**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ**

кафедра сетей и устройств телекоммуникаций

**РЕФЕРАТ**

**На тему:**

**«Модели систем массового обслуживания. Классификация систем массового обслуживания»**

**МИНСК, 2008**

**Математическое введение в теорию цепей Маркова. (Markov’s chain )**

Дискретные цепи Маркова. Будем говорить, что задана дискретная цепь Маркова, если для последовательности случайных величин выполняется равенство .



Это означает, что поток случайных величин определяется только вероятностью перехода от предыдущего значения случайной величины к последующему. Зная начальное распределение вероятностей, можно найти распределение на любом шаге. Величины *in*можно интерпретировать как номера состояний некоторой динамической системы с дискретным множеством состояний (типа конечного автомата). Если вероятности переходов не зависят от номера шага, то такая цепь Маркова называется однородной и ее определение задается набором вероятностей .



Для однородной Марковской цепи можно определить вероятности перехода из состояния *i* в состояние *j* за *m* шагов



Цепь Маркова называется неприводимой, если каждое ее состояние может быть достигнуто из любого другого состояния. Состояние *i* называется поглощающим, если для него *pii* =1.

Состояние называется возвратным, если вероятность попадания в него за конечное число шагов равна единице. В другом случае состояние относится к невозвратным. Возвратное состояние может быть периодическим и апериодическим в зависимости от наличия кратных шагов возврата. Введем вероятности возврата в состояние *i* через *n* шагов после ухода из этого состояния:



Они позволяют определить среднее число шагов или, иначе говоря, среднее время возврата:.



Состояние называется **возвратным нулевым**, если среднее время возвращения в него равно бесконечности, и **возвратным ненулевым**, если это время конечно. Известны две важные теоремы:

Теорема 1.

Состояния неприводимой цепи Маркова либо все невозвратные, либо все возвратные нулевые, либо все возвратные ненулевые. В случае периодической цепи все состояния имеют один и тот же период.

Вторая теорема рассматривает вероятности достижения состояний в стационарном (то есть не зависящем от начального распределения вероятностей) режиме. Соответствующее распределение вероятностей также называют стационарным. Нахождение стационарного распределения вероятностей достижения состояний одна из основных задач теории телетрафика.

Теорема 2.

Для неприводимой и апериодической цепи Маркова всегда существуют предельные вероятности, не зависящие от начального распределения вероятностей. Более того, имеет место одна из следующих двух возможностей:

А) все состояния цепи невозвратные или все возвратные нулевые, и тогда все предельные вероятности равны нулю и стационарного состояния не существует;

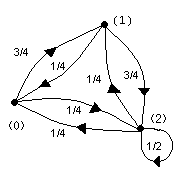
Б) все состояния возвратные ненулевые и тогда существует стационарное распределение вероятностей:



Состояние называется **эргодическим**, если оно апериодично и возвратно ненулевое. Если все состояния цепи Маркова ***эргодичны***, то вся цепь называется ***эргодической***. Предельные вероятности эргодической цепи Маркова называют вероятностями состояния равновесия, имея в виду, что зависимость от начального распределения вероятностей полностью отсутствует.

Цепь Маркова с конечным числом состояний (конечная цепь), удобно изображать в виде ориентированного графа, называемого диаграммой переходов. Вершины графа ассоциируются с состояниями, а ребра с вероятностями переходов.

Вычисления вероятностей достижения состояний производится прямыми методами или с помощью z-преобразования.



Цепь Маркова.

Введем матрицу вероятностей переходов и вектор-строку вероятностей на шаге n

.



Распределение вероятностей на произвольном шаге тогда будет подчиняться матричному соотношению:

.



Оно позволяет рекуррентно вычислять все вероятности состояний. Для нахождения предельного распределения (стационарного) нужно решить уравнение:



Его можно решать как систему линейных алгебраических уравнений, если цепь конечна.

Для примера (рис. 1) имеем:

.



и решение матричного уравнения сводится к решению системы трёх уравнений:



Коэффициенты первого уравнения в этой системе дополняют до единицы сумму коэффициентов второго и третьего уравнений; это свидетельствует о линейной зависимости между ними. Поэтому для решения системы уравнений нужно ввести дополнительное нормирующее условие. В данном примере: .



Решая систему полученных уравнений, имеем:



Уравнение для вероятности достижения состояния в переходном режиме решить значительно труднее. Некоторого упрощения можно достигнуть, используя z – преобразование. Применим его к уравнению для переходных вероятностей

.



Обозначая соответствующие преобразования, получим:



Все полученные здесь математические результаты относились к однородным Марковским процессам, где вероятности переходов не зависят от времени. В более общем случае такая зависимость имеет место.

Рассмотрим вероятности перехода системы из состояния i на m-том шаге в состояние j на n-том шаге для n > m.

Можно показать, что эти вероятности связаны между собой, так называемым уравнениями Чепмена-Колмогорова.(Chapman - Kolmogorov)

.



Для однородных цепей Маркова эти уравнения упрощаются так как

.



И сводятся к анализируемым выше.

Непрерывные цепи Маркова.

Случайный процесс X(t) с дискретным множеством значений образует непрерывную цепь Маркова, если

.



Будущие состояния зависят от прошлого только через текущее состояние. Для непрерывный цепей Маркова основным также является уравнение Чепмена –Колмогорова, для однородной цепи имеющее вид: .



Здесь матрица **H**(t)= [ pij(t)] - матрица вероятностей перехода из состояния i в состояние j в момент времени t , а матрица **Q** называется матрицей интенсивностей переходов. Ее элементы имеют следующий смысл: если в момент времени t система находилась в состоянии Ei , то вероятность перехода в течение промежутка времени (t,t+Δt) в произвольное состояние Ej задается величиной qij(t)Δt + o(Δt), а вероятность ухода из состояния Ei величиной -qiiΔt + o(Δt).

Таким образом, интенсивности переходов можно вычислять как соответствующие пределы при стремлении к нулю длительности временного интервала.

Наиболее важным для дальнейшего использования является класс непрерывных цепей Маркова называемых «процессами гибели - размножения» **(**Birth – death process**).** Для таких систем из состояния k возможны переходы только в состояния k, k-1 и k+1 в следующие моменты времени:

* в момент t объем популяции был равен k и в течение времени (t,t+Δt) не произошло изменения состояния
* в момент t объем популяции был равен k-1 и в течение времени (t,t+Δt) родился один член популяции
* в момент времени t объем популяции был равен k+1 и в течение времени (t,t+Δt) погиб один член популяции

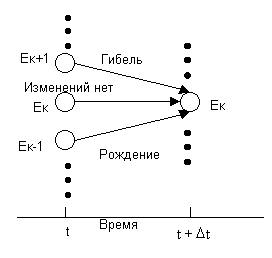


Рис. 1. Возможные переходы в состояние Ек.

Будем искать вероятность того, что в момент времени t объем популяции равен k , обозначив его Pk(t). Можно записать соотношения для вероятности достижения со­стояния k в момент времени t+Δt:

.



Определим граничные и нормирующие условия:



Выразим вероятности переходов за интервал Δt через интенсивности

Вер(+1)=λkΔt+o(Δt) ; Вер(-1)=μkΔt+o(Δt).

Вероятность нуля рождений 1- λkΔt+o(Δt) , а нуля гибелей 1- μkΔt+o(Δt).

Таким образом, вероятность того, что состояние k сохранится неизменным, будет равно произведению [1- λkΔt+o(Δt)][ 1- μkΔt+o(Δt)].

Тогда **уравнения Чепмена-Колмогорова** приобретают вид



Раскрывая скобки и проводя деление на Δt, получим:



В пределе получается система дифференциально-разностных уравнений, решение которой будут играть важную роль для практических задач.



В соответствие этой системе уравнений можно поставить наглядную диаграмму интенсивностей переходов, которая аналогична диаграмме переходов для дискретных цепей Маркова (Рис. 2)

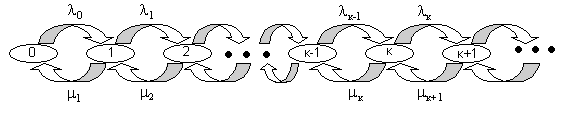


Рис. 2 Диаграмма интенсивностей переходов для процесса размножения и гибели.

Овалам здесь соответствуют дискретные состояния, а стрелки определяют интенсивности потоков вероятности (а не вероятности!) переходов от одного состояния к другому.

Имеет место своеобразный «**закон сохранения**»:

Разность между суммой интенсивностей, с которой система попадает в состояние k и суммой интенсивностей, с которой система покидает это состояние должна равняться интенсивности изменения потока в это состояние (производной по времени).

Применение закона сохранения позволяет получать уравнения для любой подсистемы Марковской цепи типа процесса «гибели-размножения». Особенно эффективным оказывается построение решений в стационарном, установившемся режиме, когда можно полагать что вероятности в произвольный, достаточно отдаленный момент времени, остаются постоянными.

Приравнивая производную по времени нулю, получаем систему разностных уравнений



Полагая, что интенсивности λ-1 =λ-2 = λ-3 =…0; μ0 = μ-1 = μ-2 = μ-3 =…=0, второе уравнение выписывать отдельно далее не потребуется. Итак, стационарный режим в цепи Маркова будет описываться системой разностных уравнений и условием нормировки для вероятностей



Нетрудно видеть, что эти уравнения легко выводятся из закона сохранения интенсивностей вероятностей. В стационарном режиме разность потоков равна нулю и полученные выше уравнения приобретают смысл уравнений равновесия или баланса, как их и называют.

.



Интенсивность потока вероятностей **в** состояние k равна интенсивности потока из этого состояния.

Решать уравнение баланса можно, сначала определив при k =0 значение

.



Затем, построив систему уравнений для k =1, можно получить .



Далее получаем



Из условия нормировки: .



Система, описываемая полученными выше выражениями, будет иметь стационарные вероятности состояний, когда она эргодическая. Это условие может быть выражено через соотношение интенсивностей. Необходимо и достаточно, чтобы существовало некоторое значение k , начиная с которого выполнялось неравенство

.



Для большинства реальных систем массового обслуживания это неравенство выполняется.

**Классификация систем массового обслуживания.**

Используется трех -, четырех -, шести – компонентное символическое обозначение системы массового обслуживания, предложенное Кендаллом (Candall) и развитое в работах Г.П.Барашина.

a/b/c :d/e/f

a – распределение поступающего потока запросов.

b – закон распределения времени обслуживания.

Типовые условные обозначения:

М – экспоненциальное (Марковское) распределение,

D – детерминированное распределение,

Ek – эрланговское распределение k-го порядка,

HMk – гиперэкспоненциальное,

HEk – гиперэрланговское распределение порядка k,

GI – произвольное распределение независимых промежутков между заявками,

G – произвольное распределение длительностей обслуживания.

c – структура системы обслуживания (обычно число серверов).

d – дисциплина обслуживания (параметры после двоеточия иногда опускают).

Обычно используется сокращенное символическое обозначение, например FF вместо FIFO, LF, PR и т.п.

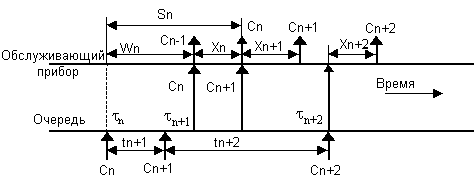
e – максимальное число запросов, воспринимаемое системой, может употребляться символ ∞.

f – максимальное число запросов к системе обслуживания.

В некоторых публикациях последними символами отражают качественные характеристики системы обслуживания. Некоторые общие результаты и основы математического аппарата, необходимого для анализа можно получить, рассматривая системы G/G/m.

Формула Литтла (Little).

Рассмотрим временную диаграмму работы системы массового обслуживания (рис. 3), отразив на ней последовательность поступления требований, помещение требований в очередь и обработки серверами системы.



Временная диаграмма работы системы массового обслуживания.

В общем случае ясно, что с увеличением числа требований растет время ожидания. Установим соотношение между средним числом требований в системе, интенсивностью потока и среднего времени пребывания в системе. Обозначим число поступающих в промежутке времени *(0 , t)* требований как функцию времени ***α(t).***

Число исходящих из системы заявок (обслуженных) на этом интервале обозначим ***δ(t).*** На рисунке 4 показаны примеры функциональных зависимостей этих двух случайных процессов от времени.

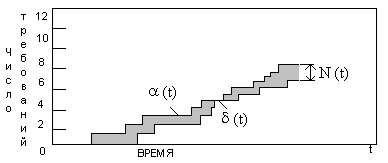


Рис. 4 Зависимость между средним числом требований в системе, интенсивностью потока и средним времени пребывания в системе.

Число требований, находящихся в системе в момент t будет равно:

.



Площадь между двумя рассматриваемыми кривыми от 0 до t - дает общее время, проведенное всеми заявками в системе за время t.

Обозначим эту накопленную величину γ(t) . Если интенсивность входного потока равна λ, а средняя интенсивность за время t: ,то время, проведенное одной заявкой в системе, усредненное по всем заявкам будет равно:



.



Наконец, определим среднее число требований в системе в промежутке (0,t): .



Из последних трех уравнений следует, что: , (где ).



Если в СМО существует стационарный режим, то при t→ ∞ , будут иметь место соотношения:



Последнее соотношение означает, что среднее число заявок в системе равно произведению интенсивности поступления требований в систему на среднее время пребывания в системе. При этом не накладывается никаких ограничений на распределения входного потока и времени обслуживания. Впервые доказательство этого факта дал Дж.Литтл и это соотношение носит название **формула Литтла.**

Интересно, что в качестве СМО можно рассмотреть только очередь из заявок в буфере. Тогда формула Литтла приобретает иной смысл - средняя длина очереди равна произведению интенсивности входного потока заявок на среднее время ожидания в очереди: .



Если наоборот рассматривать СМО только как серверы, то формула Литтла дает:

,



где – среднее число заявок в серверах, а– среднее время обработки в сервере.



В любом случае: .



Одним из основных параметров, которые используются при описании СМО, является **коэффициент использования (utilization factor)**. Это фундаментальный параметр, так как он определяется как отношение интенсивности входного потока к пропускной способности системы. Поскольку пропускная способность СМО содержащей m серверов может быть определена как: , то коэффициент использования может быть определен как:



.



Нетрудно видеть, что коэффициент использования равен в точности интенсивности нагрузки, если СМО с одним сервером и в m раз меньше для систем с m серверами. Величина коэффициента использования равна среднему значению от доли занятых серверов и .



Если в СМО типа **G/G/1** существует стационарный режим и можно определить вероятность того, что в некоторый случайный момент сервер будет свободный, то

.



**ЛИТЕРАТУРА**

1. Л.Н. Волков, М.С. Немировский, Ю.С. Шинаков. Системы цифровой радиосвязи: базовые методы и характеристики. Учебное пособие.-М.: Эко-трендз, 2005.
2. М.В. Гаранин, В.И. Журавлев, С.В. Кунегин. Системы и сети передачи информации. - М.: Радио и связь, 2001.
3. Передача дискретных сообщений./Под ред. В.П. Шувалова. – М.: Радио и связь, 1990.
4. Основы передачи дискретных сообщений./Под ред. В.М. Пушкина. – М.: Радио и связь, 1992.
5. Н.В. Захарченко, П.Я. Нудельман, В.Г. Кононович. Основы передачи дискретных сообщений. –М.: Радио и связь, 1990.