

# **1. МЕТОДЫ ПОСТРОЕНИЯ РАСПИСАНИЯ ПОТОКОВ ТРЕБОВАНИЙ В КОММУТАЦИОННЫХ УЗЛАХ**

## **1.1. Существующие методы построения расписания передачи потоков информации в коммутационных узлах сети**

С увеличением объемов передаваемой информации и числа объектов, между которыми она распределяется, роль технических средств распределения информации резко возрастает, что приводит к необходимости построения новых и оптимального использования существующих систем и сетей телекоммуникации.

Решение задач управления сложными современными техническими комплексами (разветвленные энергетические системы, большие промышленные предприятия, отраслевые системы и др.) обусловило появление понятия «большая система». Одна и та же система в зависимости от характера задач может рассматриваться как большая система и как классическая совокупность объекта и управляющего устройства. Любая большая система управления представляет собой комплекс алгоритмов, технических средств (в состав которых в общем случае входят технологические средства, выполняющие основную целевую функцию системы) и взаимосвязанных организационно-технических мероприятий, обеспечивающих эффективное достижение стоящей перед системой цели.

С помощью телекоммуникационной сети информация доводится до соответствующих объектов управления. Эта сеть состоит из взаимосвязанных комплексов технических средств, выполняющих определенные функции. К основным техническим средствам относятся:

а) сеть телекоммуникации, представляющая собой совокупность узлов и линий связи и характеризуемая надежностью, пропускной способностью, протяженностью линий связи и экономическими показателями;

б) средства передачи информации соответствующим объектам - узлам информационной сети и повышения ее достоверности;

в) средства управления потоками информации в сети телекоммуникации, осуществляющие упорядоченный обмен сообщениями между объектами системы с учетом состояния и загрузки каналов и узлов сети, важности и срочности передаваемой информации; одна из главных функций этих средств – выбор оптимальных относительно некоторого критерия путей передачи. Указанные технические средства являются важнейшей составной частью любой современной разветвленной информационной сети;

г) средства контроля и коммутации рабочих и резервных каналов связи, осуществляющие статистическую оценку состояния этих каналов относительно некоторого критерия с целью определения необходимости и момента перехода с неисправного рабочего канала связи на исправный, т.е. резервный и обратно.

При обмене информацией необходимо, обеспечение:

а) заданных вероятностно-временных характеристик доведения сообщений до потребителей, поскольку чрезмерные их задержки или потери могут привести к уменьшению эффективности управления;

б) достоверности передаваемых сообщений, поскольку использование искаженной информации может в отдельных случаях привести к большому материальному ущербу или другим нежелательным последствиям.

Одной из главных проблем, связанных с распределением и доставкой информации заданным потребителям, является рациональное использование сети телекоммуникации, объединяющей оконечное оборудование, каналы и устройства переработки информации с помощью аппаратуры, обеспечивающей ее распределение и управление сетью. Основными качественными показателями сети телекоммуникации являются надежность доставки информации по заданному адресу, скорость ее передачи и степень использования различных технических средств связи.

Процесс распределения потоков информации, поступающей от абонентов (объектов управления) в информационную сеть, происходит в коммутационных узлах и зависит от расположения сети. Для стационарных процессов это расположение может быть задано в виде матрицы расстояний между каждой

парой узлов  $\alpha = \|l_{ij}\|$  ( $ij = 1, 2, 3, \dots, N$ ). Структура (конфигурация) сети определяется наличием ветвей (магистралей) между отдельными узлами и их мощностью. Такая структура также может быть описана матрицей мощностей  $P = \|p_{ij}\|$  или графом  $P$ . Элементами  $p_{ij}$  матрицы мощностей может быть либо число каналов, идущих от узла  $i$ , к узлу  $j$ , либо общее число каналов между этими узлами  $p'_{ij}=p_{ji}=p_{ij}=p_{ji}$ . Соответствующий матрице граф, ориентированный при направленных каналах, имеет взвешенные ребра. К характеристикам сети относятся и стоимостные показатели ветвей и узлов, которые включаются в приведенные капитальные и эксплуатационные расходы. Приведенные расходы могут быть описаны матрицей  $C = \|c_{ij}\|$ , где  $c_{ij}$  — расходы по ветви между узлами  $i$  и  $j$ , а  $c_{ii} = c_i$  — расходы по узлу  $i$ .

Нагрузочные характеристики системы определяют структуру и объемы потоков информации для каждого направления обмена. Обычно они содержат общие характеристики информации, подлежащей передаче между объектами системы, и частные характеристики информации на каждом направлении обмена при ее передаче от узла  $i$  к узлу  $j$ , независимо от того, по какому пути происходит передача в реальной сети связи. Общее среднее число сообщений, поступающих в сеть в единицу времени от внешних источников, можно определить выражением  $A = \sum_{ij} \lambda_{ij}$ , где  $\lambda_{ij}$  - результирующая плотность потока (т. е. потока сообщений всех категорий срочности). Совокупность  $\lambda_{ij}$ , представленная в виде матрицы  $\|\lambda_{ij}\|$ , характеризует нагрузку информационной сети в целом.

Как известно, системы управления различных уровней и назначений создаются главным образом на базе уже существующих неавтоматизированных систем. Информационный обмен между абонентами сети телекоммуникации осуществляется благодаря использованию каналов связи. В случае отсутствия последних возникает необходимость создания (синтеза) сети связи, достаточной для передачи заданных потоков информации с требуемыми характеристиками.

Для случаев, когда в каждый момент времени в сети имеется более

одного двухполюсного потока, в настоящее время нет строгого метода решения задачи синтеза минимальной сети, но известны решения задач для случаев, когда в каждый момент времени сеть используется для передачи только одного (любого из заданных) двухполюсного потока. Можно предположить, что сеть состоит из конечного числа ветвей с ограниченными пропускными способностями каналов связи. Синтез сети связи можно описать с использованием систематического метода, основанного на последовательном разложении матрицы заключительных мощностей. Под матрицей заключительных мощностей  $\|p_{ij}\|$  понимается конечная мощность информации, передаваемой между любой парой вершин сети связи. Этот метод можно использовать только в том случае, когда матрица, характеризующая нагрузку информационной сети  $\|\lambda_{ij}\|$ , представима в виде матрицы заключительных мощностей  $\|p_{ij}\|$ , что на практике не всегда осуществимо.

Задачи синтеза сети связи с минимальной стоимостью можно решить методом линейного программирования. Более полный, но сложный подход к синтезу сетей по матрице  $\|\lambda_{ij}\|$  можно осуществлять синтезом сети относительно критерия минимума задержки сообщения при заданном ограничении по стоимости информационной сети и фиксированной процедуре выбора пути. Для различных конфигураций сетей данная задача решается на специализированных компьютерах, методом программного моделирования.

Методы синтеза информационной сети по матрице ее нагрузки не дают конструктивных рекомендаций по выбору вариантов распределения потоков для каждого возможного состояния сети. В то же время оценка достаточности сети для пропускания многополюсных потоков требует предварительного решения задачи, непосредственно связанной с управлением потоками в сети при динамическом изменении ее состояния.

Методы распределения потоков информации в коммутационных узлах делятся на динамические и программные. В свою очередь, по способу получения информации о состоянии сети связи динамические методы распределения потоков

подразделяются на детерминированные, когда информацию о состоянии сети получают в результате непосредственного контроля каналов и оборудования сети, устройствами контроля, и стохастические, когда информацию о состоянии сети получают в результате анализа данных о длине пройденного пути, времени передачи, потерях, установленных по результатам обслуживания переданных ранее сообщений.

Широкое распространение получили методы распределения потоков сообщений по кратчайшим путям. Они позволяют определять пути, не только кратчайшие по длине, но и с минимальным числом транзитов, с минимальной стоимостью, с минимальными помехами и т. п., т. е. в качестве оценки распределения потоков информации можно применять различные критерии. Чаще всего для распределения потоков по сети используются пути, кратчайшие по числу транзитных узлов.

Можно устанавливать кратчайший путь от произвольного узла  $i$  к фиксированному узлу  $N$  по числу транзитов (когда длины всех ветвей равны единице). Этот метод может быть также использован для определения путей, кратчайших по длине.

Существуют методы позволяющие определять кратчайший путь с помощью нумерации узлов и ветвей. Процесс нумерации ветвей (приписывание весов) называется формированием рельефа фиксированного узла  $N$ . В этом случае рельефы сети записываются в специальные таблицы рельефов, и на каждом узле формируется своя таблица. Для определения длины кратчайших путей методом нумерации узлов предлагается также поисковый способ распределения потоков.

Кратчайшие пути можно рассчитывать и с помощью матричных методов, где предусматривается определение длины кратчайших путей между всеми узлами сети возведением в степень  $P \leq (n - 1)$  матрицы длин ветвей  $\alpha$ , скорректированной с учетом поступающей от устройства контроля информации об исправности и занятости, ветвей сети. Такая операция рассматривается как последовательное  $(P-1)$ -кратное умножение матрицы  $\alpha$  самой на себя, т.е.

$\alpha^p = \alpha * \alpha^{p-1}$ , в результате можно получить матрицу длин кратчайших путей.

Для нахождения оптимального распределения потоков в сетях связи авторы работы [14] предлагают определять матрицы путей маршрутов на ЭВМ. Они приводят алгоритм определения всех возможных путей от некоторого узла ко всем остальным узлам сети и оценивают машинное время, необходимое для решения рассматриваемой задачи, а также объем оперативной памяти машин. Аналогичная задача для 10—20 узлов решена методом линейного программирования за несколько десятков минут машинного времени. Здесь сформулированы необходимые и достаточные условия передачи потоков информации по сети заданной системы с помощью матрицы путей, а также понятия покрытия совокупности путей сети, описаны методы оптимального распределения каналов в коммутируемых и некоммутируемых сетях связи.

Широко распространен децентрализованный метод распределения потоков в коммутируемых сетях по длине пути. Такой способ наиболее целесообразно использовать на однородных коммутируемых сетях с абонентами одной категории, поскольку учет категоричности абонентов и неоднородности сети усложняет процесс распределения потоков.

В информационных сетях для организации передачи потоков, информации применяются несколько методов пространственной коммутации. Основными из которых являются: а) метод коммутации каналов, при котором необходима предварительная установка соединений между двумя окончными абонентами, как на обычных автоматических телефонных, сетях; б) метод коммутаций сообщений или пакетов, при котором не нужна, предварительная установка соединений между абонентами, обменивающимися информацией; в этом случае сообщение, поступившее от абонента-отправителя на коммутационный узел, становится в очередь и обслуживается в соответствии с присвоенным ему приоритетом.

Эффективность использования сети в часы наибольшей нагрузки при реализации первого метода очень малая, а при реализации второго довольно высокая. Метод коммутации сообщений или пакетов имеет и другие

преимущества перед методом коммутации каналов:

1) возможность свободного выбора скорости и способа передачи сообщений от объектов-отправителей независимо от скорости и способа работы технических средств объектов-получателей;

2) лучшее использование центральной части сети вследствие более эффективного распределения трафика;

3) возможность передачи циркулярных и многоадресных сообщений;

4) гарантия определенного качества системы защиты от ошибок, входящей в состав системы коммутации сообщений;

5) допустимость использования при передаче информации между двумя абонентами каналов связи различной производительности;

6) обслуживание абонентов различных категорий даже в часы наибольшей нагрузки со значительно меньшей вероятностью отказов;

7) возможность совершенствования методов управления потоками сообщений без замены технических средств, с целью дальнейшего повышения эффективности использования сети связи, на которой базируется информационная сеть; такая возможность обусловлена тем, что обработка сообщений на коммутационных узлах осуществляется с помощью универсальных или специализированных устройств с программным управлением.

Рассмотренные методы распределения потоков информации в узлах сети целесообразно применять в тех случаях, когда сообщения между обменивающимися объектами носят непрерывный характер. При обмене информацией между вычислительными центрами и объектами управления потоки информации, поступающие в узлы сети, имеют в основном пакетный (дискретный) характер, количество многоадресных потоков в информационной сети возрастает, увеличивается также число стандартных скоростей аппаратур передачи данных. Поэтому в сетях телекоммуникации естественно использование метода коммутации пакетов при оптимальном по некоторому критерию оценки распределении потоков информации.

## 1. 2. Параметры, отражающие функционирование сети телекоммуникации при передаче данных

Известно, что информационная сеть с коммутацией сообщений или пакетов, предназначенная для направленного обмена информацией между входами и выходами методом коммутации, представляет собой совокупность взаимодействующих элементов: конечных узлов, трактов передачи данных, узлов коммутации и отдельных органов управления функционированием сети. Основными параметрами, отражающими функционирование сети, являются ее надежность, достоверность принятого сообщения, определяемая помехоустойчивостью тракта передачи данных, и время доставки сообщения, зависящее от скорости его передачи.

Нормы на надежность, время доставки и достоверность различны для каждого отдельного сообщения или определенных групп (категорий) сообщений, объединяемых по признаку важности, срочности и смыслового содержания. В связи с этим было введено понятие нормальной доставки сообщения в заданный адрес за время, не превышающее времени его передачи при соблюдении степени соответствия для сообщений данной категории. Надежность нормальной доставки сообщения зависит от требований к сети, выполнение которых обеспечивает желаемую степень гарантии необходимого качества обслуживания внешних сообщений. Время доставки – также индивидуальное требование.

Важными параметрами каналов информационной сети являются достоверность и скорость передачи информации. Достоверность принятого сообщения характеризуется степенью его соответствия переданному сообщению. Количественной мерой достоверности служит вероятность возникновения в сообщении ошибки, называемая коэффициентом потерь достоверности.

Скорость передачи информации в каналах можно определить из выражения

$$B = \lim_{T \rightarrow \infty} \frac{J(Y, X)}{T} = \lim_{T \rightarrow \infty} \frac{H(Y) - H(N)}{T},$$

где  $J(Y,X)=H(Y)-H(N)$  — количество информации, приходящееся на 1 символ;  $H(Y)$  — энтропия на 1 символ источника сообщения;  $H(N)=H(Y/X)$  — условная энтропия на 1 символ, определяемая вероятностью того, что был отправлен символ  $X$ , если принят символ  $Y$  (Шеннон определил условную энтропию как ненадежность канала, обуславливающую «рассеяние» информации в канале из-за влияния искажений). Устремление  $T \rightarrow \infty$  (длительность каждого одиночного сообщения) означает, что для определения средней скорости передачи надо рассмотреть всю совокупность возможных сообщений.

В реальных случаях характер пропускной способности различных направлений зависит от вида и числа каналов. Соответственно при функционировании сети количество каналов должно быть выбрано по заданным направлениям.

Необходимое количество каналов можно определить по формуле:

$$n = \frac{\Delta m_{n.u}}{K_{п.и} \lambda_p \Delta t B},$$

где  $\Delta m_{n.u}$  — объем полезной информации, который необходимо передать за время  $\Delta t$ ;  $K_{п.и}$  — коэффициент передачи информации, характеризующий реальные возможности каналов по передаче данных;  $B$  — скорость передачи информации;  $\lambda_p$  — среднесуточный резерв по интенсивности нагрузки каналов.

При распределении информационных потоков в узлах сети необходимо учитывать также параметры, характеризующие динамику изменения технико-экономических показателей управляемого объекта.

При функционировании сети информация доставляется не мгновенно, а вследствие ряда причин — с некоторым запаздыванием. Во многих конкретных случаях из-за упомянутого запаздывания сообщение доставляется адресату с некоторыми отклонениями от истинного значения параметра, описывающего состояние процесса в момент получения информации, ввиду динамики объекта

наблюдения. Вследствие этого определенные процессы будут протекать в системах неоптимальным образом, т. е. с какими-то потерями  $C_{п}$ , которые определяются стоимостными оценками. На рис.1.1. представлен один из возможных типов графика зависимости стоимости потерь (штрафа) от времени запаздывания информации.

Время запаздывания информации определяется главным образом суммарным временем ожидания в узлах сети и передачи по тракту связи, которое характеризуется выбором скорости работы аппаратуры.

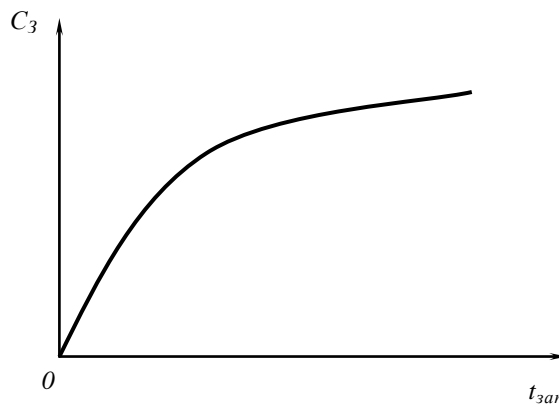


Рис.1.1. Зависимость стоимости потерь (штрафа) от времени запаздывания информации

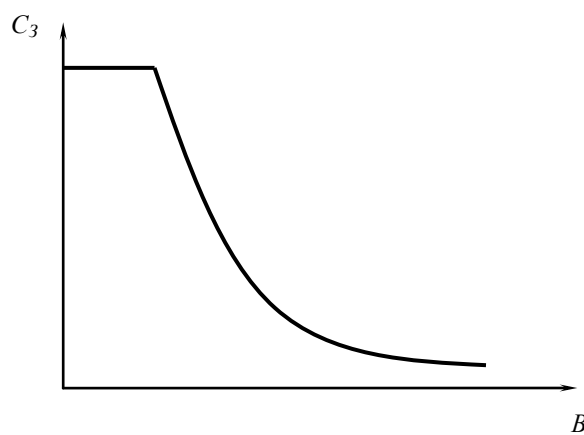


Рис. 1.2. Зависимость стоимости потерь от скорости передачи сообщения

Уменьшение стоимости потерь, зависящих от времени запаздывания, может быть достигнуто повышением скорости передачи информации по каналу, На рис.1.2. представлена одна из возможных зависимостей стоимости потерь, обусловленных запаздыванием, от скорости передачи для некоторого сообщения гипотетического объекта. Однако в процессе прохождения сообщения по тракту связи из-за помех и других причин возникают неизбежные ошибки, интенсивность которых (определяемая, например, средней вероятностью ошибки на символ  $P_{ош}$ ) возрастает с увеличением скорости передачи (при прочих равных условиях – мощности передатчика, вида модуляции и т. п.— без учета специальных мер по повышению помехоустойчивости). На рис.1.3 показан характер зависимости  $P_{ош}$  от скорости передачи информации.

Ошибки в получаемой информации приводят к отклонениям управляемого процесса от оптимума и могут так же, как и запаздывание, характеризоваться среднестатистическими стоимостными оценками потерь (штрафа). На рис.1.4. представлен один из возможных типов зависимости стоимости потерь от вероятности ошибок на символ информации.

Снижения стоимости потерь от ошибок можно добиться повышением помехоустойчивости систем передачи данных, что возможно при информационной избыточности (приводящей к снижению скорости передачи) и усложнении технических средств передачи данных. Применение более сложных технических средств для повышения помехоустойчивости приводит к дополнительным материальным затратам, которые также определяются стоимостными оценками.

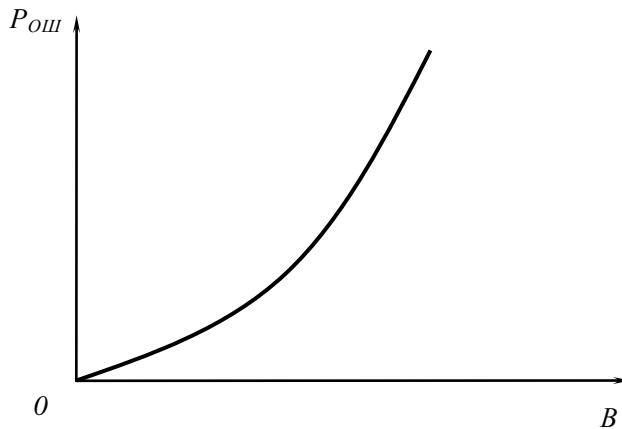


Рис.1.3. Зависимость  $P_{ош}$  от скорости передачи информации

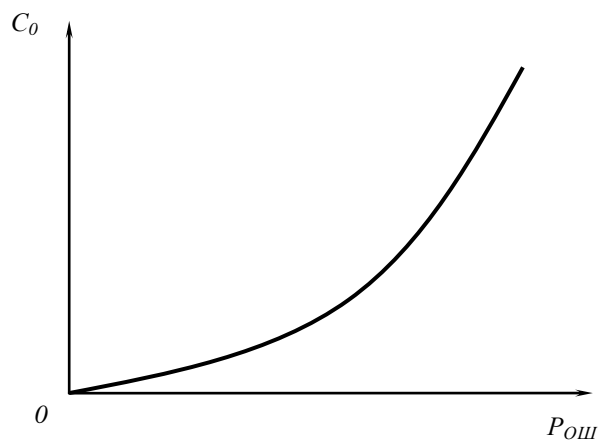


Рис.1.4. Зависимость стоимости потерь от вероятности ошибок на символ информации

В общем случае зависимость средней стоимости общих потерь от запаздывания и ошибок в сообщении можно представить в виде

$$C_{\Pi} = \sum (C_{зан} + C_{ош})$$

Общая стоимость потерь  $C_{\Pi}$  зависит от всех перечисленных выше

параметров:

$$C_{\Pi} = F[t_{зан}, P_{ош}]$$

Для определения оптимального режима использования сети необходимо найти значения указанных параметров, при которых общая стоимость потерь с учетом затрат на повышение качества функционирования сети будет минимальной.

Функциональная зависимость стоимостных потерь от времени запаздывания  $C_{зан} = f(t_{зан})$  тесно связана с ценностью информации, может иметь различный характер, и определяется конкретным назначением сообщения, динамикой и экономическими параметрами объекта и управляющей системы. Будем именовать такую зависимость временной ценностной характеристикой (ВЦХ) сообщения (порции или пакета информации). Применение такого критерия оптимальности при распределении информационных потоков в узлах сети позволяет учитывать технико-экономические показатели объектов-потребителей информации.

Стоимостные потери, связанные с запаздыванием информации ( $C_{зан} = f(t_{зан})$ ) и с ошибками ( $C_{ош} = \varphi(P_{ош})$ ), задаются потребителем информации, и для разных объектов и сообщений имеют различный характер. Такие зависимости для конкретного объекта могут быть получены экспериментальным путем или методом статистического моделирования.

### 1.3. Выбор критерия оптимальности построения расписаний

В этом параграфе основное внимание уделяется одностадийным обслуживающим системам, в которых:

а)  $Q^{(i)} = \{1, 2, \dots, M\}$  ( $\overline{1, n}$ ), т. е. каждый прибор может обслуживать любое из требований множества  $N$ ;

б) В любой момент времени каждое требование обслуживается не более

чем одним прибором, и каждый прибор обслуживает не более одного требования.

Для требования  $i \in N$  задан момент времени  $d_i \geq 0$  (момент поступления в очередь на обслуживание), начиная с которого оно может обслуживаться.

Длительности  $t_{iL} > 0$  обслуживания каждого требования  $i \in N$  каждым прибором  $1 \leq L \leq M$  предполагаются заданными. Если  $t_{iL} = a_L t_i$  ( $i = \overline{1, n}, 1, n, L = \overline{1, M}$ ), то говорят, что прибор  $L$  обладает «производительностью», равной  $1/a_L$ .

Если  $a_L = 1$  ( $L = \overline{1, M}$ ), то приборы *идентичны*.

В зависимости от характера обслуживающей системы процесс обслуживания каждого требования либо должен протекать непрерывно, либо могут допускаться *прерывания* с последующим дообслуживанием требования любым из приборов. Прерывания могут допускаться как в заданные сроки, так и в произвольные моменты времени. Обычно предполагается, что прерывания не сопряжены с временными затратами и их число конечно.

Процесс обслуживания требований может быть описан заданием совокупности  $s = \{s_1(t), s_2(t), \dots, s_M(t)\}$  кусочно-постоянных непрерывных слева функции  $s_L = s_L(t)$  ( $L = \overline{1, M}$ ), каждая из которых задана на интервале  $0 \leq t < \infty$  и принимает значения  $0, 1, \dots, n$ . Если  $s_L(t') = i \neq 0$ , то в момент времени  $t'$  прибор  $L$  обслуживает требование  $i$ . Если  $s_L(t') = 0$ , то в момент времени  $t'$  прибор  $L$  простаивает.

Поскольку каждое требование не может одновременно обслуживаться двумя и более приборами, то из условия  $s_L(t') = i \neq 0$  следует  $s_H(t') \neq i$  для всех  $1 \leq H \neq L \leq M$ . Поскольку требование  $i$  поступает в очередь на обслуживание в момент времени  $d_i$ , то  $s_L(t) \neq i$  ( $L = \overline{1, M}$ ) при  $t \leq d_i$  ( $i = 1, n$ ).

Если  $t_{iL}$  — суммарная длина временных интервалов, на которых функция  $s_L(t)$  принимает значение  $i$ , то должны выполняться соотношения

$$\sum_{L=1}^M (t'_{iL} / t_{iL}) = 1 (i = \overline{1, n}).$$

В частности, если приборы идентичны, то суммарная

длина всех временных интервалов, на которых функции  $s_L(t)$  ( $L = \overline{1, M}$ ) принимают значение  $i$ , должна быть равна  $t_i$ .

Совокупность функции  $s$ , обладающая перечисленными свойствами, называется *расписанием* обслуживания требований множества  $N$  системой, содержащей  $M$  параллельных приборов.

На рис.1.5 приведен график расписания  $s(t)$  обслуживания требований множества  $N = \{1, 2, 3, 4\}$  одним прибором, т. е. АТС. Значения  $d_1 = 0, d_2 = 2, d_3 = d_4 = 3, t_1 = 4, t_2 = 1, t_3 = t_4 = 2$ .

Если система содержит два и более приборов, то графики функций  $s_L(t)$  обычно совмещают, как это сделано, например, на рис.1.6. Здесь  $M = 3$ , приборы идентичны,  $N = \{1, 2, 3, 4, 5\}$ ;  $d_1 = d_2 = d_3 = 0, d_4 = 1, d_5 = 2$ ;  $t_1 = 1, t_2 = t_3 = 3, t_4 = t_5 = 2$ . Прибор 1 во временном интервале  $(0, 1]$  обслуживает требование 1, а в интервале  $(3, 4]$  — требование 5. Прибор 2 в интервалах времени  $(0, 1]$  и  $(3, 5]$  обслуживает требование 2, а в интервале  $(1, 3]$  — требование 4. Наконец, прибор 3 обслуживает требование 3 во временном интервале  $(0, 3]$  и требование 5 — в интервале  $(4, 5)$ .

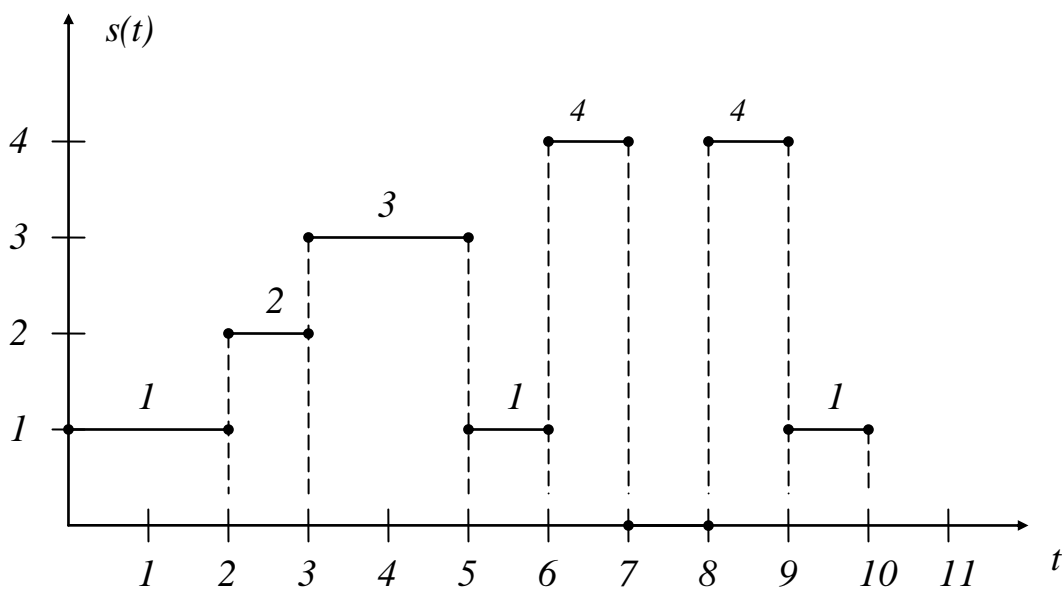


Рис.1.5. График расписания  $s(t)$  обслуживания требований множества  $N = \{1, 2, 3, 4\}$  одним прибором

Говорят, что расписание  $s = \{s_1(t), s_2(t), \dots, s_M(t)\}$  допускает прерывания в

процессе обслуживания требований, если существуют такие  $1 \leq i \leq n$ ,  $1 \leq L \neq H \leq M$  и  $0 \leq t' < t < t'' < \infty$ , что выполняется хотя бы одно из условий: 1)  $s_L(t') = s_L(t'') = i$ , но  $s_L(t) \neq i$ ; 2)  $s_L(t') = s_H(t'') = i$ . Если при этом  $s_L(t' + \delta) \neq i$  для любого достаточно малого

$\delta > 0$ , то в момент времени  $t'$  имеет место прерывание обслуживания требования  $i$  прибором  $L$ . Не исключается возможность непосредственного продолжения обслуживания требования  $i$ , но другим прибором.

Содержательно процесс обслуживания требований без прерываний удовлетворяет следующему условию. Каждое требование обслуживается только одним прибором. Если обслуживание некоторого требования  $i$  прибором  $L$  начинается в момент времени  $t_i^0$  то оно протекает непрерывно и завершается в момент времени  $\bar{t}_i = t_i^0 + t_{iL}$ . В этом случае расписание, очевидно, полностью определяется распределением требований по приборам и указанием для каждого требования  $i$  момента  $t_i^0$  начала его обслуживания. Если в процессе обслуживания требования допускаются прерывания, то каждое отдельное требование может обслуживаться «по частям» и по обязательно одним и тем же прибором. Так, при расписании, изображением на рис.1.6. допускаются прерывания в обслуживании требований 2 и 5. Обслуживание требования 2 прибором 2 прерывается в момент времени  $t = 1$  с последующим его дообслуживанием тем же прибором, начиная с момента времени  $t = 3$ . Обслуживание требования 5 прибором 1 прерывается в момент времени  $t = 4$  и продолжается с этого момента времени прибором 3. При расписании, изображенном на рис. В.1, допускаются прерывания в обслуживании требований 1 и 4. Во временном интервале (7, 8] прибор простаивает.

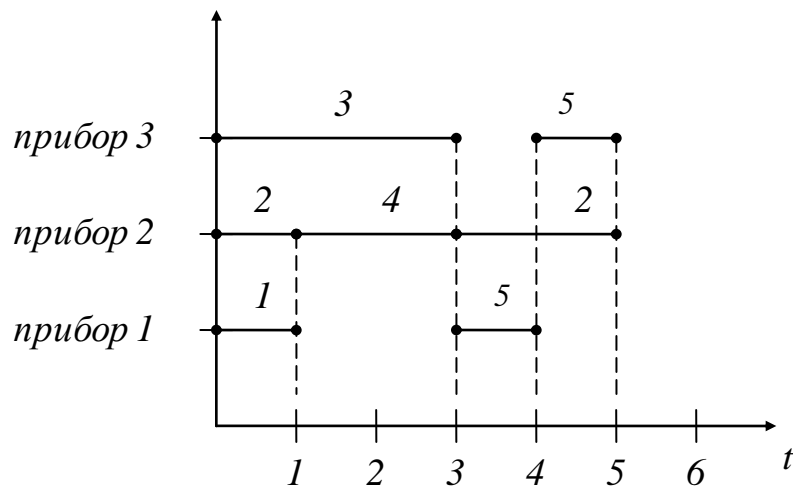


Рис. В.2.

Рис.1.6. График расписания  $s(t)$  обслуживания требований множества  $N = \{1, 2, 3, 4\}$  несколькими приборами

В приложениях числа  $d_i$  и  $t_{iL}$  являются рациональными, и при выборе соответствующей единицы измерения их можно считать целыми. В этом случае зачастую ограничиваются рассмотрением расписаний с прерываниями в целочисленные моменты времени. При этом предполагается, что моменты начала обслуживания и дообслуживания каждого требования также целочисленны. Такие расписания, очевидно, однозначно определяются заданием для каждого временного интервала единичной длины  $M$ -мерного вектора с компонентами 0, 1, ...,  $n$ . Если  $L$ -я компонента этого вектора равна  $i \neq 0$ , то в рассматриваемом единичном временном интервале прибор  $L$  обслуживает требование  $i$ . В противном случае прибор простаивает.

При разрешении прерываний в произвольные моменты времени, наряду с предположением об ограниченности числа прерываний, естественно предполагать, что длительность каждого частичного обслуживания требования является конечной величиной.

Кроме запрещения прерываний, к расписанию могут предъявляться и другие требования, вытекающие из постановки конкретно рассматриваемой

задачи. Так, для каждого требований  $i$  может быть задан *директивный срок*  $D_i$ , к которому необходимо или, во всяком случае, желательно завершить обслуживание требования  $i$ . Расписание, при котором все требования обслуживаются в заданные директивные сроки, называется *допустимым относительно этих сроков*. В общем случае такие расписания могут, очевидно, не существовать.

Весьма распространенными являются ситуации, в которых накладываются некоторые ограничения на возможную последовательность обслуживания требований. Если по условию задачи требование  $j$  может обслуживаться только после завершения обслуживания требования  $i$ , то расписание  $s$  должно удовлетворять условию: если  $s_L(t') = i$  при некотором  $1 \leq L \leq M$  и некотором  $t' > 0$ , то  $s_H(t) \neq j$  при всех  $1 \leq H \leq M$  и  $t \leq t'$ . Ситуации такого рода обычно описываются заданием на множестве требований  $N$  некоторого отношения, строгого порядка  $\rightarrow$ , согласно которому обслуживание требования  $i$  должно завершиться до начала обслуживания требования  $j$ , если  $i \rightarrow j$ . При этом говорят о расписаниях, *допустимых относительно заданного на  $N$  порядка*.

Процесс обслуживания требований может быть сопряжен с потреблением или использованием некоторых дополнительных *ресурсов*. Типичную ситуацию такого рода можно описать следующим образом. При обслуживании требований используется  $q$  видов ресурсов. В момент времени  $t$  ресурс вида  $k$  имеется в количестве  $R_k(t)$  ( $k = \overline{1, q}$ ). Для обслуживания требования  $i$  в момент времени  $t$  необходим ресурс вида  $k$  в количестве  $r_{ik}(t)$  ( $i = \overline{1, n}$ ,  $k = \overline{1, q}$ ). Если в момент времени  $t$  обслуживаются требования  $i_1, i_2, \dots, i_l$  и только они, то необходимо, чтобы 
$$\sum_{j=1}^l r_{jk}(t) \leq R_k(t)$$
 для всех  $1 \leq k \leq q$ . Расписание  $s$ , при котором указанные ресурсные ограничения выполняются для всех  $t \geq 0$ , называется *допустимым ресурсам*.

Определенный практический интерес представляют также расписания, которые удовлетворяют ограничениям, связанным с необходимостью учета

переналадок обслуживающих приборов, соблюдения условий группирования требований при их обслуживании и т. п. При рассмотрении задач такого рода под допустимым понимается такое расписание, которое удовлетворяет всем ограничениям, вытекающим из постановки конкретно рассматриваемой задачи.

Следует отметить, что построение допустимого расписания или даже выяснение того факта, существу ли оно вообще, является зачастую далеко не тривиальной задачей. Вместе с тем во многих ситуациях построение допустимых расписаний не вызывает сколь-нибудь серьезных затруднений, и речь идет о выборе среди них наилучшего в том или ином смысле расписания.

Наиболее распространенный способ оценки *качества* расписаний обслуживания требований в одностадийных системах состоит в следующем. Каждому расписанию  $s$  соответствует вектор  $\bar{t}(s) = (\bar{t}_1(s), \dots, \bar{t}_n(s))$  моментов завершения обслуживания требований при этом расписании. Здесь  $\bar{t}_i(s)$  — наибольшее значение  $t$ , при котором существует такое  $L \in \{1, 2, \dots, M\}$ , что  $s_L(t) = i$ . Задается действительная неубывающая по всем аргументам функция  $n$  переменных  $F(x) = F(x_1, x_2, \dots, x_n)$ . Качество расписания  $s$  характеризуется значением этой функции при  $x = \bar{t}(s)$ . Из двух расписаний лучшим считается то, которому соответствует меньшее значение  $F(x)$ . Расписание, которому соответствует наименьшее значение  $F(x)$  (среди всех допустимых расписаний), называется *оптимальным* расписанием.

При задании функции  $F(x)$  каждому требованию  $i$  обычно сопоставляют некоторую неубывающую функцию — функцию штрафа  $\varphi_i(t)$ , выражающую в количественном отношении «штраф», который необходимо «заплатить», если обслуживание этого требования завершится в момент времени  $t$ . Качество расписания характеризуется суммарным или максимальным штрафом, который необходимо заплатить при обслуживании требований по расписанию  $s$ , т. е.

$$F_{\Sigma}(s) = \sum_{i=1}^n \varphi_i(\bar{t}_i(s)) \text{ или } F_{\max}(s) = \max_{1 \leq i \leq n} \{\varphi_i(\bar{t}_i(s))\}$$

В частности, если  $\varphi_i(t) = t$  ( $i = \overline{1, n}$ ), то  $F_{\max}(s) = \max_{1 \leq i \leq n} \{t_i(s)\}$  — момент завершения обслуживания всех требований (общее время обслуживания). В этом случае  $F_{\max}(s)$  обозначают через  $\bar{t}_{\max}(s)$ , и расписание  $s$ , доставляющее наименьшее значение  $\bar{t}_{\max}(s)$ , называют оптимальным по быстрдействию расписанием.

Если  $\varphi_i(t) = t - D_i$ , то  $F_{\max}(s)$  обозначают через  $L_{\max}(s)$ .

Имеем  $L_{\max}(s) = \max_{1 \leq i \leq n} \{L_i(s)\}$ , где  $L_i(s) = \bar{t}_i(s) - D_i$  — временное смещение момента завершения обслуживания требования  $i$  относительно директивного срока  $D_i$ .

Если  $\varphi_i(t) = \max\{0, t - D_i\}$ , то  $F_{\max}(s)$  обозначают через  $Z_{\max}(s)$ . Имеем

$$z_{\max}(s) = \max_{1 \leq i \leq n} \{z_i(s)\} \text{ где } z_i(s) = \max\{0, \bar{t}_i(s) - D_i\} \text{ — запаздывание в}$$

обслуживании требования  $i$  относительно директивного срока  $D_i$ . В этом случае

$$F_{\Sigma}(s) = \sum_{i=1}^n z_i(s) \text{ — суммарное запаздывание.}$$

$$\text{Если } \varphi_i(t) = \text{sign}(\max\{0, t - D_i\}), \text{ то } F_{\Sigma}(s) = \sum_{i=1}^n u_i(s),$$

где  $u_i(s) = \text{sign}(z_i(s))$ , представляет собой число запаздывающих (относительно директивных сроков) требований.

В ряде случаев каждому требованию  $i$  сопоставляется некоторое число  $\alpha_i$  — «вес» требования  $i$ , и говорят о взвешенной сумме моментов завершения

обслуживания требований  $\sum_{i=1}^n \alpha_i \bar{t}_i(s)$  взвешенном суммарном запаздывании

$\sum_{i=1}^n \alpha_i z_i(s)$  и о сумме весов запаздывающих требований (взвешенном числе запаздывающих требований)  $\sum_{i=1}^n \alpha_i u_i(s)$ .

Перечисленные критерии оптимальности отражают стремление скорейшего завершения обслуживания каждого требования с учетом их возможной неоднородности. В этих условиях при поиске оптимального расписания можно ограничиться рассмотрением класса расписаний, при которых не допускаются неоправданные простои обслуживающих приборов. Если прерывания процесса обслуживания требований запрещены либо допускаются только в целочисленные моменты времени, то указанный класс содержит конечное число расписаний.

Действительно, пусть  $s$  — расписание обслуживания требование одним прибором без прерываний, при котором требование обслуживаются в последовательности  $\pi = (i_1, i_2, \dots, i_n)$  где  $\pi$  — некоторая перестановка элементов множества  $N$ . Момент  $t_{i_j}^0(s)$  начала обслуживания требования  $i_j$  при этом расписании удовлетворяет, очевидно, неравенству  $t_{i_j}^0(s) \geq \max\{\bar{t}_{i_{j-1}}(s), d_{i_j}\}$ , а момент завершения обслуживания этого требования равен  $\bar{t}_{i_j}(s) = t_{i_j}^0(s) + t_{i_j}$  ( $j = \overline{1, n}$ );  $\bar{t}_{i_0}(s) = 0$ . Рассмотрим расписание  $s'$ , при котором требования обслуживаются в той же последовательности  $\pi$ ,

$$t_{i_j}^0(s') = \{\max \bar{t}_{i_{j-1}}(s'), d_{i_j}\}, \bar{t}_{i_j}(s') = t_{i_j}^0(s') + t_{i_j} \quad (j = \overline{1, n}), \bar{t}_{i_0}(s') = 0.$$

Нетрудно видеть, что  $\bar{t}_{i_j}(s') \leq \bar{t}_{i_j}(s)$  ( $j = \overline{1, n}$ ), и поскольку функция  $F(x)$  является неубывающей, то  $F(\bar{t}(s')) \leq F(\bar{t}(s))$ . Расписание  $s'$  однозначно определяется заданием перестановки  $\pi$ , и, следовательно, при поиске оптимального расписания достаточно рассмотреть  $n!$  расписаний.

Аналогично, пусть  $s$  — расписание обслуживания требований  $M$  параллельными приборами без прерываний, при котором прибор  $L$  обслуживает

требования множества  $N_L$  в последовательности  $\pi^L = (i_1^L, i_2^L, \dots, i_{n_L}^L)$ ,  $L = \overline{1, M}$ .  
Здесь  $N_1 \cup N_2 \cup \dots \cup N_M = N$ ,  $N_H \cap N_R = \emptyset$ ,  $1 \leq H \neq R \leq M$  и необязательно  $N_L = \emptyset$ .  
Рассмотрим расписание  $s'$ , при котором момент времени начала обслуживания  
требования  $i_j^L$  равен  $t_{i_j^L}^0 = \max\{\bar{t}_{i_{j-1}^L}, d_{i_j^L}\}$ , а момент завершения обслуживания  
этого требования  $\bar{t}_{i_j^L} = t_{i_j^L}^0 + d_{i_j^L}$ ,  $(i = \overline{1, n})$ ,  $\bar{t}_{i_0^L} = 0$ ,  $(L = \overline{1, M})$ . Очевидно,  
 $F(\bar{t}(s')) \leq F(\bar{t}(s))$ . Расписание  $s'$  однозначно определяется заданием разбиения  
множества  $N$  на подмножества  $N_1, N_2, \dots, N_M$  (некоторые из них могут быть  
пустыми) и заданием перестановок элементов этих множеств. Следовательно, при  
поиске оптимального расписания достаточно рассмотреть не более  $n! \cdot C_{M+n-1}^n$   
расписаний, где  $C_q^p$  — число сочетаний из  $q$  по  $p$  элементов. Если приборы  
идентичны, то число рассматриваемых расписаний уменьшается в  $M!$  раз.

Если прерывания процесса обслуживания требований допускаются только в  
целочисленные моменты времени, то, как уже отмечалось, расписание однозначно  
определяется заданием для каждого временного интервала единичной длины  $M$ -  
мерного вектора с компонентами  $0, 1, \dots, n$ . При этом, очевидно, достаточно  
рассматривать временной диапазон (так называемый интервал планирования) от

$$\min_{1 \leq i \leq n} d_i \text{ до } \max_{1 \leq i \leq n} d_i + \sum_{i=1}^n \max_{1 \leq L \leq M} t_{iL}.$$

Обозначая длину этого диапазона через  $T$ , приходим к заключению, что при поиске оптимального расписания достаточно  
рассмотреть не более  $T * C_{n+1}^M$  расписаний.

В случае, когда прерывания допускаются в произвольные моменты времени,  
при поиске оптимального расписания, вообще говоря, нельзя ограничиться  
рассмотрением конечного множества расписаний. Однако при определенных  
условиях и в этом случае может быть выделено конечное множество расписаний,  
содержащее хотя бы одно оптимальное расписание.

Аналогичные рассуждения можно провести и относительно допустимых в  
том или ином смысле расписаний.

Таким образом, оптимальное расписание может быть найдено, как правило, в результате перебора конечного множества возможных вариантов. Основное затруднение состоит в том, что число таких вариантов обычно оказывается исключительно большим (так, например, уже  $10! = 3\,628\,800$ ) и растет по меньшей мере экспоненциально с ростом размерности задачи. Стремлению максимально сократить этот перебор, построить оптимальное расписание в результате проведения возможно меньшего объема вычислений подчинены все работы в области теории расписаний.

Если, объем вычислений ограничен некоторым полиномом от размерности задачи, то говорят, что задача относится к классу полиномиально разрешимых. Соответствующий алгоритм при этом называют полиномиальным. Наряду с полиномиально разрешимыми задачами известны так называемые *NP*-трудные задачи, для которых полиномиальных алгоритмов, по-видимому, не существует.

Важнейшую роль при проектировании всякой сложной системы играет правильный выбор критерия качества работы. Заметим, что, как бы ни была сложна система и каким бы большим числом параметров она ни характеризовалась, критерии качества работы должен выражаться скалярным числом, ибо о различных вариантах одной и той же системы требуется сказать, какой вариант лучше, и желательно сказать насколько.

Остановимся на критериях и алгоритмах, наиболее соответствующих использованию метода датаграмм. В соответствии с этим каждый пакет рассматривается как независимая единица, а критерии работы сети являются средними характеристиками качества обслуживания множества независимых пакетов.

Рассмотрим два критерия. В соответствии с первым критерием качество работы системы связи оценивается по среднему времени доставки пакетов. Этот критерий используется практически по всех известных сетях ЭВМ. Второй критерий учитывает ценность пакетов, их старение—изменение ценности пакетов с течением времени их пребывания в системе.

Качество работы всей системы в целом оценивается по величине

потерянной ценности пакетов. Соответственно и описываемые ниже методы делятся на две группы: первая направлена на уменьшение среднего времени доставки пакетов, вторая — на снижение потерь ценности пакетов.

Пусть требуется организовать доставку пакетов единичной длины между узлами коммутационной сети. Наиболее простой способ состоит в том, что каждый пакет направляется к адресату по пути с минимальным числом транзитных участков, чему в ненагруженной сети соответствует, очевидно, минимальное время доставки.

Задать такой план распределения можно с помощью матриц, хранящихся в каждом узле и указывающих на минимальное время доставки от данного узла до всех других узлов назначения по каждому исходящему направлению. Таким образом, будем предполагать, что в каждом узле имеется матрица, число строк которой равно числу узлов сети и число столбцов равно числу исходящих из узла направлений, точнее, числу драйверов каналов. Элемент  $t_{ij}^k$ , матрицы  $T_k$  узла  $k$  указывает минимальное время, необходимое для доставки сообщения из узла  $k$  к узлу  $i$ , по исходящему направлению к соседнему с узлом  $k$  узлу  $j$ .

Очевидно, если сеть не подвергается повреждениям и не загружена, то, используя информацию, хранящуюся в матрице  $T_k$ , появившийся пакет дойдет до узла назначения за кратчайшее время. Оставляя пока вне рассмотрения возможные повреждения в сети, отметим, что характерной чертой систем коммутации пакетов является их большая загруженность и возможность образования очередей. Возможны различные дисциплины организации очередей. Для методов управления, направленных на минимизацию среднего времени доставки, остановимся только на двух — прямом и инверсионном способах построения очередей.

Инверсионный способ построения очереди состоит в том, что вновь поступивший пакет ставится в очередь первым, однако передача пакета, уже занявшего канал, не прерывается. При прямом способе обслуживания новый пакет ставится в очередь последним.

Первый, наиболее простой способ управления распределением потоков состоит в том, что поступивший в исходящий или транзитный узел пакет направляется на исходящее из узла направление, которому соответствует минимальное значение элемента в строке, номер которой совпадает с номером узла назначения  $\min_{j \in I(k)} \{t_{ij}^k\}$  где  $I$  - номер узла назначения, а  $I(k)$ —множество узлов, соседних с узлом  $k$ . Если направление  $j \in I(k)$  занято, то пакет ставится в очередь в соответствии с принятой дисциплиной построения очереди. Так как матрицы  $T_k$  неизменны, то, по сути дела, для каждого узла назначения раз и навсегда задано исходящее направление. Этот вариант управления будем обозначать в дальнейшем ВР1П (прямой) или ВР1И (инверсионный), в зависимости от принятой дисциплины построения очереди.

Применение варианта ВР1 может оказаться целесообразным лишь при очень малой загрузке сети и низком использовании каналов, поскольку при распределении потоков не учитывается возможность удлинения пути за счет образования очередей и не используются обходные направления.

ВР2 отличается от описанного варианта ВР1 тем, что при выборе исходящего направления в узле учитывается величина очереди на каждом исходящем из данного узла направлении. Обозначим через  $W_j^k$  число пакетов, ожидающих отправки из узла  $k$  по направлению  $j$ , и  $0 \leq \tau \leq 1$  — время, оставшееся до окончания передачи пакета, занявшего канал  $k_j$ . Тогда поступивший пакет направляется к узлу  $j$ , соседнему с узлом  $k$ , для которого

$$t_{ij}^k + \tau_j^k + W_j^k = \min_{l \in I(k)} \{t_{il}^k + \tau_l^k + W_l^k\} \quad (1.1)$$

Очевидно, при прямом построении очереди, когда пакет ставится в очередь последним, соотношение (1.1) означает, что пакет будет доставлен в узел назначения через время  $t_i^k = t_{ij}^k + \tau_j^k + W_j^k$  в предположении, что вся остальная сеть свободна. В случае же инверсионного построения очереди ожидаемое время доставки равно  $t_{ij}^k + \tau_j^k$ , а член  $W_j^k$  учитывает, что каждому пакету, стоящему в очереди, придется ожидать отправки на единицу времени дольше, т. е.

учитываются увеличившиеся задержки для других пакетов. При использовании ВР2 принимается во внимание загрузка исходящих из узла направлений и допускается возможность обходных направлений, но не учитывается ситуация на сети в целом.

Для сравнения методов и изучения поведения сетей, управляемых с помощью различных методов и в разных условиях, можно воспользоваться методом моделирования.

В дальнейшем изложение будет иллюстрироваться рядом примеров моделирования сети, показанной на рис.1.7.

Доля нагрузки, создаваемая каждым узлом сети, приведена в табл.1.1. Распределение тяготений, т. е. доля пакетов, направляемая от некоторого узла  $k$  к любому другому узлу, для всех узлов одинакова (см. табл.1.1). Нагрузка в сети изменяется путем задания общего числа пакетов, возникающих в сети за единицу времени. Доля же нагрузки, создаваемой каждым узлом, и распределение тяготений сохраняются неизменными. При этом, конечно, в узлах не генерируются пакеты, направленные к тому же узлу.

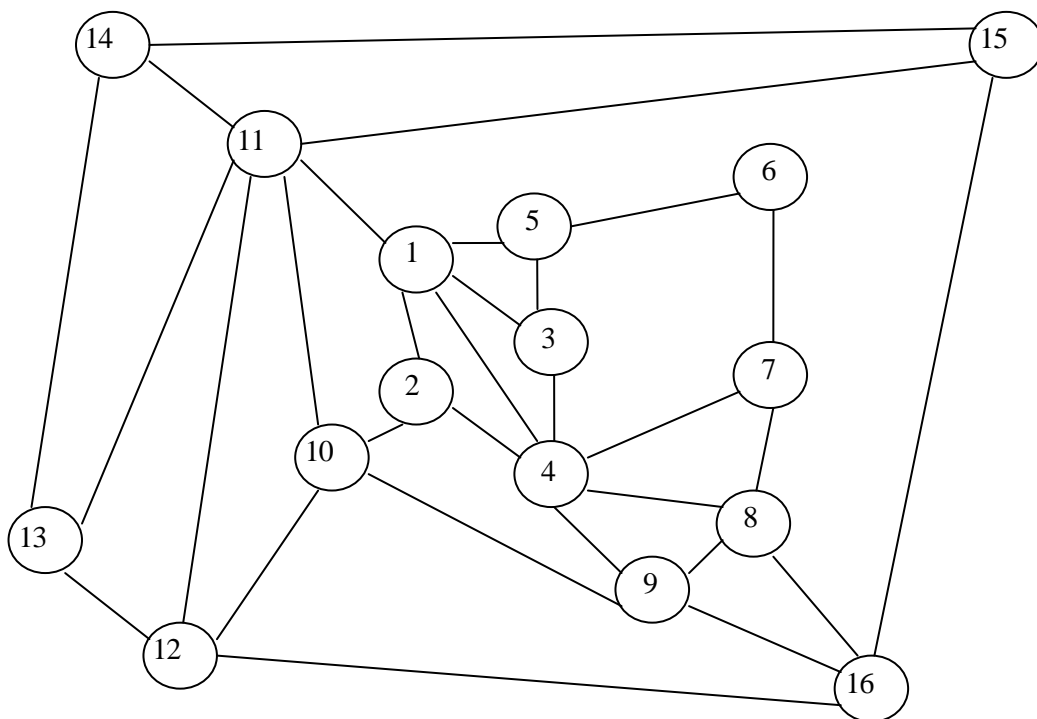


Рис.1.7. Конфигурация моделируемой сети  
цифры – номера узлов

Для сети, показанной на рис.1.7. величина предельной пропускаемой сетью нагрузки при указанных распределениях нагрузки и тяготений весьма мала для варианта ВР1И. Так, при нагрузке, составляющей 7,9 пакета, генерируемых в сети в единицу времени, система переполняется за время, равное 4201,4 единицам времени (эксперимент 1). Переполнением системы условно считается образование в системе очередей с общим числом ожидающих пакетов, равным 2000. Всего к указанному моменту времени в сеть поступило 33 414 пакетов. Среднее время доставленных пакетов составило 3,8 единицы времени. Поскольку использовалась инверсионная дисциплина построения очередей, то следует ожидать, что пакеты, находящиеся в системе, имеют относительно большой возраст.

Таблица 1.1

**Распределение нагрузки, создаваемой каждым узлом сети,  
и тяготений к каждому из узлов**

<b>Номер узла</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
<b>Нагрузка</b>	<b>3,75</b>	<b>3,00</b>	<b>3,20</b>	<b>4,00</b>	<b>2,50</b>	<b>1,50</b>	<b>2,00</b>	<b>3,00</b>
<b>Тяготения</b>	<b>0,08</b>	<b>0,06</b>	<b>0,06</b>	<b>0,1</b>	<b>0,06</b>	<b>0,02</b>	<b>0,05</b>	<b>0,06</b>
<b>Номер узла</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>
<b>Нагрузка</b>	<b>2,00</b>	<b>2,00</b>	<b>4,00</b>	<b>3,00</b>	<b>2,00</b>	<b>2,50</b>	<b>1,00</b>	<b>2,50</b>
<b>Тяготения</b>	<b>0,06</b>	<b>0,06</b>	<b>0,1</b>	<b>0,06</b>	<b>0,05</b>	<b>0,05</b>	<b>0,06</b>	<b>0,06</b>

О скорости переполнения системы говорит также то, что за время, в течение которого в систему поступили последние 2000 пакетов, число находящихся в системе пакетов увеличилось на 84.

При уменьшенной в два раза нагрузке переполнения системы не происходит и среднее время доставки пакетов  $t_d=2,5$  ед. вр. Образования очередей практически не происходит (эксперимент 2).

Использование варианта ВР2 для распределения потоков существенно повышает пропускную способность системы. Так, при нагрузке, составляющей 12 пакетов в единицу времени, при применении варианта ВР2И –  $t_d = 6,8$  ед. вр., варианта ВР2П —  $t_d = 25,0$  ед. вр.

Такое большое отличие во временах доставки при использовании двух сходных вариантов объясняется, по-видимому, следующим обстоятельством. При построении очереди по инверсионному принципу новый пакет, который становится в начало очереди, будет отправлен по оптимальному пути с большей вероятностью, чем при прямом построении очереди. Очевидно, что с уменьшением длины очередей различие во временах доставки для вариантов

ВР2И и ВР2П будет уменьшаться.

В табл.1.2 представлены результаты экспериментов, проведенных для вариантов ВР2И и ВР2П при трех величинах нагрузки.

Отметим, что моделируемый метод был несколько грубее описанного выше, так как время  $\tau_i^k$  оставшееся до окончания передачи пакета, занявшего канал, округлялось до 1.

Таблица 1.2.

### Результаты расчета сети

Номер эксперимента	Варианты	Нагрузка	Число сообщений в системе	Время доставки $t_d$
3	ВР2И	10,5	53	4,2
4	ВР2П	10,5	56	4,3
5	ВР2И	12,0	402	6,8
6	ВР2П	12,0	504	25,0
7	ВР2И	-	-	-
8	ВР2П	12,5	2 235	49,8

Следующие два варианта ВР3И и ВР3П основаны на принципе рельефа и позволяют управлять распределением потоков пакетов с учетом ситуации на всей сети. При использовании этих вариантов соседние узлы обмениваются управляющей информацией и матрицы  $\|T_k\|$  строятся с учетом информации, полученной от соседних узлов.

Если бы матрицы  $\|T_k\|$  были построены с полным учетом мгновенной ситуации на сети, то должно было бы выполняться соотношение

$$t_{ij}^k = \min_{l \in I(k)} \{t_{il}^j + \tau_l^j + W_l^j + 1\} \quad (1.2)$$

В этом случае элемент  $t_{ij}^k$  представлял бы время, необходимое для доставки

пакета из узла  $k$  в узел  $j$ , если пакет отправлен в настоящий момент из  $k$  в соседний с ним узел  $j$ , а ситуация не изменится за все время прохождения пакета. Очевидно, что предположение о неизменности ситуации не выполняется, однако время  $t_{ij}^k$  можно рассматривать как некоторую оценку времени доставки.

Взаимодействие соседних узлов происходит аналогично тому, как это для сетей коммутации каналов.

Столбец с номером  $j$  в матрице узла  $k$  вычисляется в соседнем с узлом  $k$  узле  $j$  в соответствии с соотношением (1.2) и отправляется и узел  $k$ . В приводимых ниже результатах моделирования время  $\tau_j^k$  округлялось до 1.

Отметим, что поскольку время доставки управляющей информации конечно (будем обозначать его ТАУ), то передаваемая информация устаревает.

## ВЫВОДЫ

По первой главе можно сделать следующие выводы:

- при обмене информацией необходимо, обеспечение заданных вероятностно-временных характеристик доведения сообщений до потребителей, поскольку чрезмерные их задержки или потери могут привести к уменьшению эффективности управления;

- методы распределения потоков сообщений по кратчайшим путям позволяют определять пути, не только кратчайшие по длине, но и с минимальным числом транзитов, с минимальной стоимостью, с минимальными помехами и т.д;

- время запаздывания информации определяется главным образом суммарным временем ожидания в узлах сети и передачи по тракту связи, которое характеризуется выбором скорости работы аппаратуры;

- в зависимости от характера обслуживающей системы процесс обслуживания каждого требования либо должен протекать непрерывно, либо могут допускаться *прерывания* с последующим дообслуживанием требования любым из приборов.

- если каждый прибор может обслуживать любое из требований множества  $N$ ; каждое требование обслуживается не более чем одним прибором, и каждый прибор обслуживает не более одного требования, то совокупность функции  $s$ , обладающая перечисленными свойствами, называется *расписанием* обслуживания требований множества  $N$  системой, содержащей  $M$  параллельных приборов;

- одним из способов оценки *качества* расписаний обслуживания требований является представление каждого расписания вектором  $\bar{t}(s) = (\bar{t}_1(s), \dots, \bar{t}_n(s))$  моментов завершения обслуживания требований при этом расписании. задается действительная неубывающая по всем аргументам функция  $n$  переменных  $F(x) = F(x_1, x_2, \dots, x_n)$ . Качество расписании  $s$  характеризуется значением этой функции при  $x = \bar{t}(s)$ . Из двух расписаний лучшим считается то, которому соответствует меньшее значение  $F(x)$ . Расписание, которому соответствует наименьшее

значение  $F(x)$  (среди всех допустимых расписаний), называется *оптимальным* расписанием;

- качество расписания характеризуется суммарным или максимальным штрафом, который необходимо заплатить при обслуживании требований по расписанию;

- оптимальное расписание может быть найдено, как правило, в результате перебора конечного множества возможных вариантов. Основное затруднение состоит в том, что число таких вариантов обычно оказывается исключительно большим;

- для сравнения методов и изучения поведения сетей, управляемых с помощью различных методов и в разных условиях, можно воспользоваться методом моделирования.