. Величина $r_{t\lambda}$ изменяется от *–1* до *1*. Если $r_{t\lambda} > 0$, то линейная зависимость между признаками прямая, если же $r_{t\lambda} < 0$, то зависимость обратная. Чем ближе по абсолютному значению коэффициент корреляции к *1*, тем теснее связь между изучаемыми признаками. Используя таблицу Чеддока, можно перейти от количественной оценки тесноты связи между признаками к качественной (см. табл. 9).

Таким образом, определив вид и степень тесноты связи между временем обслуживания требований и интенсивностью их потока, входящего в систему, можно установить "запас прочности" СОТ, возможные перегрузки персонала по обслуживанию покупателей, а также предупредить естественное снижение качества обслуживания, возникающее при этом.

Рассмотренные информационные технологии позволяют осуществить более детальный анализ работы СОТ, минимизировать необходимый для проведения исследования объем данных и в то же время повысить качество его информационного обеспечения.

2.2. Уточнение математической модели СОТ

Если в результате проведенного статистического анализа данных наблюдений удается доказать, что значения показателей работы СОТ подчиняются определенным законам распределения, то необходимо уточнить исходную математическую модель СОТ. Это позволит дать более точную оценку качества ее функционирования путем расчета дополнительных показателей. Так, если удалось доказать, что входящий в систему поток требований является простейшим (пуассоновским), а время обслуживания требований распределено по показательному закону, то можно вычислить ряд важных показателей работы СОТ в стационарном режиме. При этом система торгового обслуживания будет соответствовать общей классической модели однофазной однородной СМО, работающей в режиме с ожиданием и ограничением на длину очереди.

Представим основные показатели работы СОТ в стационарном режиме.

1. Вероятность того, что в системе отсутствуют требования, вычисляется по формуле Эрланга:

$$P(0)=\left[\sum_{k=0}^{n}\frac{\rho^{k}}{k!}+\frac{\rho^{n+1}}{n!\,n}\left(\frac{\left(\frac{\rho}{n}\right)^{m}-1}{\frac{\rho}{n}-1}\right)\right]^{-1},\quad \text{где}\quad \rho=\frac{\lambda}{\mu}$$

- приведенная плотность потока требований.

2. Вероятность отказа:

$$P_{отк}=\frac{\rho^{L}}{n!\,n^{m}}P(0).$$

3. Вероятность обслуживания требования или относительная пропускная способность системы:

$$P_{обсл}=1-P_{отк}.$$

4. Абсолютная пропускная способность системы, или число фактически обслуженных требований в единицу времени:

$$A=\lambda P_{обсл}.$$

5. Среднее число занятых каналов:

$$\bar{z}=\frac{A}{\mu}=\rho P_{обсл}\ .$$

6. Коэффициент занятости (использования) каналов:

$$K_{з}=\frac{\bar{z}}{n}.$$

7. Коэффициент простоя каналов

$$K_{пр}=1-K_{з}=1-\frac{\bar{z}}{n}.$$

8. Среднее число простаивающих каналов:

$$\bar{N}=n-\bar{z}.$$

9. Среднее число требований в очереди:

$$\bar{r}=\frac{\rho^{n+1}}{n!\,n}\frac{1-\left(\frac{\rho}{n}\right)^{m}\left(m+1-m\frac{\rho}{n}\right)}{\left(1-\frac{\rho}{n}\right)^{2}}P(0).$$

Среднее число требований, связанных с системой

$$\bar{k}=\bar{z}+\bar{r}.$$

11. Среднее время ожидания требования в очереди:

$$\bar{t}_{оч}=\frac{\rho^{n}}{n!\,n\mu}\frac{1-\left(\frac{\rho}{n}\right)^{m}\left(m+1-m\frac{\rho}{n}\right)}{\left(1-\frac{\rho}{n}\right)^{2}}P(0).$$

12. Среднее время обслуживания требования в системе

$$t_{обсл}=\frac{P_{обсл}}{\mu}.$$

13. Среднее время пребывания требований в системе:

$$\bar{t}_{сист}=\bar{t}_{оч}+\bar{t}_{обсл}.$$

14. Имеют место следующие соотношения:

$$\bar{r}=\lambda\bar{t}_{оч},\quad \bar{k}=\lambda\bar{t}_{сист},$$

которые называются формулами Литтла.

15. Вероятность образования очереди:

$$\theta=\frac{\rho^{n}}{n!}\frac{1-\left(\frac{\rho}{n}\right)^{m}}{1-\frac{\rho}{n}}P(0).$$

Кроме того, на основе общей модели СОТ можно построить частные модели систем с отказами, с ограниченным временем ожидания в очереди, без ограничения на длину очереди, с замкнутым потоком требований и рассчитать соответствующие показатели их работы.

2.3. Оптимизация работы СОТ на основе выбранного критерия качества работы системы обслуживания

Экономико-математический анализ качества работы СОТ предполагает не только оценку существующего состояния системы обслуживания, но и оптимизацию ее функционирования.

Этап оптимизации функционирования СОТ начинается, прежде всего, с формулирования критерия качества работы системы, относительно которого и будут определяться оптимальные значения показателей работы системы.

Выработка критерия качества работы системы является наиболее ответственным и сложным этапом в процессе проведения экономико-математического анализа работы СОТ. Любая СОТ представляет собой сложную социально-техническую систему, характеризующуюся рядом показателей ее функционирования: интенсивностью потока требований, входящих в систему, временем их обслуживания, стоимостью каналов обслуживания, издержками, связанными с отказами, очередями и простоем каналов. Существенное значение имеет также и "человеческий фактор".

Степень важности различных показателей работы СОТ в отдельные периоды времени может существенно варьироваться. Поэтому, от *лица, принимающего решения (ЛПР),* в качестве которого обычно выступает руководитель предприятия торговли, требуется правильно определить приоритетные показатели в данное время и в сложившихся условиях работы системы. Это позволит корректно сформулировать критерий качества работы СОТ, относительно которого будет осуществляться оптимизация.

В простейшем случае в качестве критерия для систем с отказами можно потребовать, чтобы вероятность отказа $P_{отк}$ не превышала заранее заданной величины. Например, требование $P_{отк} < 0.1$ означает, что не менее чем в *90 %* случаев система должна справляться с обслуживанием потока требований при заданной интенсивности λ. Можно ограничить среднее время пребывания требования в очереди или в системе. При этом в качестве показателя, подлежащего определению, могут выступать либо число каналов n при заданной интенсивности обслуживания μ, либо интенсивность μ при заданном числе каналов, либо наилучший вариант из $n_1, n_2, ..., n_q$ каналов с интенсивностями $\mu_1, \mu_2, ..., \mu_q$.

Кроме того, необходимо учитывать не только потери, связанные с отказами или с возникновением очередей, но и потери, обусловленные издержками от простоя каналов в ожидании обслуживания, а также издержками эксплуатации системы. Поэтому очень часто критерий качества представляется в виде функции издержек (потерь), связанных с работой системы.

От выбора критерия качества обслуживания зависит и выбор того или иного математического метода для проведения оптимизации. В процессе анализа работы СОТ для проведения оптимизации применяются методы математической статистики, теории игр, математического программирования и имитационные методы. Краткие характеристики указанных методов уже были представлены выше на рис. 8.

Вначале рассмотрим статистические методы анализа работы СОТ. Очень часто в торговой практике не удается осуществить уточнение математической модели СОТ. Поэтому предположим, что за время T на вход однородной n – канальной СОТ поступает поток требований, имеющий произвольный закон распределения числа λ_i поступающих требований за i – й шаг процесса – в единицу времени Δt_i, $i = 1,\dots,q$;

$\Delta t_1 = \Delta t_2 = \dots = \Delta t_q = \Delta t$. При этом полагается, что в пределах каждого из отрезков времени Δt_i, $i = 1,\dots,q$ требования поступают в систему равномерно.

Пусть известна интенсивность μ обслуживания требований каналами системы за шаг процесса. Обозначим через $x = n\mu$ объем обслуживания требований системой за время Δt, где n - количество каналов СОТ.

Введем понятие *уровня качества обслуживания* p требований системой, который определяется как вероятность того, что случайная величина Λ - число требований, поступающих в систему за время Δt - не превысит значения объема обслуживания x, т. е. $P = P(\Lambda < x)$. Таким образом, например, если принять $p = 0.95$, то это будет означать, что в *95%* случаев все требования за интервал времени Δt должны быть обслужены каналами системы. Подобным показателем пользуются в теории управления товарными запасами для определения оптимального объема партии поставки товара при случайном спросе. В данном случае уровень качества обслуживания требований может использоваться как критерий качества работы СОТ для определения оптимального объема обслуживания и, следовательно, оптимального числа каналов системы.

Пусть Λ - дискретная случайная величина с заданным законом распределения $p_i = P(\Lambda = \lambda_i)$, $i = 1,\dots,s$. Построим функцию распределения случайной величины Λ:

$$F(x) = P(\Lambda < x) = \begin{cases} 0, & x \le \lambda_1; \\ p_1, & \lambda_1 < x \le \lambda_2; \\ p_1 + p_2, & \lambda_2 < x \le \lambda_3; \\ \dots & \dots \\ p_1 + p_2 + \dots + p_{s-1}, & \lambda_{s-1} < x \le \lambda_s; \\ 1, & x > \lambda_s. \end{cases}$$

Введем следующие обозначения для полученных интервалов вероятностей:

$\Delta_1 = (0; p_1]$, $\Delta 2 = (p_1; p_1+p_2]$, …, $\Delta_s = (p_1+p_2+\dots+p_{s-1};1]$.

Определим уровень качества обслуживания p. Учитывая, что $p = F(x)$, получаем формулу для вычисления объема обслуживания x: если $p \in \Delta_i$, то $x = \lambda_i$, $i = 1,\dots,s$.

Определив объем обслуживания x и зная интенсивность обслуживания μ, вычисляем число каналов $n=x/\mu$, которое обеспечит заданный уровень качества обслуживания входящего потока требований.

Аналогично оптимальный объем обслуживания x для заданного уровня качества обслуживания p можно найти, если закон распределения числа требований, поступающих в систему за шаг процесса, задается аналитически.

Например, в случае, если поток требований имеет равномерное распределение с функцией плотности

$$f(\lambda) = \frac{1}{\lambda_{max} - \lambda_{min}}, \lambda_{min} < \lambda \le \lambda_{max}$$

и функцией распределения

$$F(x) = P(\Lambda < x) = \frac{x - \lambda_{min}}{\lambda_{max} - \lambda_{min}}, \lambda_{min} < x \le \lambda_{max},$$

то необходимый объем обслуживания x можно определить из соотношения

$$p = P(\Lambda < x) = \frac{x - \lambda_{min}}{\lambda_{max} - \lambda_{min}}.$$

Тогда $x = \lambda_{min} + p(\lambda_{max} - \lambda_{min})$ для заданного уровня качества обслуживания p.

В рассмотренных методах определения оптимального объема обслуживания x значение p устанавливалось достаточно произвольно, исходя из субъективных соображений. Выбор значения уровня обслуживания можно осуществлять более объективно, если ввести функцию издержек $C_{сист}(x, \lambda)$, связанных с ее работой за шаг процесса. Выделим два вида условных издержек, возникающих в процессе эксплуатации системы: $c_{н/обсл.}$ - издержки, связанные с недостаточным объемом обслуживания; $c_{и/обсл.}$ - издержки, связанные с избыточным объемом обслуживания.

На практике значения издержек могут определяться экспертным путем и выступать в качестве весовых коэффициентов, отражая значимость для предприятия торговли потерь, связанных с недостаточным (очередь и отказы) или избыточным (простои каналов) обслуживанием потребителей.

Тогда общесистемные издержки $C_{сист.}(x,\lambda)$ за шаг процесса будут определяться следующим образом:

$$C_{сист}(x, \lambda) = \begin{cases} c_{н/обсл.}(\lambda - x), & x < \lambda; \\ c_{и/обсл.}(x - \lambda), & x \ge \lambda, \end{cases}$$

где λ - общее количество требований, пришедших в систему за шаг процесса, а x - общий объем обслуживания за тот же период.

Если в качестве критерия эффективности работы СОТ выбрать требование минимизации математического ожидания общесистемных издержек $M\{C_{сист}(x,\lambda)\} \to min$, то уровень качества обслуживания будет определяться следующим образом:

$$p = \frac{c_{н/обсл}}{c_{и/обсл} + c_{н/обсл}} \quad (3).$$

Следовательно, задав значения издержек, связанных с недостаточным и избыточным объемом обслуживания, можно более объективно определить значение уровня качества обслуживания, а затем указанными выше способами рассчитать необходимый объем обслуживания и число каналов СОТ.

Если лицом, принимающим решения, был сформулирован единый критерий качества работы системы, то для определения оптимальных значений параметров функционирования СОТ можно воспользоваться методами теории игр и математического программирования. При этом появляется возможность рассматривать более сложные модели СОТ, например, неоднородные и многофазные системы, и использовать более сложные критерии для оценки качества их работы.

Модели теории игр и математического программирования, применяемые для оптимизации функционирования СОТ, можно разделить на модели, в которых не учитывается флуктуация интенсивности входящего в систему потока требований, и модели, в которых флуктуация учитывается. В первом случае задача заключается в определении наилучших, относительно выбранного критерия, значений показателей работы системы на весь рассматриваемый период времени *T*. Во втором случае предполагается, что значение интенсивности входящего потока требований может существенно изменяться на каждом шаге процесса Δt. Поэтому в таких моделях оптимизация работы СОТ проводится для каждого шага процесса.

Как уже указывалось выше, в отличие от технических систем обслуживания, для которых характерна стабильность (стационарность) происходящих в них процессов, СОТ являются социально-техническими и поэтому в значительной степени нестабильными по своим характеристикам. Следовательно, рассматривая даже небольшой период времени, необходимо учитывать возможность существенных изменений в процессах поступления требований и их обслуживания. Например, удобнее и правильнее сразу осуществить распределение торгового персонала в магазине в течение дня с учетом изменения интенсивности потока покупателей в различные периоды времени, нежели проводить оптимизацию для каждого из периодов дня отдельно или же ориентироваться на среднее значение интенсивности потока покупателей.

Поэтому очевидно, что модели СОТ, в которых учитывается флуктуация интенсивности входящего потока требований, имеют большое практическое значение. Однако если шаг процесса является небольшим, то возникающая при этом флуктуация объема обслуживания может свести на нет эффект от проведенной оптимизации. Поэтому возникает необходимость в оптимизации работы СОТ с учетом флуктуации как интенсивности потока требований, так и объема обслуживания. В этом случае задача оптимизации сводится к выбору оптимального объема обслуживания для каждого шага процесса и одновременному сглаживанию возникающих при этом колебаний от шага к шагу.

Вероятностный характер процессов, происходящих в СОТ, позволяет применять методы теории игр для расчета оптимальных значений параметров работы системы обслуживания. Для оптимизации функционирования однофазных СОТ воспользуемся критерием Байеса теории игр с "природой".

Рассмотрим модель однофазной СОТ в предположении, что на рассматриваемом отрезке времени *T* входящий в систему поток требований распределен по закону Пуассона с параметром λ, а время их обслуживания в системе - по показательному закону с параметром μ. Тогда, используя формулы, приведенные выше на этапе уточнения математической модели СОТ, можно рассчитать показатели функционирования системы в стационарном режиме. В частности, можно определить среднее число требований, находящихся в системе, среднее число простаивающих каналов, среднее число отказов и среднее число требований, находящихся в очереди.

В качестве критерия для оценки эффективности функционирования СОТ определим следующую функцию издержек, связанных с работой системы:

$$C_{сист} = C_{экспл}\,\bar{z} + C_{пр}\,(n - \bar{z}) + C_{отк}\,P_{отк}\,\lambda + C_{оч}\,\bar{r} \to min \qquad (4)$$

Здесь:

$C_{экспл}$ - издержки, связанные с эксплуатацией одного канала системы;

$C_{пр}$ - издержки, связанные с простоем одного канала;

$C_{отк}$ - издержки, связанные с одним отказом;

$C_{оч}$ - издержки, связанные с пребыванием одного требования в очереди;

$C_{сист}$ - общесистемные издержки.

Все издержки определяются в расчете на шаг процесса Δt.

Таким образом, задача заключается в минимизации общих издержек, связанных с функционированием системы.

На практике наиболее точно можно определить только значения издержек, связанных с эксплуатацией и простоем каналов СОТ. Значения же издержек, связанных с отказами и ожиданием обслуживания в очереди, можно только предполагать. Поэтому часто, вместо конкретных значений издержек, выраженных в денежных единицах, удобнее и правильнее использовать весовые коэффициенты, определяемые по методу экспертных оценок.

Если значения интенсивностей λ и μ на рассматриваемом отрезке времени ***T*** считаются постоянными, то общие издержки системы являются функцией от числа каналов, т. е. $C_{сист} = C_{сист}(n)$.

Пользуясь критерием *(4),* можно решать более сложные задачи оптимизации работы СОТ в условиях, когда на отрезке времени ***T*** возможны различные интенсивности входящего потока требований λ, интенсивности обслуживания требований в системе μ и число каналов ***n***: $C_{сист} = C_{сист}(\lambda, \mu, n)$. Например, хотя предполагается, что входящий поток требований является пуассоновским, если время ***T*** значительно, сам параметр распределения λ может быть случайной величиной с заданным законом распределения.

Рассмотрим некоторые возможные ситуации, в которых может осуществляться обслуживание требований, и соответствующие методы оптимизации работы СОТ с использованием критерия Байеса.

2.3.1. Выбор оптимального числа каналов для однородной СОТ

Пусть в период времени ***T*** на вход обслуживающей системы подается поток требований, интенсивность которого λ является дискретной случайной величиной с заданным законом распределения

$p_i = P(\Lambda = \lambda_i),\ i = 1,\ldots,s.$

Предположим, что интенсивность μ обслуживания требований одним каналом является величиной постоянной. Для обслуживания поступающего потока требований возможно использовать $n_1, n_2,\dots,n_q$ – канальные системы. Требуется выбрать оптимальный вариант обслуживающей системы на основе критерия качества ее работы *(4)*.

Назовем *i–й стратегией лица, принимающего решение*, выбор им для обслуживания n_i – канальной системы, $i = 1,\dots,q$, а *j–м состоянием "природы"* - входящий поток требований, имеющий в период времени T интенсивность λ_j, $j = 1,\dots,s$. Построим модель теории игр с "природой" - *матрицу выигрышей*

$$A = (a_{ij})_{i=1,\dots,q;\ j=1,\dots,s},$$

где a_{ij} - выигрыш лица, принимающего решение, если он выбирает i–ю стратегию - n_i – канальную систему, а "природа" принимает j–е состояние - входящий поток требований в течение периода времени T имеет интенсивность λ_j.

Учитывая, что выбранный критерий является функцией издержек (проигрышей), будем иметь следующее представление выигрышей:

$$a_{ij} = -C_{сист}(\lambda_j, \mu, n_i),\ i = 1,\dots,q;\ j = 1,\dots,s,$$

Для определения оптимальной стратегии воспользуемся критерием Байеса, согласно которому из множества стратегий $1, 2,\dots, q$ выбирается такая i_{max}, которая обеспечивает максимум математического ожидания выигрыша:

$$i_{max} \in \{1,\dots, q\}: \quad B_{i_{max}} = \max_{i=1,\dots,q} \sum_{j=1}^{s} a_{ij} p_j .$$

Замечание. При использовании указанного критерия необходимо учитывать, что, вообще говоря, получаемое оптимальное решение может быть не единственным.

2.3.2. Выбор оптимальной интенсивности обслуживания для n – канальной однородной СОТ

Пусть в период времени T на вход обслуживающей системы подается поток требований, интенсивность которого λ, является дискретной случайной величиной с заданным законом распределения

$$p_i = P(\Lambda = \lambda_i),\ i = 1,\dots,s.$$

Предположим, что для обслуживания поступающего потока требований возможно использовать q различных n – канальных систем, с интенсивностями обслуживания соответственно $\mu_1, \mu_2, \dots, \mu_q$. Требуется выбрать оптимальный вариант обслуживающей системы на основе критерия качества ее работы (4).

Назовем i–й стратегией лица, принимающего решение, выбор им для обслуживания n – канальной системы с интенсивностью обслуживания μ_i, $i = 1,\dots,q$, а j – м состоянием "природы" - входящий поток требований, имеющий в период времени T интенсивность λ_j, $j= 1,\dots,s$.

Имеет место следующее представление выигрышей: $a_{ij} = -C_{сист}(\lambda_j, \mu_i, n)$, $i = 1,\dots,q;\ j=1,\dots,s$. Как и в предыдущем варианте СОТ, для решения задачи оптимизации используется критерий Байеса.

2.3.3. Выбор оптимального числа каналов для адаптивной СОТ

Рассмотрим еще один вариант оптимизации СОТ, при котором интенсивность обслуживания требований каналами может изменяться в зависимости от интенсивности входящего потока требований, т. е. $\mu = \mu(\lambda)$. Подобная зависимость характерна для большинства социально-технических систем и выше, при рассмотрении первого этапа, был рассмотрен метод анализа такой зависимости. Действительно, вольно или невольно, но продавец в магазине будет стараться обслуживать покупателей быстрее, если наблюдает резкое увеличение очереди у прилавка.

Пусть в период времени T на вход обслуживающей системы подается поток требований, интенсивность которого λ является дискретной случайной величиной. Предположим, что интенсивность обслуживания μ зависит от интенсивности входящего потока требований λ и также является дискретной случайной величиной. Тогда можно рассмотреть дискретную двумерную случайную величину (M, Λ), возможные значения которой таковы:

$$\mu_1, \mu_2, \dots, \mu_u;\ \lambda_1, \lambda_2, \dots, \lambda_s.$$

Пусть заданы условные законы распределения составляющей M при $\Lambda = \lambda_j$, $j=1,\dots,s$ - совокупности условных вероятностей $p(\mu_1 | \lambda_j), p(\mu_2 | \lambda_j), \dots, p(\mu_u | \lambda_j)$, вычисленных в предположении, что событие $\Lambda = \lambda_j$ уже наступило. Следовательно, можно вычислить условные математические ожидания случайной величины M при

$$\Lambda = \lambda_j,\ j=1,\dots,s:$$

$$M_j = M\{M \mid \Lambda = \lambda_j\} = \sum_{i=1}^{u} \mu_i p(\mu_i \mid \lambda_j), \qquad j = 1,\dots, s .$$

Далее предположим, что для обслуживания поступающего потока требований возможно использовать $n_1, n_2,\dots,n_q$ – канальные системы. Требуется выбрать оптимальный вариант обслуживающей системы на основе критерия качества ее работы (4).

Назовем i–й стратегией лица, принимающего решение, выбор им для обслуживания n_i – канальной системы, $i = 1,\dots,q$, а j–м состоянием "природы" - входящий поток требований, имеющий в период времени T интенсивность λ_j, $j = 1,\dots,s$.

Для данного варианта СОТ имеет место следующее представление выигрышей:

$$a_{ij} = -C_{сист}(\lambda_j, M_j, n_i),\ i = 1,\dots,q;\ j = 1,\dots,s.$$

Как и в предыдущих вариантах СОТ, для решения задачи оптимизации используется критерий Байеса.

На основе предложенных вариантов можно рассматривать более сложные случаи оптимизации процесса обслуживания. А теперь рассмотрим общую математическую модель СОТ и методы оптимизации ее работы.

2.3.4. Выбор оптимального объема обслуживания для СОТ с произвольным законом распределения интенсивности входящего потока требований

Рассмотрим еще один вариант СОТ с произвольным законом распределения числа поступающих в систему требований за шаг процесса. В качестве критерия для выбора оптимального объема обслуживания x воспользуемся показателем уровня качества обслуживания p, значение которого определяется с использованием функции общесистемных издержек по формуле (3). Для определения объема обслуживания x, соответствующего заданному уровню качества обслуживания p, воспользуемся критерием Байеса.

Пусть число требований, входящих в систему за шаг процесса Λ - дискретная случайная величина с заданным законом распределения

$p_i = P(\Lambda = \lambda_i), i = 1,\dots,s$. Определим значения издержек $c_{н/обсл}$ и $c_{и/обсл}$ и вычислим соответствующее значение p, используя формулу (3).

Таблица 10

Матрица выигрышей

Стратегии ЛПР	Состояния "природы			
	λ_1	λ_2	..	λ_s
$x_1 = \lambda_1$	0	$-c_{н/обсл}(\lambda_2 - x_1)$	..	$-c_{н/обсл}(\lambda_s - x_1)$
$x_2 = \lambda_2$	$-c_{и/обсл}(x_2 - \lambda_1)$	0	..	$-c_{н/обсл}(\lambda_s - x_2)$
…	…	…	..	…
$x_s = \lambda_s$	$-c_{и/обсл}(x_s - \lambda_1)$	$-c_{и/обсл}(x_s - \lambda_2)$	..	0
Вероятности состояний "природы" p	p_1	p_2	..	p_s

Примем в качестве j – го состояния "природы" значение интенсивности входящего потока требований $\Lambda = \lambda_j$, $j = 1,\dots,s$, а в качестве i – ой стратегии лица, принимающего решение - объем обслуживания x_i, соответствующий интенсивности входящего потока требований λ_i, т. е. $x_i = \lambda_i$, $i = 1, \dots ,s$.

Построим матрицу выигрышей $A = (a_{ij})_{i,j = 1,\dots,s}$, элементы которой вычисляются следующим образом:

$a_{ij} = - C_{сист\ ij}$,

где

$$C_{сист\ ij} = \begin{cases} c_{н/обсл}(\lambda_j - x_i), & x_i < \lambda_j \\ c_{и/обсл}(x_i - \lambda_j), & x_i \ge \lambda_j \end{cases}$$

(см. табл. 10).

Применяя критерий Байеса, выбираем оптимальную стратегию $k = 1,\dots,s$ - необходимый объем обслуживания x_k, соответствующий уровню качества обслуживания p.

2.3.5. Оптимизация работы СОТ с учетом флуктуации интенсивности входящего потока требований

Усложним предыдущий вариант СОТ, предполагая, что на каждом шаге процесса обслуживания интенсивность входящего потока требований существенно изменяется, т. е. наблюдается ее флуктуация. При этом необходимо учитывать не только флуктуацию входящего потока требований, но и возникающую при этом флуктуацию объема обслуживания.

Рассмотрим самый простой вариант, в котором учитывается возможное сильное изменение объема обслуживания на последовательных шагах процесса.

Пусть на вход СОТ поступает поток требований с переменной интенсивностью на каждом шаге процесса $t = 1,\dots,T$. Задается вектор возможных интенсивностей потока $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s)$, а также матрица их вероятностей на каждом из шагов процесса:

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1s} \\ p_{21} & p_{22} & \dots & p_{2s} \\ \dots & \dots & \dots & \dots \\ p_{T1} & p_{T2} & \dots & p_{Ts} \end{bmatrix},$$

где p_{tj} - вероятность того, что на шаге процесса с номером $t = 1,\dots,T$ интенсивность входящего потока примет значение λ_j, $j = 1,\dots,s$. Очевидно, что должны выполняться следующие равенства:

$$\sum_{j=1}^{s} p_{tj} = 1, \qquad t = 1,\dots, T .$$

Далее, используя методику расчетов, представленную в предыдущем варианте, находим вектор значений оптимальных объемов обслуживания на каждом из шагов процесса: $X = (X_1, X_2, \dots, X_T)$. При этом флуктуация интенсивностей входящего потока требований может быть столь значительной, что ориентироваться на полученные оптимальные значения объемов обслуживания при планировании режима работы торгового и обслуживающего персонала будет просто недопустимо. Поэтому необходимо продолжить оптимизацию с использованием более эффективных в этом случае методов математического программирования. Например, для учета флуктуации интенсивности потока требований можно воспользоваться методами нелинейного программирования.

Определим план задачи нелинейного программирования - вектор объемов обслуживания на шагах процесса $x = (x_1, x_2, \dots, x_T)$.

Обозначим через $e_t = | X_t - x_t |$ ошибку, связанную с неправильным выбором объема обслуживания на t – м шаге процесса, $t = 1, \dots, T$. Далее будем считать, что наибольшие потери вызывают изменения объема обслуживания на последовательных шагах процесса. Через $f_t = | x_t - x_{t-1}|$ обозначим флуктуацию объема обслуживания на шагах $t = 2, \dots, T$.

Очевидно, что значимость ошибки, связанной с неправильным выбором объема обслуживания и его флуктуации, неодинакова. Поэтому, введем весовой коэффициент $w > 0$, который будет характеризовать значимость ошибки выбора объема обслуживания.

Чтобы избавиться от абсолютных величин при проведении оптимизации и, вместе с тем, усилить значимость ошибки выбора e_t и флуктуации f_t, целевую функцию $Z(x)$ представим как квадратичную:

$$Z(x) = w\sum_{t=1}^{T} e_t^2 + \sum_{t=2}^{T} f_t^2 = w\sum_{t=1}^{T}(X_t - x_t)^2 + \sum_{t=2}^{T}(x_t - x_{t-1})^2 .$$

Очевидно, что задача будет заключаться в том, чтобы найти такой план обслуживания $x = (x_1, x_2, \dots, x_T)$, чтобы минимизировать значение функции $Z(x)$, т. е. необходимо решить задачу:

$Z(x) \rightarrow min.$

Эта задача на безусловный экстремум решается путем нахождения частных производных функции $Z(x)$ по x_t и приравнивания их к нулю:

$$\frac{\partial Z(x)}{\partial x_t} = 0, \qquad t = 1,\dots, T .$$

Получаем систему из T линейных уравнений с T неизвестными:

$$\begin{aligned} &(1 + w)x_1 - x_2 = wX_1, && при \quad t = 1; \\ &-x_{t-1} + (2 + w)x_t - x_{t+1} = wX_t, && при \quad t = 2,\dots, T - 1; \\ &-x_{T-1} + (1 + w)x_T = wX_T, && при \quad t = T. \end{aligned}$$

В результате решения данной системы определяем значения элементов вектора x - значения объемов обслуживания на каждом из шагов процесса с учетом флуктуации интенсивности потока требований, входящих в систему.

Необходимо отметить, что использование методов линейного и нелинейного программирования позволяет осуществлять оптимизацию работы самых сложных СОТ. В то же время применение компьютерной техники и современных программных средств, таких как табличный процессор *Microsoft Excel*, позволяет снять проблему, связанную с реализацией указанных методов в торговой практике.

Полученные в результате проведения экономико-математического анализа работы СОТ данные используются для принятия соответствующих управленческих решений по повышению качества обслуживания. Однако реальные результаты от проведения оптимизации могут быть получены не сразу. Кроме того, как уже говорилось выше, качество обслуживания зависит от целого комплекса условий. Поэтому анализ работы СОТ, проведенный с использованием самого совершенного математического обеспечения, не сможет компенсировать возможное негативное воздействие на систему множества не учтенных при его проведении факторов и, следовательно, не обеспечит повышения качества обслуживания.

Следовательно, необходимо осуществлять ретроспективный анализ эффективности реализованных решений для выработки рекомендаций по дальнейшему совершенствованию работы СОТ, для формирования типовых решений по различным вопросам ее функционирования, а также для своевременного выявления возможных негативных воздействий на процесс обслуживания.

В заключение необходимо отметить, что постоянный контроль качества функционирования СОТ со стороны руководства предприятия торговли и своевременное проведение системного анализа на практике обеспечивает:

1) оптимальный режим работы торгового и обслуживающего персонала предприятия и оптимальное распределение нагрузки между работниками, повышение эффективности их взаимодействия в процессе обслуживания покупателей. Тем самым создаются условия для повышения качества торгового обслуживания, поддержания нормального микроклимата на предприятии, сокращения числа болезней работников, связанных с перегрузками, и уменьшения текучести кадров;

2) оптимальную нагрузку на технические средства, используемые в процессе торговли, и оптимальный режим их обслуживания. Это позволяет сократить издержки, связанные с их арендой, содержанием и ремонтом, а также оплатой труда обслуживающего персонала.

За счет улучшения организации системы обслуживания формируется позитивный образ предприятия торговли в глазах покупателей, повышается его конкурентоспособность и, таким образом, создается прочная основа для его стабильного развития в современных условиях рынка.

д. э. н., профессор М. И. Баканов