# Системы массового обслуживания

### 1 Введение

Эта теория представляет особый раздел теории случайных процессов и использует, в основном, аппарат теории вероятностей. Первые публикации в этой области относятся к 20-м гг. XX в. и принадлежат датчанину А. Эрлангу, занимавшемуся исследованиями функционирования телефонных станций - типичных СМО, где случайны моменты вызова, факт занятости абонента или всех каналов, продолжительность разговора. В дальнейшем теория очередей нашла развитие в работах К.Пальма, Ф.Поллачека, А.Я.Хинчина, Б.В.Гнеденко, А.Кофмана, Р.Крюона, Т. Cаати и других советских и зарубежных математиков.

Теория массового обслуживания – область прикладной математики, занимающаяся анализом процессов в системах производства, обслуживания, управления, в которых однородные события повторяются многократно, например, на предприятиях бытового обслуживания; в системах приема, переработки и передачи информации; автоматических линиях производства и др.

Предметом теории массового обслуживания является установление зависимостей между характером потока заявок, числом каналов обслуживания, производительностью отдельного канала и эффективным обслуживанием с целью нахождения наилучших путей управления этими процессами.

### 2 Понятие системы массового обслуживания

***Системы массового обслуживания*** (**СМО**) – это такие системы, в которые в случайные моменты времени поступают заявки на обслуживание, при этом поступившие заявки обслуживаются с помощью имеющихся в распоряжении системы каналов обслуживания.

Примерами систем массового обслуживания могут служить: магазины, банки, ремонтные мастерские, почтовые отделения, посты технического обслуживания и ремонта автомобилей, персональные компьютеры, обслуживающие поступающие заявки или требования на решение тех или иных задач, аудиторские фирмы, отделы налоговых инспекций, занимающиеся приемкой и проверкой текущей отчетности предприятий, телефонные станции и т.д.

В теории систем массового обслуживания (СМО) обслуживаемый объект называют *требованием*. В общем случае под требованием обычно понимают запрос на удовлетворение некоторой потребности, например, разговор с абонентом, посадка самолета, покупка билета, получение материалов на складе.

Средства, обслуживающие требования, называются*обслуживающими устройствами* или*каналами обслуживания****.*** Например, к ним относятся каналы телефонной связи, посадочные полосы, мастера-ремонтники, билетные кассиры, погрузочно-разгрузочные точки на базах и складах.

Совокупность однотипных обслуживающих устройств называется *системой массового обслуживания***.** Такими системами могут быть телефонные станции, аэродромы, билетные кассы, ремонтные мастерские, склады и базы снабженческо-сбытовых организаций и т.д.

Основной задачей теории СМО является изучение режима функционирования обслуживающей системы и исследование явлений, возникающих в процессе обслуживания. Так, одной из характеристик обслуживающей системы является время пребывания требования в очереди. Очевидно, что это время можно сократить за счет увеличения количества обслуживающих устройств. Однако каждое дополнительное устройство требует определенных материальных затрат, при этом увеличивается время без­действия обслуживающего устройства из-за отсутствия требований на обслуживание, что также является негативным явлением. Следовательно, в теории СМО возникают задачи оптимизации: каким образом достичь определенного уровня обслуживания (максимального сокращения очереди или потерь требований) при минимальных затратах, связанных с простоем обслуживающих устройств.

### 3 Основные элементы СМО

Основными элементами СМО являются:*входящий поток требований*, *очередь требований*, *обслуживающие устройства*, (каналы) и *выходящий поток требований. Потоком* называют последовательность событий. Поток, состоящий из требований на обслуживание, называют *потоком требований*. Поток требований, поступающих в обслуживающую систему, называют *входящим потоком*. Поток требований, которые обслужены, называют *выходящим потоком*. Совокупность очередей и приборов (каналов) обслуживания называются *системой обслуживания*.

### 4 Структура СМО

|  |
| --- |
| http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_50fa9c2b.png |
| Рисунок 1 – Структурная схема типичной системы массового обслуживания. |

**Источник.**Источник определяется как устройство или множество, из которого требования поступают в систему для обслуживания. Источник называют бесконечным или конечным в зависимости от того, бесконечное или конечное число требований содержится в нем. Будем всегда предполагать, что источник, генерирующий требования, неисчерпаем. Например, хотя абонентов некоторого телефонного узла конечное число, предполагаем, что они образуют бесконечный источник.

**Входящий поток***.* Требования, поступающие из источника на обслуживание, образуют входящий поток. Само требование можно рассматривать как запрос на удовлетворение какой-то потребности. Примеров входящих потоков можно привести множество. Это — поток информации, поступающей на обработку в ЭВМ; поток заявок на АТС; поток клиентов, приходящих в ателье, и больных в поликлинику, поток прибывающих в порт судов; налетающие на объект удара самолеты и ракеты противника и т. д.

**Обслуживающая система.**Под обслуживающей системой понимают множество технических средств или производственного персонала (различного рода установки, приборы, устройства, тоннели, взлетно-посадочные полосы, линии связи, продавцы, бригады рабочих или служащих, кассиры и т. д.), выполняющих функции обслуживания. Все перечисленное выше, как уже говорилось, объединяется одним названием «канал обслуживания» (обслуживающий прибор). Состав системы определяется количеством каналов (приборов, линий). По количеству каналов системы можно подразделить на одноканальные и многоканальные.

**Дисциплина очереди.** Это важный компонент системы массового обслуживания, он определяет принцип, в соответствии с которым поступающие на вход обслуживающей системы требования подключаются из очереди к процедуре обслуживания. Чаще всего используются дисциплины очереди, определяемые следующими правилами:

* первым пришел - первый обслуживаешься;
* пришел последним - обслуживаешься первым;
* случайный отбор заявок;
* отбор заявок по критерию приоритетности;
* ограничение времени ожидания момента наступления обслуживания (имеет место очередь с ограниченным временем ожидания обслуживания, что ассоциируется с понятием «допустимая длина очереди»).

**Механизм обслуживания** определяется характеристиками самой процедуры обслуживания и структурой обслуживающей системы. К характеристикам процедуры обслуживания относятся: продолжительность процедуры обслуживания и количество требований, удовлетворяемых в результате выполнения каждой такой процедуры. Для аналитического описания характеристик процедуры обслуживания оперируют понятием «вероятностное распределение времени обслуживания требований».

Следует отметить, что время обслуживания заявки зависит от характера самой заявки или требований клиента и от состояния и возможностей обслуживающей системы. В ряде случаев приходится также учитывать вероятность выхода обслуживающего прибора по истечении некоторого ограниченного интервала времени.

Структура обслуживающей системы определяется количеством и взаимным расположением каналов обслуживания (механизмов, приборов и т. п.). Прежде всего, следует подчеркнуть, что система обслуживания может иметь не один канал обслуживания, а несколько; система такого рода способна обслуживать одновременно несколько требований. В этом случае все каналы обслуживания предлагают одни и те же услуги, и, следовательно, можно утверждать, что имеет место параллельное обслуживание.

Система обслуживания может состоять из нескольких разнотипных каналов обслуживания, через которые должно пройти каждое обслуживаемое требование, т. е. в обслуживающей системе процедуры обслуживания требований реализуются последовательно. Механизм обслуживания определяет характеристики выходящего (обслуженного) потока требований.

Рассмотрев основные компоненты систем обслуживания, можно констатировать, что функциональные возможности любой системы массового обслуживания определяются следующими основными факторами:

* вероятностным распределением моментов поступлений заявок на обслуживание (единичных или групповых);
* вероятностным распределением времени продолжительности обслуживания;
* конфигурацией обслуживающей системы (параллельное, последовательное или параллельно-последовательное обслуживание);
* количеством и производительностью обслуживающих каналов;
* дисциплиной очереди;
* мощностью источника требований.

В качестве основных критериев эффективности функционирования систем массового обслуживания, в зависимости от характера решаемой задачи могут выступать:

* вероятность немедленного обслуживания поступившей заявки;
* вероятность отказа в обслуживании поступившей заявки;
* относительная и абсолютная пропускная способность системы;
* средний процент заявок, получивших отказ в обслуживании;
* среднее время ожидания в очереди;
* средняя длина очереди;
* средний доход от функционирования системы в единицу времени и т.п.

**Выходящий поток.**Выходящий поток – это поток требований, покидающих систему после обслуживания. Сюда могут входить и требования, которые покинули систему, не пройдя обслуживания.

Предметом теории массового обслуживания является установление зависимости между факторами, определяющими функциональные возможности системы массового обслуживания, и эффективностью ее функционирования. В большинстве случаев все параметры, описывающие системы массового обслуживания, являются случайными величинами или функциями, поэтому эти системы относятся к стохастическим системам.

На рисунке 2 представлена классификация компонентов системы массового обслуживания.

|  |
| --- |
| Классификация СМО |
| Рисунок 2 – Классификация СМО. |

Системы массового обслуживания могут быть классифицированы по ряду признаков.

*В зависимости от условий ожидания начала обслуживания различают*

* ***системы с отказами***, в которых заявка, поступившая в систему в момент, когда все каналы заняты, получает отказ и сразу же покидает очередь;
* ***системы с ожиданием*** (очередью), в которых заявка, поступившая в момент, когда все каналы обслуживания заняты, становится в очередь и ждет, пока не освободится один из каналов.

Системы массового обслуживания с ожиданием делятся на системы с ограниченным ожиданием и системы с неограниченным ожиданием.

В системах с ограниченным ожиданием может ограничиваться:

* длина очереди;
* время пребывания в очереди.

В системах с неограниченным ожиданием заявка, стоящая в очереди, ждет обслуживание неограниченно долго, т.е. пока не подойдет очередь.

*По кругу обслуживаемых объектов различают два вида*.

**Замкнутые СМО**. Замкнутая система массового обслуживания – это система массового обслуживания, в которой обслуженные требования могут возвращаться в систему и вновь поступать на обслуживание. Примерами замкнутой СМО являются ремонтные мастерские, сберегательные банки.

**Открытые (разомкнутые) СМО**. Для открытой СМО предполагается, что исходная совокупность на столько велика, что изменение ее размеров, вследствие прибытия или возвращения обслуженной заявки в исходную совокупность не оказывает существенного влияния на вероятность появления очередной заявки.

*По типу приоритета* *различают два вида:*

О**днородные -**  требования одинаковые по уровню приоритетности.

**Приоритетные –**­ требования, имеющие предпочтение по уровню важности перед другими.

*По количеству каналов обслуживания СМО подразделяются на следующие группы*.

**Одноканальные СМО**. Она состоит из одной очереди и одного устройства обслуживания. Термин "одноканальная" говорит о том, что к устройству обслуживания ведет только один путь.

**Многоканальные СМО**. Обслуживание очередной заявки может начаться до окончания обслуживания предыдущей заявки. Каждый канал действует как самостоятельное обслуживающее устройство.

Если приборы обслуживания соединяются параллельно, то такое обслуживание называется однофазным, а если приборы соединяются последовательно, то многофазным, (ряд последовательных операций).

**Однофазные СМО** – это однородные системы, которые выполняют одну и ту же операцию обслуживания.

**Многофазные СМО** – это системы, в которых каналы обслуживания расположены последовательно и выполняют различные операции обслуживания. Примером многофазной СМО являются станции технического обслуживания автомобилей.

### 5. Основные характеристики СМО

Значение P0 определяет вероятность того, что все каналы обслуживания свободны (находятся в состоянии простоя). Значение *Pk* определяет вероятность того, что в системе (в очереди и на обслуживании) находятся *k* заявок. Если *k* не превышает числа каналов *N*, то все заявки находятся на обслуживании и очередь отсутствует; в противном случае все каналы заняты и *k-N* заявок находится в очереди.

Вероятность *Pотк* отказа в обслуживании определяется ситуацией занятости всех *N* каналов и всех *m* мест в очереди и равн а*PN+m*.

Среднее число занятых каналов *Nзан* определяется математическим ожиданием дискретной случайной величины

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_m1dd19ab4.png

Среднее число свободных каналов

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_49a7c3ad.png

Коэффициент простоя каналов

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_7c276f2c.png

Коэффициент занятости каналов

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_2e58d460.png

Относительная пропускная способность (доля обслуженных заявок в общем числе поступавших в систему) определяется величиной

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_45417835.png

Абсолютная пропускная способность (среднее число заявок, обслуживаемых в единицу времени) определяется величиной

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_m29c71cb0.png

Средняя длина очереди

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_m2b984570.png

Среднее число заявок, находящихся в системе, складывается из средних значений занятости каналов и длины очереди

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_m16e3721e.png

Среднее время пребывания заявки в очереди равно

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_m3f89934b.png

Общее время пребывания заявки в очереди будет складываться из T*очер* и среднего времени обслуживания

http://www.studfiles.ru/html/2706/672/html_JgkvhzkYy4.1B7P/htmlconvd-XZPW0D_html_26a811ad.png

Полученные характеристики дают возможность анализа замкнутых и разомкнутых систем с отказами (*m=0*), с очередью или с ожиданием (*m →∞*) при простейшем входном потоке и однотипных параллельных каналах обслуживания с показательным законом длительности обслуживания (в частности, с фиксированной длительностью).