**Основы теории массового обслуживания: Марковские случайные процессы**

При исследовании операций часто приходится сталкиваться с системами, предназначенными для многоразового использования при решении однотипных задач. Возникающие при этом процессы получили название *процессов обслуживания,*а  
системы – *систем массового обслуживания (СМО).*Примерами таких систем являются телефонные системы, ремонтные мастерские, вычислительные комплексы, билетные кассы, магазины, парикмахерские и т.п.

Каждая СМО состоит из определенного числа обслуживающих единиц (приборов, устройств, пунктов, станций), которые будем называть *каналами* обслуживания. Каналами могут быть линии связи, рабочие точки, вычислительные машины, продавцы и др. По числу каналов СМО подразделяют на *одноканальные* и *многоканальные.*

Заявки поступают в СМО обычно не регулярно, а случайно, образуя так называемый *случайный поток заявок (требований).* Обслуживание заявок также продолжается какое-то случайное время. Случайный характер потока заявок и времени обслуживания приводит к тому, что СМО оказывается загруженной неравномерно: в какие-то периоды времени скапливается очень большое количество заявок (они либо становятся в очередь, либо покидают СМО необслуженными), в другие же периоды СМО работает с недогрузкой или простаивает.

*Предметом теории массового обслуживания* является построение математических моделей, связывающих заданные условия работы СМО (число каналов, их производительность, характер потока заявок и т.п.) с показателями эффективности СМО, описывающими ее способность справляться с потоком заявок.

В качестве **показателей эффективности** СМО используются: среднее(здесь и в дальнейшем средние величины понимаются как математические ожидания соответствующих случайных величин) число заявок, обслуживаемых в единицу времени; среднее число заявок в очереди; среднее время ожидания обслуживания; вероятность отказа в обслуживании без ожидания; вероятность того, что число заявок в очереди превысит определенное значение и т.п.

**СМО делят на два основных типа (класса)**: [СМО с отказами](https://math.semestr.ru/cmo/cmo_otkaz.php) и СМО с ожиданием (очередью). В СМО с отказами заявка, поступившая в момент, когда все каналы заняты, получает отказ, покидает СМО и в дальнейшем процессе обслуживания не участвует (например, заявка на телефонный разговор в момент, когда все каналы заняты, получает отказ и покидает СМО необслуженной). В СМО с ожиданием заявка, пришедшая в момент, когда все каналы заняты, не уходит, а становится в очередь на обслуживание.  
СМО с ожиданием подразделяются на разные виды в зависимости от того, как организована очередь: с ограниченной или неограниченной длиной очереди, с ограниченным временем ожидания и т.п.

Процесс работы СМО представляет собой *случайный процесс.* Под *случайным (вероятностным или стохастическим) процессом* понимается процесс изменения во времени состояния какой-либо системы в соответствии с вероятностными закономерностями.

Процесс называется *процессом с дискретными состояниями,*если его возможные состояния S1, S2, S3… можно заранее перечислить, а переход системы из состояния в состояние происходит мгновенно (скачком). Процесс называется *процессом с непрерывным временем,*если моменты возможных переходов системы из состояния в состояние не фиксированы заранее, а случайны.

Процесс работы СМО представляет собой случайный процесс c дискретными состояниями и непрерывным временем. Это означает, что состояние СМО меняется скачком в случайные моменты появления каких-то событий (например, прихода новой заявки, окончания обслуживания и т.п.).

Математический анализ работы СМО существенно упрощается, если процесс этой работы – марковский. Случайный процесс называется *марковским* или *случайным процессом без последствия,*если для любого момента времени t0 вероятностные характеристики процесса в будущем зависят только от его состояния в данный момент t0 и не зависят от того, когда и как система пришла в это состояние.

Пример марковского процесса: система S – счетчик в такси. Состояние системы в момент t характеризуется числом километров (десятых долей километров), пройденных автомобилем до данного момента. Пусть в момент t0 счетчик показывает S0. Вероятность того, что в момент t > t0 счетчик покажет то или иное число километров (точнее, соответствующее число рублей) S1, зависит от S0, но не зависит от того, в какие моменты времени изменялись показания счетчика до момента t0.

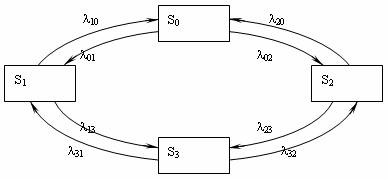
Многие процессы можно приближенно считать марковскими. Например, процесс игры в шахматы; система S*–*группа шахматных фигур. Состояние системы характеризуется числом фигур противника, сохранившихся на доске в момент t0. Вероятность того, что в момент t > t0 материальный перевес будет на стороне одного из противников, зависят в первую очередь от того, в каком состоянии находится система в данный момент t0, а не того, когда и в какой последовательности исчезли фигуры с доски до момента t0*.*

В ряде случаев предысторией рассматриваемых процессов можно просто пренебречь и применять для их изучения марковские модели.

При анализе случайных процессов с дискретными состояниями удобно пользоваться геометрической схемой – так называемым *графом состоянии.*Обычно состояния системы изображаются прямоугольниками (кружками), а возможные переходы из состояния в состояние – стрелками (ориентированными дугами), соединяющими состояния.

***Задача.*** Построить граф состояний следующего случайного процесса: устройство S состоит из двух узлов, каждый из которых в случайный момент времени может выйти из строя, после чего мгновенно начинается ремонт узла, продолжающийся заранее неизвестное случайное время.

***Решение.*** Возможные состояния системы: S0 – оба узла исправны; S1 – первый узел ремонтируется, второй исправен; S2 – второй узел ремонтируется, первый исправен;  
S3 – оба узла ремонтируются. Граф системы приведен на рис. 1.

  
Рисунок 1 – Граф системы

Стрелка, направленная, например, из S0 в S1 означает переход системы в момент отказа первого узла, из S1 в S0 – переход в момент окончанияремонта этого узла.  
На графе отсутствуют стрелки из S0, в S3 и из S1 в S2. Это объясняется тем, что выходы узлов из строя предполагаются независимыми друг от друга и, например, вероятностью одновременного выхода из строя двух узлов (переход из S0 в S3) или одновременного окончания ремонтов двух узлов (переход из S3 в S0) можно пренебречь.