**§1. Основные понятия теории марковских процессов**

В большинстве случаев не удается построить простую математическую модель, позволяющую в явном (аналитическом) виде найти интересующие нас величины (показатели эффективности) в зависимости от условий операции α и элементов решения х. Однако в некоторых особых случаях такую математическую модель удается построить. Это — когда исследуемая операция представляет собой так называемый Марковский случайный процесс.

Пусть имеется некоторая физическая система S, которая с течением времени меняет свое состояние (переходит из одного состояния в другое), причем заранее неизвестным, случайным образом. Тогда будем говорить, что в системе S протекает случайный процесс.

Под «физической системой» можно понимать что угодно: техническое устройство, группу таких устройств, предприятие, отрасль промышленности, живой организм, популяцию и т. д. Большинству процессов, протекающих в реальных системах, свойственны, в той или иной мере, черты случайности, неопределенности.

Например: система S — техническое устройство, состоящее из ряда узлов, которые время от времени выходят из строя, заменяются или восстанавливаются. Процесс, протекающий в этой системе, безусловно, случаен.

Случайный процесс, протекающий в системе, называется Марковским, если для любого момента времени t0 вероятностные характеристики процесса в будущем зависят только от его состояния в данный момент t0 u не зависят от того, когда и как система пришла в это состояние.

Пусть в настоящий момент t0 (см. рисунок 14.1) система находится в определенном состоянии S0.

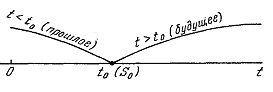


Рисунок 14.1

Мы наблюдаем процесс со стороны и в момент t0 знаем состояние системы S0 и всю предысторию процесса, все, что было при t < t0. Нас интересует будущее (t > t0). Так как в точности его предугадать мы не можем - наш процесс случайный, значит — непредсказуемый. Но какие-то вероятностные характеристики процесса в будущем мы найти можем. Например, вероятность того, что через некоторое время τ система S окажется в состоянии S1 или сохранит состояние S0, и т. п.

Если процесс — Марковский, то предсказывать можно, только учитывая настоящее состояние системы S0 и забыв о его «предыстории» (поведении системы при t < t0). Само состояние S0, разумеется, зависит от прошлого, но как только оно достигнуто, о прошлом можно забыть. Иначе формулируя, в Марковском процессе «будущее зависит от прошлого только через настоящее».

В исследовании операций большое значение имеют так называемые Марковские случайные процессы с дискретными состояниями и непрерывным временем. Процесс называется процессом с дискретными состояниями, если его возможные состояния S1, S2, S3,., . можно заранее перечислить (перенумеровать), и переход системы из состояния в состояние происходит «скачком», практически мгновенно. Процесс называется процессом с непрерывным временем, если моменты возможных переходов из состояния в состояние не фиксированы заранее, а неопределенны, случайны, если переход может осуществиться, в принципе, в любой момент. Мы будем рассматривать только процессы с дискретными состояниями и непрерывным временем.

Пример такого процесса: техническое устройство S состоит из двух узлов, каждый из которых в случайный момент времени может выйти из строя (отказать), после чего мгновенно начинается ремонт узла, тоже продолжающийся заранее неизвестное, случайное время. Возможные состояния системы можно перечислить:

So — оба узла исправны,

S1 — первый узел ремонтируется, второй исправен,

S2 — второй узел ремонтируется, первый исправен,

S3 — оба узла ремонтируются.

Переходы системы S из состояния в состояние происходят практически мгновенно, в случайные моменты выхода из строя того или другого узла или окончания ремонта.

При анализе случайных процессов с дискретными состояниями удобно пользоваться геометрической схемой — так называемым графом состояний.

Состояния системы изображаются прямоугольниками, а возможные переходы из состояния в состояние — стрелками, соединяющими состояния. Мы будем изображать состояния прямоугольниками, в которых записаны обозначения состояний: S1, S2, ...., Sn.

Построим граф состояний для рассмотренного выше примера (см. рисунок 14.2).

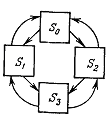


Рисунок 14.2

Стрелка, направленная из S0 в S1, означает переход в момент отказа первого узла; стрелка, направленная обратно, из S1 в S0,— переход в момент окончания ремонта этого узла. Остальные стрелки объясняются аналогично.

Если процесс, протекающий в системе с дискретными состояниями и непрерывным временем, является Марковским, то для его описания можно построить довольно простую математическую модель.

**§2. Потоки событий**

Потоком событий называется последовательность однородных событий, следующих одно за другим в какие-то случайные моменты времени. Например: поток вызовов на телефонной станции; поток отказов (сбоев) ЭВМ; поток железнодорожных составов, поступающих на сортировочную станцию; поток частиц, попадающих на счетчик Гейгера, и т.д.

Поток событий можно наглядно изобразить рядом точек на оси времени Ot (рисунок 15.1); положение каждой из них случайно, и на рисунке 15.1 изображена только какая-то одна реализация потока.



Рисунок 15.1

Важной характеристикой потока событий является его интенсивность λ — среднее число событий, приходящееся на единицу времени. Интенсивность потока может быть как постоянной (λ = const), так и переменной, зависящей от времени t. Например, поток автомашин, движущихся по улице, днем интенсивнее, чем ночью, в часы пик — интенсивнее, чем в другие часы.

Поток событий называется регулярным, если события следуют одно за другим через определенные, равные промежутки времени. На практике чаще встречаются потоки не регулярные, со случайными интервалами.

Поток событий называется стационарным, если его вероятностные характеристики не зависят от времени. В частности, интенсивность λ стационарного потока должна быть постоянной.

На практике часто встречаются потоки событий, которые (по крайней мере, на ограниченном участке времени) могут считаться стационарными. Например, поток вызовов, поступающих на АТС между 13 и 14 часами, практически стационарен; тот же поток в течение суток уже не стационарен.

Поток событий называется потоком без последействия, если для любых двух непересекающихся участков времени τ1 и τ2 (см. рисунок 15.2) число событий, попадающих на один из них, не зависит от того, сколько событий попало на другой.



Рисунок 15.2

Это означает, что события, образующие поток, появляются в те или другие моменты времени независимо друг от друга, вызванные каждое своими собственными причинами. Например, поток пассажиров, входящих в метро, практически не имеет последействия. А вот поток покупателей, отходящих от прилавка с купленными товарами, уже имеет последействие (т.к. интервал времени между отдельными покупателями не может быть меньше, чем минимальное время t0 обслуживания каждого из них).

Поток событий называется ординарным, если события в нем появляются поодиночке, а не группами по несколько сразу. Например, поток клиентов, направляющихся в парикмахерскую или к зубному врачу, обычно ординарен, чего нельзя сказать о потоке клиентов, направляющихся в загс для регистрации брака. Поток поездов, подходящих к станции, ординарен, а поток вагонов — неординарен. Если поток событий ординарен, то вероятностью попадания на малый участок времени Δt двух или более событий можно пренебречь.

Поток событий называется простейшим (или стационарным пуассоновским), если он обладает сразу тремя свойствами: стационарен, ординарен и не имеет последействия. Название «простейший» связано с тем, что процессы, связанные с простейшими потоками, имеют наиболее простое математическое описание.

**§3. Уравнения Колмогорова для вероятностей состояний**

Рассматривая Марковские процессы с дискретными состояниями и непрерывным временем, представим, что все переходы системы S из состояния в состояние происходят под действием каких-то потоков событий. Если все потоки событий, переводящие систему S из состояния в состояние,— простейшие, то процесс, протекающий в системе, будет Марковским.

Если система S находится в каком-то состоянии Si, из которого есть непосредственный переход в другое состояние Sj (стрелка, ведущая из Si, в Sj на графе состояний), то представим что на систему, пока она находится в состоянии Si, действует простейший поток событий, переводящий ее по стрелке Si->Sj. Как только появится первое событие этого потока, происходит «перескок» системы из Si в Sj.

Для наглядности на графе состояний у каждой стрелки проставим интенсивность того потока событий, который переводит систему по данной стрелке. Обозначим λ0 интенсивность потока событий, переводящего систему из состояния Si в Sj. На рисунке 16.1 дан граф состояний с проставленными у стрелок интенсивностями (будем называть такой граф размеченным).

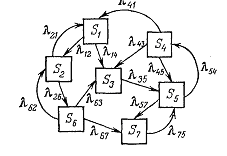


Рисунок 16.1

Построим размеченный граф состояний для технического устройства из двух узлов. Напомним состояния системы:

S0 — оба узла исправны,

S1 — первый узел ремонтируется, второй исправен,

S2 — второй узел ремонтируется, первый исправен,

S3 — оба узла ремонтируются.

Интенсивности потоков событий, переводящих систему из состояния в состояние, будем вычислять, предполагая, что среднее время ремонта узла не зависит от того, ремонтируется ли один узел или оба сразу.

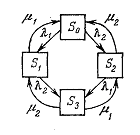


Рисунок 16.2

Найдем все интенсивности потоков событий, переводящих систему из состояния в состояние. Пусть система находится в состоянии S0. Поток отказов первого узла переводит ее в состояние S1. Его интенсивность λ1 равна единице, деленной на среднее время безотказной работы первого узла. Поток «окончаний ремонтов» первого узла переводит систему обратно из S1 в S0. Его интенсивность μ1 равна единице, деленной на среднее время ремонта первого узла. Аналогично вычисляются интенсивности потоков событий, переводящих систему по всем стрелкам графа рисунка 16.2.

Имея в своем распоряжении размеченный граф состояний системы, построим математическую модель данного процесса.

Пусть рассматривается система S, имеющая n возможных состояний S1, S2, ..., Sn. Назовем вероятностью i-го состояния вероятность pi(t) того, что в момент t система будет находиться в состоянии Si. Очевидно, что для любого момента сумма всех вероятностей состояний равна единице:

n

Σ pi (t) = 1 (16.1)

i = 1

Имея в своем распоряжении размеченный граф состояний, можно найти все вероятности состояний pi(t) как функции времени. Для этого составляются и решаются так называемые уравнения Колмогорова — особого вида дифференциальные уравнения, в которых неизвестными функциями являются вероятности состояний.

Пусть система S имеет четыре состояния: S1, S2, S3, S4, размеченный граф которых показан на рисунке 16.3.

Рассмотрим одну из вероятностей состояний, например p1(t). Это—вероятность того, что в момент t система будет в состоянии S1. Придадим t малое приращение Δt и найдем p1 (t + Δt) — вероятность того, что в момент t + Δt система будет в состоянии S1.

Это может произойти двумя способами:

1. в момент t система уже была в состоянии S1, а за время Δt не вышла из него;
2. в момент t система была в состоянии S2, а за время Δt перешла из него в S1.

Найдем вероятность первого варианта. Вероятность того, что в момент t система была в состоянии S1, равна p1(t). Эту вероятность нужно умножить на вероятность того, что, находившись в момент t в состоянии S1, система за время Δt не перейдет из него ни в S2, ни в S3. Суммарный поток событий, выводящий систему из состояния S1, тоже будет простейшим, с интенсивностью λ12+λ13. Значит, вероятность того, что за время Δt система выйдет из состояния S1, равна (λ12 + λ13)Δt; вероятность того, что не выйдет:

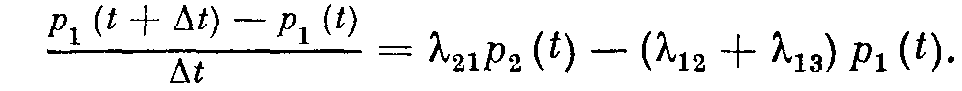
1 — (λ12 + λ13)Δt. Отсюда вероятность первого варианта равна p1(t) [1 — (λ12 + λ13)Δt].

Найдем вероятность второго варианта. Она равна вероятности того, что в момент t система будет в состоянии S2, а за время Δt перейдет из него в состояние S1, т. е. она равна p2(t)λ21Δt.

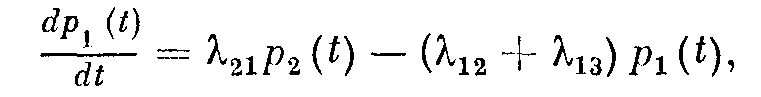
Складывая вероятности обоих вариантов (по правилу сложения вероятностей), получим:

p1(t +Δt ) = p1(t) [1 — (λ12 + λ13)Δt] + p2(t)λ21Δt .

Раскроем квадратные скобки, перенесем p1(t) в левую часть и разделим обе части на Δt:



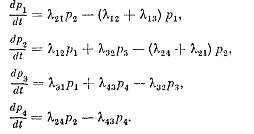
Устремим Δt к нулю; слева получим в пределе производную функции p1(t). Таким образом, запишем дифференциальное уравнение для p1(t):



или, короче, отбрасывая аргумент t у функций p1, p2 :



Рассуждая аналогично для всех остальных состоянии, напишем еще три дифференциальных уравнения. Присоединяя к ним уравнение (16.2), получим систему дифференциальных уравнений для вероятностей состояний:



(16.3)

Это — система четырех линейных дифференциальных уравнений с четырьмя неизвестными функциями p1 p2, p3, p4. Одно из них (любое) можно отбросить, пользуясь тем, что p1 + p2 + p3 + p4 = 1, т.е. выразить любую из вероятностей pi через другие, это выражение подставить в (16.3), а соответствующее уравнение с производной dpi/dt отбросить.

Сформулируем общее правило составления уравнений Колмогорова. В левой части каждого из них стоит производная вероятности какого-то (i-го) состояния. В правой части — сумма произведений вероятностей всех состояний, из которых идут стрелки в данное состояние, на интенсивности соответствующих потоков событий, минус суммарная интенсивность всех потоков, выводящих систему из данного состояния, умноженная на вероятность данного (i-ro) состояния.

Пользуясь этим правилом, запишем уравнения Колмогорова для системы S, размеченный граф состояний которой дан на рисунке 16.2:

dp0

---- = μ1p1 + μ2p2 – (λ1 + λ2) p0

dt

dp1

---- = λ1p0 + μ2p3 – (λ2 + μ1) p1

dt (16.4)

dp2

---- = λ2p0 + μ1p3 – (λ1 + μ2) p2

dt

dp3

---- = λ2p1 + λ1p2 – (μ1 + μ2) p3

dt

Чтобы решить уравнения Колмогорова и найти вероятности состояний, прежде всего надо задать начальные условия. Если мы точно знаем начальное состояние системы Si, то в начальный момент (при t = 0) р0(0) = 1, а все остальные начальные вероятности равны нулю. Так, например, уравнения (17.4) естественно решать при начальных условиях р0(0) = 1, p1(0) = р2(0) = р3(0) = 0 (в начальный момент оба узла исправны).

Уравнения Колмогорова дают возможность найти все вероятности состояний как функции времени.

**§4. Финальные вероятности состояний**

Если при t 🡪 ∞ вероятности состояний p1(t), p2(t),... будут стремиться к каким-то пределам и если эти пределы существуют и не зависят от начального состояния системы, то они называются финальными вероятностями состояний. В теории случайных процессов доказывается, что если число n состояний системы конечно и из каждого из них можно (за конечное число шагов) перейти в любое другое, то финальные вероятности существуют.

Предположим, что это условие выполнено и финальные вероятности существуют:

lim pi (t) = pi (i = 1, 2, . . . , n) (17.1)

t -> ∞

Финальные вероятности будем обозначать теми же буквами p1, р2, ..., что и сами вероятности состояний, но понимая под ними уже не переменные величины (функции времени), а постоянные числа. Очевидно, они тоже образуют в сумме единицу:

n

Σ pi (t) = 1 (17.2)

i = 1

При t 🡪 ∞ в системе S устанавливается предельный стационарный режим, в ходе которого система случайным образом меняет свои состояния, но их вероятности уже не зависят от времени. Финальную вероятность состояния Si можно рассматривать как среднее относительное время пребывания системы в этом состоянии. Например, если система S имеет три состояния S1, S2, S3 и их финальные вероятности равны 0,2, 0,3 и 0,5, это значит, что в предельном, стационарном режиме система в среднем две десятых времени проводит в состоянии S1, три десятых—в состоянии S2 и половину времени — в состоянии S3.

Если вероятности p1, p2, ... постоянны, то их производные равны нулю. Значит, чтобы найти финальные вероятности, нужно все левые части в уравнениях Колмогорова положить равными нулю и решить полученную систему уже не дифференциальных, а линейных алгебраических уравнений. Если перенести отрицательный член каждого уравнения из правой части в левую, то получим сразу систему уравнений, где слева стоит финальная вероятность данного состояния рi, умноженная на суммарную интенсивность всех потоков, ведущих из данного состояния, а справа — сумма произведений интенсивностей всех потоков, входящих в i-е состояние, на вероятности тех состояний, из которых эти потоки исходят.

Пользуясь этим правилом, напишем линейные алгебраические уравнения для финальных вероятностей состояний системы, граф состояний которой дан на рисунке 16.2:

(λ1 + λ2) p0 = μ1p1 + μ2p2

(λ2 + μ1) p1 = λ1p0 + μ2p3

(λ1 + μ2) p2 = λ2p0 + μ1p3 (17.3)

(μ1 + μ2) p3 = λ2p1 + λ1p2

При решении этой системы четырех уравнений с четырьмя неизвестными р0, р1, р2, р3, воспользуемся так называемым нормировочным условием:

p0 + p1 + p2 + p3 = 1 (17.4)

и с его помощью решить систему. При этом одно (любое) из уравнений можно отбросить (оно вытекает как следствие из остальных).

Зададимся численными значениями интенсивностей λ1 = 1, λ2=2, μ1 = 2, μ2 = 3 и решим систему (17.3). Вместо четвертого уравнения добавим нормировочное условие (17.4). Уравнения примут вид:

3p0 = 2p1 + 3p2

4p1 = p0 + 3p3

4p2 = 2p0 + 3p3 (17.5)

p0 + p1 + p2 + p3 = 1

Решая их, получим:

p0 = 6/15=0,40; p1 =3/15 =0,20; p2 = 4/15 = 0,27; P3=2/15 =0,13,

т. е. в предельном, стационарном режиме система S в среднем 40% времени будет проводить в состоянии S0 (оба узла исправны), 20% — в состоянии S1 (первый узел ремонтируется, второй работает), 27% —в состоянии S2 (второй узел ремонтируется, первый работает) и 13% — в состоянии S3 полной негодности (оба узла ремонтируются). Знание этих финальных вероятностей может помочь оценить среднюю эффективность работы системы и загрузку ремонтных органов. Предположим, что система S в состоянии S0 (полностью исправная) приносит в единицу времени доход 8 (условных единиц), в состоянии S1—доход 3, в состоянии S2—доход 5, в состоянии S3 — вообще не приносит дохода. Тогда в предельном, стационарном режиме средний доход в единицу времени будет

W = 0,40 • 8 + 0,20 • 3 + 0,27 • 5 = 5,15.

Теперь оценим загрузку рабочих, занятых ремонтом узлов 1 и 2. Узел 1 ремонтируется долю времени, равную p1 + р3 = 0,20 + 0,13 = 0,33. Узел 2 ремонтируется долю времени p2 + р3 = 0,40.