Московский государственный технический университет
гражданской авиации
Факультет прикладной математики и вычислительной техники
Кафедра прикладной математики

Коновалов В.М.

Математические модели и алгоритмы управления информационными системами

ПОСОБИЕ

к выполнению курсовой работы

для студентов IV курса
специальности 230401
дневного обучения

Москва - 2009

Рецензент: проф. Кузнецов В.Л..

Коновалов В.М.

Пособие к выполнению курсовой работы по дисциплине "Математические модели и алгоритмы управления информационными системами". – М.: МГТУ ГА, 2009

Методическое пособие содержит необходимый для выполнения курсовой работы теоретический и практический материал. Издается в соответствии с Учебным планом подготовки студентов специальности 230401 - ”Прикладная математика”, специализации “Математическое и программное обеспечение информационных систем на воздушном транспорте” в рамках спецкурса “Математические модели и алгоритмы управления информационными системами”.

Рассмотрено и рекомендовано к изданию в РИО МГТУ ГА на заседаниях кафедры ПМ и методического Совета ФПМиВТ 24 02 2009 г.

СОДЕРЖАНИЕ

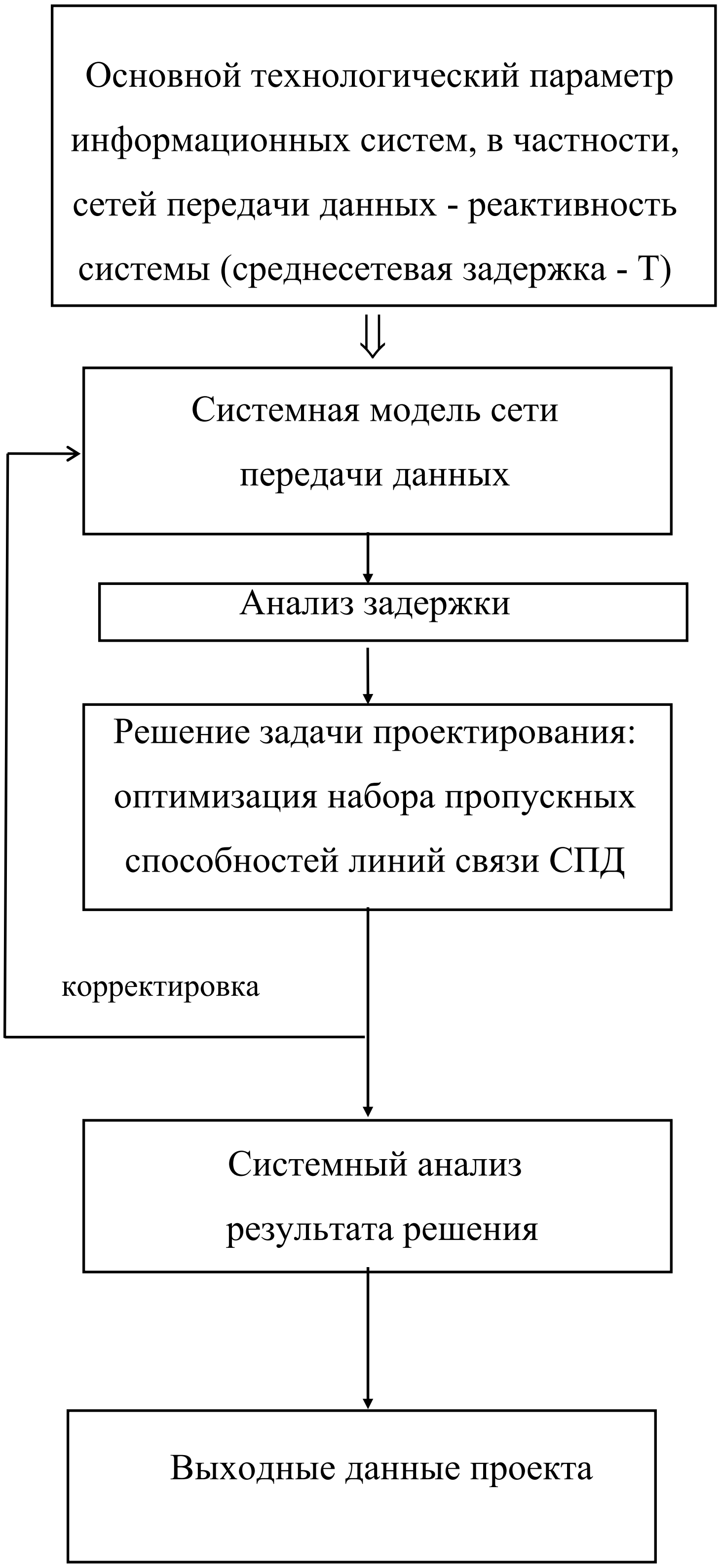
ПРЕДИСЛОВИЕ

Выполнение курсовой работы связано с практическим применением знаний и умений, полученных Вами при изучении дисциплин, связанных с математическим моделированием объектов, относящихся к классу информационных систем. В качестве представителя объектов этого класса для курсового проектирования выбрана сеть передачи данных (СПД), являющаяся важнейшей подсистемой информационно-вычислительной сети. Разработка модели СПД и процессов управления передачей данных в ней прямо или косвенно отражает практически все этапы технологии моделирования, начиная с выбора цели моделирования и заканчивая системным анализом результатов, полученных в ходе решения оптимизационной задачи.

В этой методической разработке Вы найдете необходимый для проектирования теоретический материал, примеры его использования, доведенные “до числа”, выходные формы для представления конечных результатов, используемых в последующем анализе. Удобным средством для ориентации в последовательности изложения может оказаться логическая схема, приводимая на следующей странице. Основные требования к содержанию и оформлению расчетно-пояснительной записки изложены в техническом задании на выполнение курсового проекта.

Следует обратить внимание на тот факт, что методу моделирования, как таковому, свойственна погрешность отображения объекта-оригинала. Поэтому, не должны оставаться без внимания те моменты в разработке модели, которые являются источниками вносимых погрешностей. В конечном счете их влияние может привести к недопустимой неадекватности модели оригиналу и к необходимости выполнить разработку с самого начала, начиная с выбора исходных данных.

ЛОГИЧЕСКАЯ СХЕМА ИЗУЧАЕМОГО МАТЕРИАЛА



1. ФОРМАЛИЗАЦИЯ ОПИСАНИЯ СЕТИ ДЛЯ ОПРЕДЕЛЕНИЯ ВРЕМЕННЫХ ХАРАКТЕРИСТИК ПЕРЕДАЧИ СООБЩЕНИЙ.

Объектом рассмотрения является распределенная сеть передачи данных (СПД) с коммутацией пакетов. Математической моделью топологической структуры такой сети является неориентированный граф, изображенный ниже.

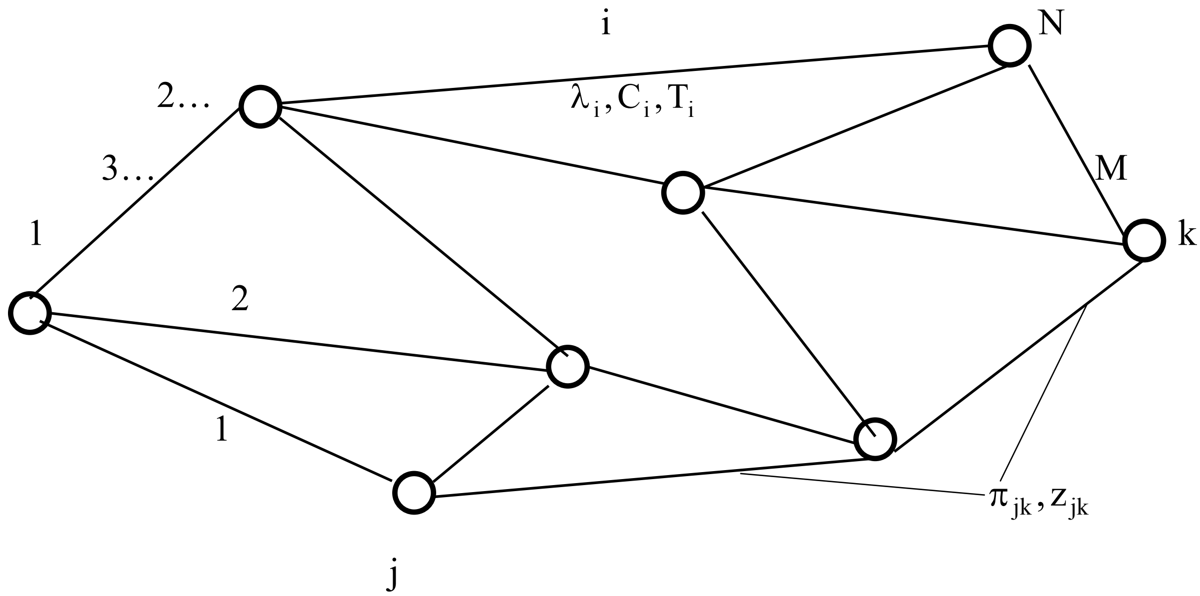


Рис.1

Примем следующие обозначения и условия, которыми в дальнейшем будем пользоваться при моделировании структуры СПД и процессов, связанных с передачей данных через сеть и представляющих непосредственный интерес для нашего рассмотрения.

Пусть

M - количество линий связи в сети;

N - количество узлов коммутации в сети.

Тогда математической моделью топологической структуры сети является граф $G = < V,E,q >$, где $V = \{ v_1, v_2, \dots, v_N \}$ - множество вершин графа;

$E = \{ e_1, e_2, \dots, e_M \}$ - множество ребер графа;

q - весовая функция на множестве ребер графа ,

$$q : E \to \mathbf{R}^{+}$$

$$e \mapsto q(e) = r^{+}\ ,\ r^{+} \in \mathbf{R}^{+}\ .$$

Допустим ряд упрощений, идеализирующих реальные каналы связи. Будем считать, что каналы связи бесшумны и абсолютно надежны.

C_i [бит/с] - пропускная способность линии i (i = 1,2, ... ,M), состоящей из пучка каналов связи. Пропускная способность определяет максимальную производительность канала по передаче данных.

Все N узлов сети абсолютно надежны и выполняют функции коммутации пакетов. Время обработки сообщений в узле пренебрежимо мало в сравнении со временем передачи по каналу связи. В силу конечности значения C_i, канал может иметь очередь на входе. Последнее является причиной возможных задержек при передаче сообщения через сеть.

Входящий в сеть трафик (внешний трафик сети) создает сетевую нагрузку от внешних источников - терминалов, генерирующих пуассоновские потоки запросов на передачу своих сообщений через сеть, и образует пуассоновский процесс.

γ [1/с] - интенсивность входного (в сеть) потока сообщений.

γ_{jk} [1/с] - интенсивность входного (в узел j сети) потока сообщений, требующих передачи в узел k сети.

В этом случае, вероятность появления в узле j за время t в точности s сообщений для передачи в узел k, определяется по формуле Пуассона :

$$P_s(t) = \frac{(\gamma_{jk} t)^s}{s!} \cdot e^{-\gamma_{jk} t \cdot}\ .$$

L - длина сообщения, представляющая собой последовательность бит, обрабатываемых в приемо-передающих модемах. Является случайной величиной, распределенной по показательному закону с параметром μ :

$$F(\ell) = P(\ L < \ell\ ) = 1 - e^{-\mu \ell}\ ;$$

$$f(\ell) = \mu \cdot e^{-\mu \ell}.$$

Понятно, что определенная таким образом “длина” сообщения косвенно характеризует длительность передачи по каналу, связанную с данным сообщением.

$1/\mu$ [бит] - средняя длина сообщения.

λ_i [1/c] - интенсивность потока (внутреннего) сообщений, передаваемых в линии i сети.

T [c] - среднее время задержки при передаче сообщения через сеть (среднесетевая задержка).

π_{jk} - путь передачи сообщения из узла j в узел k сети, состоящий из последовательности тандемных каналов.

Z_{jk} - средняя задержка передачи сообщения из узла j в узел k сети.

T_i - средняя задержка передачи сообщения по каналу i .

С учетом введенных обозначений, можно записать очевидные модельные соотношения для ряда характеристик.

Полный входящий в сеть поток сообщений (внешний поток) :

$$\gamma = \sum_{j=1}^{N} \sum_{k=1}^{N} \gamma_{jk} . \qquad (1)$$

Полный внутренний поток в сети, получающийся в результате работы процедур марщрутизации :

$$\Lambda = \sum_{i=1}^{M} \lambda_i . \qquad (2)$$

Среднесетевая задержка передачи сообщений :

$$T = \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\gamma_{jk}}{\gamma} \cdot Z_{jk} . \qquad (3)$$

В последнем выражении величина $\frac{\gamma_{jk}}{\gamma}$ - доля полного входящего трафика, передающегося по пути π_{jk} и имеющего задержку z_{jk}. Соотношение (3) представляет разложение среднесетевой задержки по парам “источник - адресат”.

Для передачи сообщения от источника к адресату будем считать, что в сети реализованы процедуры фиксированной маршрутизации (или маршрутизация от источника), при которой найденный кратчайший путь π_{jk} для передачи трафика из узла j в узел k не изменяется во времени.

Поток γ_{jk}, передаваемый по пути (маршруту) π_{jk}, называется маршрутным потоком (или (j-k)- трафиком).

Поток, передаваемый в линии i, называется линейным потоком и определяется выражением:

$$\lambda_i = \sum_j \sum_k \gamma_{jk} \quad , \qquad j,k : C_i \in \pi_{jk} . \tag{4}$$

Из последнего выражения следует, что линейный поток является суммой маршрутных потоков, проходящих через линию (канал) i .

Средняя задержка передачи сообщения по маршруту π_{jk} может быть определена , как сумма средних задержек в линиях (каналах), составляющих этот маршрут.

$$z_{jk} = \sum_i T_i \quad , \qquad i : C_i \in \pi_{jk} . \tag{5}$$

Условие $i : C_i \in \pi_{jk}$ означает, что маршрут π_{jk} проходит через канал i с пропускной способностью C_i.

2. ВЫВОД СООТНОШЕНИЙ ДЛЯ ОЦЕНКИ СРЕДНЕСЕТЕВОЙ ЗАДЕРЖКИ

В качестве основного технологического параметра для информационной системы, в данном случае таковой является распределенная СПД, используется какая-либо временная характеристика. В таком случае логично выбрать в качестве искомой характеристики среднесетевую задержку передачи сообщений - T.

Конечной целью курсовой работы является решение оптимизационной задачи – задачи структурного синтеза распределенной СПД , оптимальной по набору пропускных способностей линий связи сети. В качестве оптимизационного критерия (критерия эффективности) решаемой задачи будем использовать величину среднесетевой задержки. В качестве параметров-переменных будем использовать значения пропускных способностей линий связи - C_i. Необходимо получить явное выражение для критерия T, которое содержало бы переменную-параметр C_i. Такое выражение можно получить из соотношения (3), учитывая соотношение (5) :

$$T = \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\gamma_{jk}}{\gamma}\sum_{i} T_i = \sum_{i=1}^{M}\frac{T_i}{\gamma}\sum_{j}\sum_{k}\gamma_{jk} \quad , \quad i : C_i \in \pi_{jk} \ ; \ \ j,k : C_i \in \pi_{jk} \ .$$

Последние преобразования основаны на следующем. Изменен порядок суммирования. Поскольку ищется средняя оценка для всей сети, то внешнее суммирование ведется по всем каналам сети - от 1 до M. Для фиксированного i внутреннее суммирование ведется по тем (j - k)-маршрутам, которые включают в свой состав канал с пропускной способностью C_i.

Используя соотношение (4), получим :

$$T = \sum_{i=1}^{M}\frac{\lambda_i}{\gamma}T_i \quad . \qquad (6)$$

Выражение (6) называют разложением среднесетевой задержки на компоненты, относящиеся к отдельным каналам. Тем не менее, данное выражение является совершенно общим. Необходимо выразить среднесетевую задержку через конструктивные характеристики каналов сети, в частности, через пропускные способности каналов связи. Для этой цели рассмотрим канал связи, “глубоко погруженный в сеть”. Такой тип каналов характерен для распределенных сетей, топология которых имеет связность, большую единицы, т.е. узлы таких сетей имеют более, чем один входной канал и более, чем один выходной канал. В действительности топологические структуры сетей передачи данных имеют связность, большую единицы для обеспечения требуемой структурной надежности.

Канал, глубоко погруженный в сеть, можно рассматривать как канал с пуассоновским потоком на входе, интенсивность которого равна интенсивности, задаваемой сетью. Времена обслуживания сообщений отдельными каналами являются независимыми случайными величинами. В силу этого, интервалы между моментами поступления сообщений в канал не зависят от времени их обслуживания в предыдущих (тандемных) каналах.

Возможность подобной модели канала связи в сети обоснована Л. Клейнроком [л.1] и сформулирована в форме “предположения о независимости”. В поддержку такого решения выступает следующее рассуждение.

К зависимости интервалов между поступлениями сообщений в канал от времени их обслуживания в предыдущих каналах и, следовательно, к нарушению пуассоновости входного потока в канал, приводит поступление сообщений в данный узел в некоторой последовательности и уход к некоторому другому узлу в той же последовательности. Вместе с тем, если сообщения, уходящие из узла по данному каналу, пришли к узлу из разных каналов, или сообщения, пришедшие по одному и тому же каналу, отправляются из узла по разным каналам, или то и другое вместе, то можно считать, что указанная зависимость уменьшается.

Последнее согласуется с известным условием пуассоновского свойства случайного потока: чем больше источников-потребителей, тем ближе реальный поток к пуассоновскому. Таким образом, для каналов, глубоко погруженных в сеть, характеризуемых числом входов-выходов, большим единицы, применимо предположение о независимости. Используя это предположение, одиночный канал можно представить моделью в виде системы массового обслуживания (СМО) типа M | M | 1.

На входе i-ого канала - пуассоновский поток сообщений :

$$P_s(t) = \frac{(\lambda_i t)^s}{s!} e^{-\lambda_i t} \quad .$$

Время обслуживания - Θ сообщения каналом определяется длиной сообщения и распределено по показательному закону :

$$F_i(\vartheta) = P(\Theta < \vartheta) = 1 - e^{-\mu C_i \vartheta} \quad .$$

Почему так? Как уже говорилось выше, длина - L передаваемого сообщения является случайной величиной, распределенной показательно :

$$F(\ell) = P(\, L < \ell\,) = 1 - e^{-\mu \ell},$$

из чего следует, что средняя длина сообщений есть величина $1/\mu$ [ед.длины].

Теоретическая модель в виде СМО исходит из того, что пропускная способность СМО есть величина C=1[ед.скорости], например, 1[бит/с], 1[сообщ./с]. Вообще, не привязываясь к какой либо единице измерения длины, считают, что для СМО величина C=1[c^{-1}]. В таком случае среднее время обслуживания сообщений длины $1/\mu$ в СМО будет $1/\mu$: C = $1/\mu$[c]. Одновременно эта величина определяет средний интервал времени между моментами окончания обслуживания. Поэтому интенсивность обслуживания составит 1[c] : $1/\mu$[c] = μ. Теперь для СМО M | M | 1 функцию распределения времени обслуживания Θ можно записать в виде:

$$F(\vartheta) = P(\Theta < \vartheta) = 1 - e^{-\mu \vartheta} \quad .$$

В отличие от теоретической модели СМО, i-ый канал сети имеет вполне конкретное значение пропускной способности $C_i\,[c^{-1}]$, отличное от единицы. Поэтому среднее время обслуживания сообщений длины $1/\mu$ в канале будет в C_i раз меньше, чем в СМО, т.е. определяется величиной $1/\mu C_i$ [c]. Отсюда интенсивность обслуживания в канале сети будет μC_i, что и отражено выше в записи показательного закона для $F_i(\vartheta)$.

Средняя задержка T_i при передаче сообщения по каналу i определяется двумя составляющими :

$$T_i = [\text{время ожидания} + \text{время обслуживания}] = n\frac{1}{\mu C_i} + \frac{1}{\mu C_i} ,$$

где n - количество сообщений, стоящих в очереди перед данным сообщением. По формуле Литтла $n = \lambda_i T_i$. Тогда, подставляя значение n в выражение для T_i, получим :

$$T_i = \frac{\lambda_i T_i}{\mu C_i} + \frac{1}{\mu C_i} , \text{ откуда}$$

$$T_i = \frac{1}{\mu C_i - \lambda_i} . \tag{7}$$

Подставляя значение T_i в выражение (6), окончательно получим :

$$T = \sum_{i=1}^{M} \frac{\lambda_i}{\gamma}\left(\frac{1}{\mu C_i - \lambda_i}\right) . \tag{8}$$

3. АНАЛИЗ ЗАВИСИМОСТИ СРЕДНЕСЕТЕВОЙ ЗАДЕРЖКИ ОТ НАГРУЗКИ ДЛЯ СЕТИ СМО И РЕАЛЬНОЙ СПД

Моделирование объектов и процессов реального мира требует ясного представления о том, какова адекватность используемых в ходе разработки математических моделей реальным физическим процессам. Это означает, что необходим анализ условий, в которых получено то или иное решение.

Наша задача теперь состоит в анализе зависимости среднесетевой задержки от рабочих параметров и режимов функционирования сети. Для этой цели будем вести сравнительный анализ сети систем массового обслуживания и сети каналов передачи данных.

Выражение (7) характеризует среднюю задержку передачи сообщений по каналу. Аналогичная характеристика для СМО M | M | 1 может быть получена из (7), положив $C_i = C = 1[c^{-1}]$:

$$T_i = \frac{1}{\mu - \lambda_i} \quad . \tag{9}$$

Тогда среднесетевая задержка для сети СМО примет вид :

$$T = \sum_{i=1}^{M} \frac{\lambda_i}{\gamma}\left(\frac{1}{\mu - \lambda_i}\right) \quad . \tag{10}$$

По определению, нагрузка (загрузка) ρ системы является отношением интенсивности потока заявок на входе системы к интенсивности потока обслуженных заявок на выходе системы.

Для СМО M | M | 1 нагрузка есть отношение

$$\rho_i = \frac{\lambda_i}{\mu} \quad , \qquad 0 \le \rho_i < 1 \quad .$$

Подставив в формуле (9) вместо λ_i его эквивалент $\mu\rho_i$,получим зависимость $T(\rho)$ для сети СМО :

$$T(\rho) = \sum_{i=1}^{M} \frac{\lambda_i}{\mu\gamma}\left(\frac{1}{1 - \rho_i}\right) \quad . \tag{10a}$$

Аналогично, для канала СПД , имеем :

$$\rho_i = \frac{\lambda_i}{\mu C_i} \quad ,$$

а соответствующая зависимость $T(\rho)$ для сети каналов передачи данных примет вид :

$$T(\rho) = \sum_{i=1}^{M} \frac{\lambda_i}{\mu C_i \gamma} \left(\frac{1}{1-\rho_i} \right) . \tag{8a}$$

Анализируя соотношения (10а) и (8а), можно сделать количественные предсказания и некоторые выводы в отношении среднесетевой задержки сообщений в сети СМО и СПД.

В том случае, если множество $\{C_i\}$ относительно однородно, что характерно для сети СМО - все C_i=1, то при увеличении нагрузки на сеть никакое слагаемое в выражении для среднесетевой задержки не будет доминирующим. Кривая $T(\rho)$ плавно вписывается в область асимптотических значений T - область насыщения сети.

В том случае , если множество $\{C_i\}$ неоднородно, что характерно для СПД - каждый канал имеет собственное значение пропускной способности, возможна ситуация, когда поток в одном из каналов достигнет пропускной способности этого канала, являющегося “узким” местом в сети. В этом случае соответствующая компонента суммы в выражениях (8) или (8а) становится доминирующей вследствие бесконечного значения задержки передачи сообщений по данному каналу. В результате, вся сумма - среднесетевая задержка, бесконечно растет.

Таким образом, выражение для среднесетевой задержки во втором случае определяется компонентой T_i. Величина T имеет “порог”, до которого она остается относительно постоянной, а при приближении к нему резко возрастает.

Выводы :

- из соотношений (8а) и (10а) следует, что при $\rho \to 1$ среднесетевая задержка $T \to \infty$, что является следствием попытки использовать систему в режиме, близком к ее пропускной способности;

- более резкий переход в область асимптотических значений T для СПД в сравнении с сетью СМО обусловлен неоднородностью значений во множестве $\{C_i\}$;
- существует диапазон нагрузки на сеть, в котором среднесетевая задержка остается относительно постоянной; величина этого диапазона определяется степенью неоднородности во множестве $\{C_i\}$.

Пороговое поведение значений среднесетевой задержки приводит к упрощенной модели зависимости $T = T(\rho)$, что иллюстрируется на рис.2 .

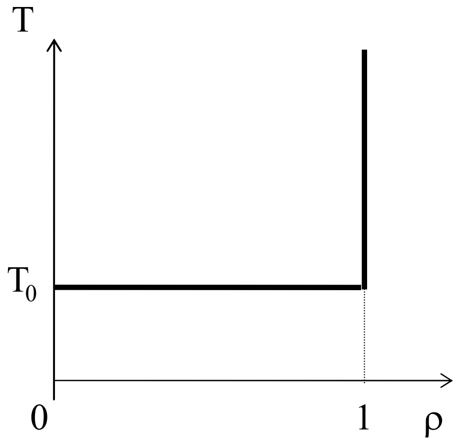


Рис.2

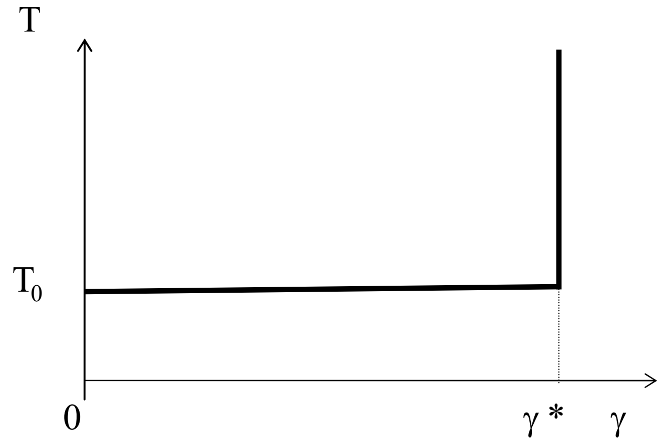


Рис.3

Существенным для такой модели является то, что среднесетевая задержка до некоторого порогового значения нагрузки относительно постоянна, а затем неограниченно возрастает. Здесь T_0- постоянная среднесетевая задержка, характеризующая область устойчивой работы сети, т.е. при $\rho < 1$.

Очевидно, нагрузка ρ сети определяется скоростью поступления сообщений в сеть, т.е. интенсивностью γ входного потока сообщений. Тогда можно перейти к зависимости $T = T(\gamma)$, которая будет отличаться от зависимости $T = T(\rho)$ лишь масштабом изображения (рис. 3).

В самом деле :

$$\rho = \frac{\lambda}{\mu C} = \frac{1}{\mu C}\sum_{i=1}^{M}\lambda_i = \frac{1}{\mu C}\sum_{i=1}^{M}\sum_{j}\sum_{k}\gamma_{jk} = \frac{1}{\mu C}H\gamma \ , \qquad j,k\!:C_i \in \pi_{jk} \ .$$

Здесь C - общая пропускная способность сети, т.е.

$$C = \sum_{i=1}^{M} C_i \ .$$

H - некоторая константа, характеризующая отношение полного внутреннего потока - λ в сети к внешнему потоку γ, обусловленная маршрутизацией потоков γ_{jk}.

В таком случае $\rho = \mathrm{const} \cdot \gamma$, а зависимости $T(\rho)$ и $T(\gamma)$ являются качественно идентичными. На рис.3 параметр γ^* представляет критическое значение скорости поступления сообщений в сеть, обусловливающее нагрузку насыщения сети, при которой $T \to \infty$.

4. РАСЧЕТ ЗНАЧЕНИЙ T_0 и γ^* В ОСОБЫХ ТОЧКАХ ПОРОГОВОЙ МОДЕЛИ ЗАВИСИМОСТИ $T(\gamma)$.

Обозначим через n_{jk} длину (в числе каналов связи) пути π_{jk}, определяемого процедурой маршрутизации в сети. Пусть также n - средняя длина пути в сети. Тогда можно записать :

$$n = \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\gamma_{jk}}{\gamma} n_{jk} \ . \qquad (11)$$

Определим внутренний поток λ_{jk} в сети, направляемый из узла j в узел k. Очевидно

$$\lambda_{jk} = \gamma_{jk} n_{jk} \ . \qquad (12)$$

Тогда полный внутренний поток в сети можно выразить соотношением :

$$\lambda = \sum_{j=1}^{N}\sum_{k=1}^{N}\lambda_{jk} = \sum_{j=1}^{N}\sum_{k=1}^{N}\gamma_{jk} n_{jk} \ . \qquad (13)$$

Тогда

$$n = \frac{\lambda}{\gamma} \ . \qquad (11a)$$

Из последнего равенства следует, что $\lambda = n\gamma$. Значит константа Н, введенная в рассмотрение выше и характеризующая соотношение между полным внутренним потоком λ в сети и внешним потоком γ, есть нечто иное, как средняя длина пути n, которая , в свою очередь, определяется топологией сети и процедурой маршрутизации.

Учитывая введенные обозначения, задержку T_0 будем вычислять как задержку в отсутствие нагрузки, используя предельную форму при γ, λ_i, стремящихся к нулю.

$$T_0 = \lim_{\lambda_i,\gamma\to 0}\sum_{i=1}^{M}\frac{\lambda_i}{\gamma}\left(\frac{1}{\mu C_i - \lambda_i}\right) = \sum_{i=1}^{M}\frac{1}{\mu C_i}\lim_{\lambda_i,\gamma\to 0}\frac{\lambda_i}{\gamma} =$$

$$\sum_{i=1}^{M}\frac{1}{\mu C_i}\lim_{\lambda_i,\gamma\to 0}\frac{\lambda_i}{\lambda} n = n\sum_{i=1}^{M}\frac{\lambda_i}{\lambda}\frac{1}{\mu C_i} \ . \qquad (14)$$

Последнее равенство возможно потому, что отношение λ_i / λ остается постоянным и не зависит от нагрузки в сети - γ. Оно является лишь функцией процедуры маршрутизации. Это подтверждается соотношением (11a), из которого следует: $\lambda = n\gamma$, т.е. при фиксированной (n = const) маршрутизации изменению внешнего потока в некоторое число раз соответствует изменение внутреннего потока во столько же раз, а отношение λ_i / λ будет постоянным, независящем от γ.

Значение γ^* находится из следующих рассуждений.

1. При заданной фиксированной процедуре маршрутизации, с помощью формулы (4) находится множество $\{\lambda_i\}$ потоков в сети при некотором значении γ.

2. Просматриваются все отношения $\lambda_i / \mu C_i$ и определяется некоторый индекс i_0 по наибольшему из этих отношений, характеризующий наиболее "узкое" место в сети, т.е. $\max_i \rho_i$:

$$\max_i \left(\frac{\lambda_i}{\mu C_i} \right) = \frac{\lambda_{i_0}}{\mu C_{i_0}} \quad .$$

3. Критическая точка, определяющая потерю устойчивого режима сети, описывается уравнением:

$$h\lambda_{i_0} = \mu C_{i_0} \ , \qquad h>1 \ .$$

Здесь h - коэффициент нагрузки, показывающий, во сколько раз можно увеличить действующую нагрузку в наиболее узком месте (канале i_0), чтобы сеть потеряла устойчивость ($T \to \infty$).

4. Так как для λ_i, λ и γ коэффициент нагрузки будет одним и тем же, то искомое значение γ^* находится в результате решения уравнения :

$$h = \frac{\mu C_{i_0}}{\lambda_{i_0}} \Big|\gamma = \gamma^* \ .$$

Окончательно находим искомое значение критической нагрузки :

$$\gamma^* = h\gamma \ . \tag{15}$$

Отметим, что рассмотренная двухпараметрическая пороговая модель для оценки среднесетевой задержки, описывает поведение сети в первом приближении. Использованный метод вычисления максимальной (критической) нагрузки на сеть не эквивалентен отысканию максимального потока, который может перенести сеть в действительности. Величина γ является несколько заниженной и вот почему.

В реальной сети передачи данных процедура маршрутизации обычно зависит от текущего трафика, т.е. является адаптивной. В результате, если какой-либо отдельный канал достигает точки насыщения, трафик направляется к месту назначения по другим путям в сети.

Тем не менее, оценка γ^* изложенным методом может оказаться полезной на этапе проектирования сети с целью выявления “узких” мест.

5. ФОРМАЛИЗАЦИЯ ПОСТАНОВКИ ЗАДАЧИ ВЫБОРА ОПТИМАЛЬНОГО НАБОРА ПРОПУСКНЫХ СПОСОБНОСТЕЙ ЛИНИЙ СВЯЗИ РАСПРЕДЕЛЕННОЙ СПД (ЗАДАЧА ВПС)

Задача оптимального выбора пропускных способностей (ВПС) линий связи сети из конечного набора пропускных способностей каналов связи, образующих соответствующую линию связи, является одной из наиболее трудных проблем проектирования (синтеза) сетей передачи данных.

Прежде всего отметим, что выпускаемая промышленностью каналообразующая аппаратура ориентирована на фиксацию конечного числа дискретных значений пропускной способности каналов связи, определяющих максимальные значения производительности канала по передаче данных. Так для цифровых каналов связи, используемых для организации сетей передачи данных, ряд стандартных значений пропускных способностей, к примеру, может быть таким: 64, 128, 192, 256, 320, 384,…,1984, 2048 [Кбит/с] с шагом, равным 64 .

Для получения оптимального набора пропускных способностей при решении задачи ВПС, будем считать, что они могут принимать значения не только из дискретного множества, но и иметь любые неотрицательные значения (из $\mathbf{R}^+$). В силу используемого допущения, результат решения задачи ВПС подлежит аппроксимации значениями из стандартного ряда пропускных способностей, причем погрешность аппроксимации должна быть минимальной. Иначе - непрерывное значение пропускной способности линии, полученное в ходе решения задачи, заменяется одним или суммой нескольких дискретных (стандартных) значений так, чтобы получить максимально близкую к непрерывному значению величину реальной пропускной способности линии связи.

Следующая особенность задачи связана со стоимостью сети. Будем считать, что стоимость сети определяется стоимостью эксплуатации линий связи. Это вполне согласуется с действительностью, так как стоимость эксплуатации линий связи во много раз превышает эксплуатационную стоимость коммутационного оборудования узла сети.

Обозначим

$S_i(C_i)$ - стоимость [ед.стоимости] эксплуатации линии i .

Тогда стоимость всей сети составит величину

$$D = \sum_{i=1}^{M} S_i(C_i) \ . \qquad (16)$$

В практике стоимостная функция $S(C)$ имеет дискретный характер и представляется набором значений стоимости, устанавливаемых действующей системой тарифов служб связи (рис.4).

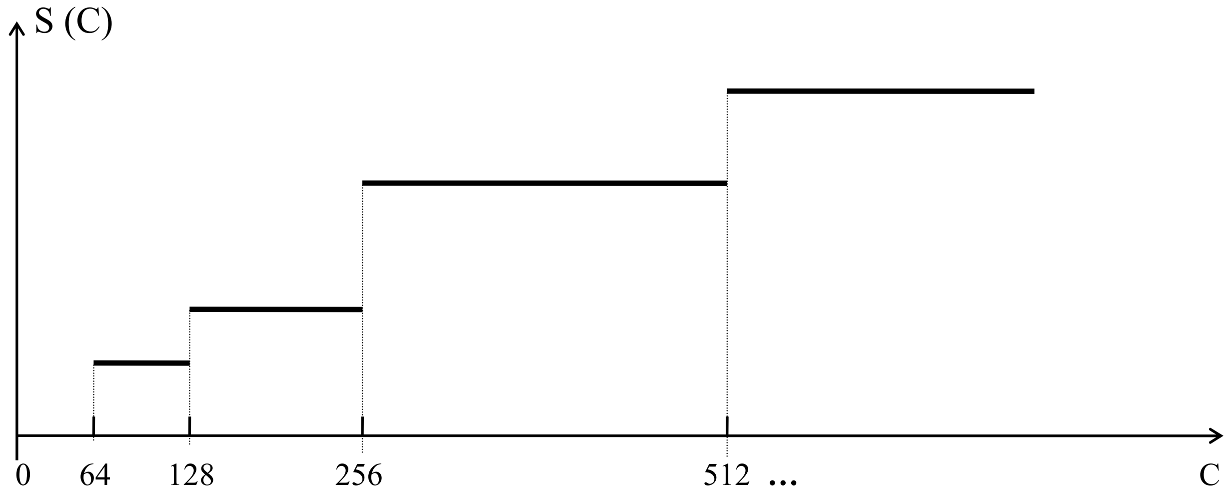


Рис.4

В целях решения задачи ВПС будем использовать непрерывную аппроксимацию стоимостной функции, имеющую вид линейной стоимостной функции пропускной способности:

$$S_i(C_i) = d_i C_i \quad , \qquad (17)$$

где d_i - стоимостной коэффициент, который может произвольно меняться в зависимости от какого-либо параметра линии, например, часто d_i зависит от физической длины линии, но он должен линейно зависеть от пропускной способности. В таком случае d_i имеет смысл стоимости единицы пропускной способности.

Графически теоретическая функция стоимости, представляющая линейную зависимость стоимости линии i от возможных значений ее пропускной способности, показана на рис.5.

Очевидно, переход от дискретной стоимостной функции к непрерывной порождает погрешность аппроксимации. Для ее уменьшения используют обобщения, выходящие за рамки линейных стоимостных функций, что позволяет найти более удачные непрерывные аппроксимации для истинного дискретного множества значений стоимостной функции.

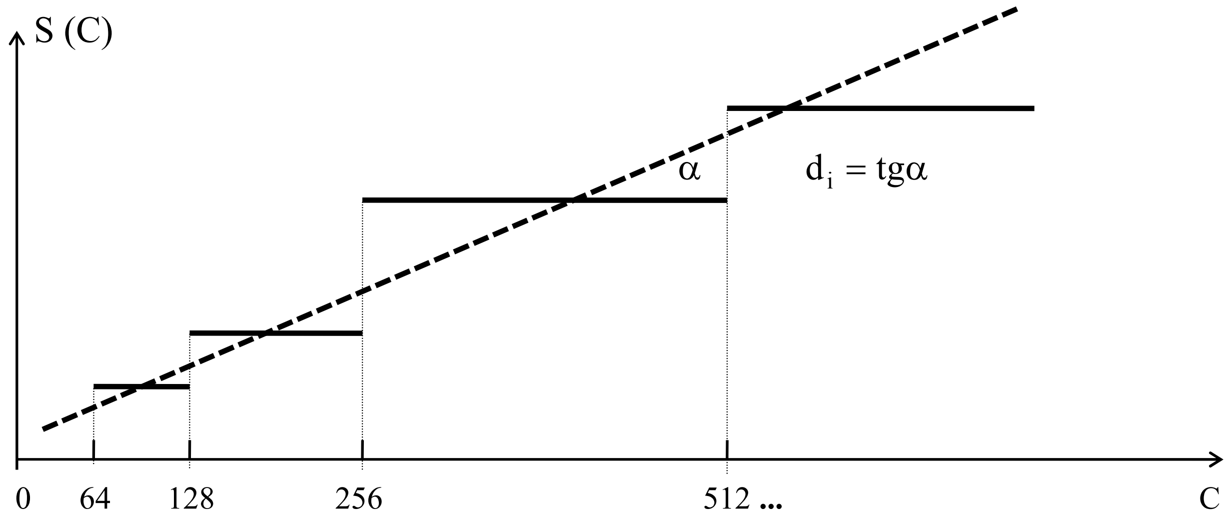


Рис.5

В качестве примера можно привести логарифмическую стоимостную функцию

$$D = \sum_{i=1}^{M} d_i \log \alpha C_i \quad ,$$

где α - константа, определяемая системой действующих тарифов служб связи.

Другой пример - степенная стоимостная функция:

$$D = \sum_{i=1}^{M} d_i C_i^{\alpha} \quad .$$

Теперь можно формализовать постановку задачи ВПС в следующим виде. Для заданных:

- топологии сети

$$TOP: \quad V \times V \to E$$

$$(v_i, v_j) \mapsto e_{ij} \ ;$$

- маршрутизации

$$SHP: \quad \gamma \to PATH$$

$$\gamma_{jk} \mapsto \pi_{jk} \quad ,$$

определяющих множество $\{\lambda_i\}$линейных потоков в сети, найти множество $\{C_i\}$ пропускных способностей линий связи сети, минимизирующих среднесетевую задержку

$$T = \sum_{i=1}^{M} \frac{\lambda_i}{\gamma} \left(\frac{1}{\mu C_i - \lambda_i} \right)$$

передачи сообщений через сеть, при ограничениях на стоимость используемой пропускной способности, которая определяется линейной стоимостной функцией

$$D = \sum_{i=1}^{M} d_i C_i \quad . \tag{16a}$$

Таким образом, имеем оптимизационную задачу, в ходе решения которой минимизируется скалярный критерий T. Значения переменных $\{C_i\}$, при которых достигается минимум T, являются оптимальными.

6. РАСЧЕТ ОПТИМАЛЬНЫХ ЗНАЧЕНИЙ C_i В СЛУЧАЕ ЛИНЕЙНОЙ ФУНКЦИИ СТОИМОСТИ

Решение задачи ВПС является реализуемым, если $C_i > \lambda_i / \mu$, $\forall i \quad (i = \overline{1, M})$, что следует из формулы (8).

Применим один из методов вариационного исчисления - метод неопределенных множителей Лагранжа. В соответствии с ним, для минимизации T составляется функция Лагранжа:

$$G = T + \beta \left[\sum_{i=1}^{M} d_i C_i - D \right] ,$$

где β - неопределенный множитель Лагранжа. Выражение в квадратных скобках учитывает соблюдение ограничений и тождественно равно нулю.

Поэтому минимизация функции T равносильна поиску минимума для G, в ходе которого определяется значение β. Минимум для G означает, что

$$\frac{\partial G}{\partial C_i} = 0 \;, \quad i = \overline{1, M} \;.$$

Дифференцирование приводит к выражению:

$$-\frac{\lambda_i}{\gamma} \cdot \frac{\mu}{(\mu C_i - \lambda_i)^2} + \beta\, d_i = 0 \;,$$

откуда

$$C_i = \frac{\lambda_i}{\mu} + \frac{1}{\sqrt{\beta\gamma\mu}} \left(\frac{\lambda_i}{d_i} \right)^{1/2} \;, \qquad i = \overline{1, M} \qquad (18)$$

Если найти постоянную β, то последнее выражение будет искомым решением. Для этого умножим обе части уравнения (18) на d_i и возьмем сумму по i :

$$\sum_{i=1}^{M} d_i C_i = \sum_{i=1}^{M} \frac{\lambda_i}{\mu} d_i + \frac{1}{\sqrt{\beta\gamma\mu}} \sum_{i=1}^{M} \sqrt{\lambda_i d_i} \;,$$

Левая часть последнего равенства есть D, поэтому

$$\frac{1}{\sqrt{\beta\gamma\mu}} = \frac{D - \sum_{i=1}^{M} \frac{\lambda_i}{\mu} d_i}{\sum_{i=1}^{M} \sqrt{\lambda_i d_i}} \;.$$

Определим величину

$$D_e = D - \sum_{i=1}^{M} \frac{\lambda_i}{\mu} d_i \qquad (19)$$

как “добавочную” стоимость. Смысл этого выражения станет понятным ниже. Используя два последних равенства, из уравнения (18) получим:

$$C_{i\,opt} = \frac{\lambda_i}{\mu} + \frac{D_e}{\sum_{i=1}^{M} \sqrt{\lambda_i d_i}} \cdot \sqrt{\frac{\lambda_i}{d_i}} \;.$$

Окончательно, оптимальное решение линейной задачи ВПС имеет вид:

$$C_{i\,opt} = \frac{\lambda_i}{\mu} + \frac{D_e}{d_i} \cdot \frac{\sqrt{\lambda_i d_i}}{\sum_{j=1}^{M} \sqrt{\lambda_j d_j}} \quad , \qquad i = \overline{1, M} \tag{20}$$

В формуле (20) числитель второго слагаемого-дроби относится к данному i, а знаменатель - ко всем индексам. Чтобы отразить этот факт, произведена замена индекса суммирования i на j в знаменателе.

Из выражения (20) следует, что каждая линия в сети будет иметь пропускную способность, по крайней мере, равную величине λ_i / μ - первое слагаемое, и кроме того, некоторую дополнительную пропускную способность, величина которой определяется вторым слагаемым. Следовательно, величину λ_i / μ можно рассматривать, как минимально необходимое значение пропускной способности линии i .

Стоимость минимальной пропускной способности для линии i есть величина $\lambda_i d_i / \mu$. Аналогичная величина для всей сети может быть представлена в виде:

$$\sum_{i=1}^{M} \frac{\lambda_i}{\mu} d_i \quad .$$

Очевидно, для того чтобы иметь возможность получить конечное значение среднесетевой задержки в сети, полная ее стоимость D должна быть больше вышевзятой суммы. В таком случае разность

$$D - \sum_{i=1}^{M} \frac{\lambda_i}{\mu} d_i = D_e$$

и будет той добавочной стоимостью, определенной выражением (19), которая направлена на поддержание требуемой характеристики сети в реальных условиях эксплуатации. Эта добавочная стоимость, после нормирования с помощью стоимостного коэффициента d_i (см. формулу (20)), распределяется по всем линиям сети пропорционально корню квадратному из взвешенной интенсивности трафика в линии - $\lambda_i d_i$.

Такой оптимальный набор пропускных способностей называется **набором пропускных способностей по правилу квадратного корня.**

Подставив выражение для C_i в формулу (8), получим экстремум (минимум) критерия оптимальности:

$$T_{min} = \frac{n}{\mu D_e} \cdot \left[\sum_{i=1}^{M} \sqrt{\frac{\lambda_i d_i}{\lambda}} \right]^2 . \qquad (21)$$

Здесь n - средняя длина пути в сети, определяемая по формуле (11а), как λ / γ .

При $D_e \to 0$, среднесетевая задержка передачи сообщений неограниченно возрастает.

При $D_e > 0$, задача ВПС имеет реализуемое решение, т.е. $T < \infty$.

При $D_e \le 0$, задача не имеет реализуемого решения.

Равенства (20) и (21) дают полное аналитическое решение задачи ВПС в случае линейной стоимостной функции.

7. ПРИВЕДЕНИЕ РЕЗУЛЬТАТОВ РЕШЕНИЯ ЗАДАЧИ ВПС К ВИДУ, УДОБНОМУ ДЛЯ СИСТЕМНОГО АНАЛИЗА

Важный частный случай, позволяющий провести системный анализ проблемы, состоит в том , что $d_i = d$, т.е. стоимость единицы пропускной способности одна и та же для любой линии сети. При этом, без потери общности, можно положить d = 1 ед.стоимости. В этом случае

$$D = \sum_{i=1}^{M} d_i C_i = \sum_{i=1}^{M} C_i \equiv C .$$

Здесь C - общая пропускная способность сети.

Из последнего равенства видно, что в случае линейной зависимости, сохранение общей стоимости D на фиксированном уровне будет эквивалентно поддержанию общей пропускной способности также на некотором постоянном уровне.

Для рассматриваемого частного случая формулы (20) и (21) примут, соответственно, вид:

$$C_{i\,opt} = \frac{\lambda_i}{\mu} + C(1 - n\,r)\frac{\sqrt{\lambda_i}}{\sum_{j=1}^{M}\sqrt{\lambda_j}} \quad , \qquad i = \overline{1, M} \;; \tag{20a}$$

$$T_{min} = \frac{n\left[\sum_{i=1}^{M}\sqrt{\frac{\lambda_i}{\lambda}}\right]^2}{\mu C(1 - n\,r)} \quad , \tag{21a}$$

где параметр

$$r = \frac{\gamma}{\mu C} \tag{22}$$

можно рассматривать, как нагрузку на сеть, создаваемую трафиком γ (внешним) на входах сети, в отличие от нагрузки $\rho = \lambda / \mu C$, создаваемой внутренним трафиком λ в сети.

Из равенств (20а) и (21а) следует, что

$$D_e = C(1 - nr) \; . \tag{23}$$

Действительно:

$$C(1 - nr) = C\left(1 - \frac{\lambda}{\gamma}\cdot\frac{\gamma}{\mu C}\right) = C - \frac{\lambda}{\mu} = C - \frac{\sum_{i=1}^{M}\lambda_i}{\mu} = D - \sum_{i=1}^{M}\frac{\lambda_i}{\mu} = D_e \;\Bigg|\; d = 1 \text{ ед.стоим.}$$

Из цепочки последних рассуждений также следует, что

$$D_e = C\left(1 - \frac{\lambda}{\mu C}\right) = C(1 - \rho) = D(1 - \rho) \; , \tag{23a}$$

так как, в силу рассматриваемого частного случая, $C = D$.

Кроме того, сопоставляя равенства (23) и (23а), видим, что

$$D_e = C(1 - nr) = C(1 - \rho) \ ,$$

откуда следует, что

$$\rho = n \cdot r \ . \qquad (23б)$$

Последнее означает, что нагрузка, создаваемая внутренним трафиком в n раз превышает нагрузку, создаваемую внешним трафиком. Это превышение обусловлено результатом работы процедуры маршрутизации, которая увеличивает, в среднем, в n раз длину пути передачи сообщений в сети по сравнению с сетью, имеющей полносвязную топологическую структуру.

8. СИСТЕМНЫЙ АНАЛИЗ РЕЗУЛЬТАТОВ РЕШЕНИЯ ЗАДАЧИ ВПС

Анализируя выражения (20а), (21а), можно сделать ряд полезных выводов.

1. T_{min} является строго возрастающей функцией средней длины n пути в сети. Следовательно, топология сети должна выбираться так, чтобы получить минимальную среднюю длину пути. Последнее, естественно, достигается в полносвязной сети, т.е. такой, где каждая пара узлов соединена линией связи.

2. Минимизация суммы в числителе выражения (21а) связана с минимизацией множества $\{\lambda_i / \lambda\}$. Сумма минимизируется по множеству $\{\lambda_i / \lambda\}$, когда одна из компонент равна единице, а все остальные - нулю.

Пример. Пусть имеется некоторое распределение потока λ по шести линиям связи:

$$\lambda_1 = \lambda_2 = \cdots = \lambda_5 = 4 \ , \ \lambda_6 = 16 \ ; \qquad \sum_{i=1}^{6} \lambda_i = 36 = \lambda \ .$$

При таком распределении сумма для числителя выражения (21а) примет значение:

$$\sum_{i=1}^{6}\sqrt{\frac{\lambda_i}{\lambda}} = 5\sqrt{\frac{4}{36}} + \sqrt{\frac{16}{36}} = \frac{10}{6} + \frac{4}{6} = 2\frac{1}{3} \ .$$

Перераспределим потоки, сохраняя при этом величину общего потока, следующим образом:

$\lambda_1 = \lambda$, $\lambda_2 = \lambda_3 = \cdots \lambda_6 = 0$.

Теперь сумма для числителя выражения (21а) примет минимальное значение:

$$\sum_{i=1}^{6}\sqrt{\frac{\lambda_i}{\lambda}} = \sqrt{\frac{36}{36}} + 5 \cdot 0 = 1 \to \min \ .$$

Следовательно, минимум суммы, при условии сохранения значения $\sum_i \lambda_i = \lambda$, достигается тогда, когда для некоторого i выполняется $\lambda_i = \lambda$, а для любого другого $j \neq i$ имеем $\lambda_j = 0$.

Физически это означает, что весь трафик следует направить по одной линии, а все остальные линии будут простаивать.

Естественно, это заведомо нереальный путь, однако минимизация суммы подсказывает, что надо большую часть трафика посылать по небольшому числу линий связи и малую его часть - по остальным линиям, т.е. здесь содержится идея концентрации трафика на небольшом числе линий связи.

Таким образом, второй теоретический экстремум может обеспечить топология, концентрирующая трафик на малом числе линий связи и обладающая, к тому же, невысоким значением средней длины пути. Последнее достигается в звездообразной (радиальной) топологии. В такой топологии можно получить относительно высокую степень концентрации трафика на каждой линии, а средняя длина пути n =2.

Следовательно, две идеи, характеризующие “крайние” случаи выбора топологии сети, приводят к следующим выводам:

1. Выбирать топологическую структуру сети с минимальной средней длиной пути - n .

2. Концентрировать поток на небольшом числе линий связи, а не распределять его.

В практических случаях следует делать выбор топологии от полносвязной до радиальной, используя промежуточные конфигурации, что определяется в таком случае значением величины r в выражении (22). При $r \to 0$ наиболее предпочтительна радиальная сеть, а при $r \to 1$ - полносвязная сеть.

Кроме рассмотренных, в основном интуитивных, суждений, связанных с топологией сети, анализ формулы для минимальной среднесетевой задержки T затрагивает вопрос выбора процедуры маршрутизации в сети.

Полученные формулы относились к фиксированным процедурам выбора маршрутов передачи сообщений. В отличие от них процедуры маршрутизации, допускающие альтернативы (адаптивная маршрутизация, к примеру), предлагают более, чем один единственный путь для трафика, направляемого к данному узлу. Подобные процедуры приводят к увеличению длины пути. Кроме того, они стремятся распределить трафик по большему числу линий связи, а не концентрировать его лишь на нескольких линиях.

Следовательно, при допускающем альтернативы выборе маршрутов, нарушаются оба вывода, которые следуют из анализа выражения (21а).

Исследования на основе имитационного моделирования приводят к выводу о том, что более предпочтительны фиксированные процедуры выбора маршрутов, чем процедуры, допускающие альтернативы. Данный вывод следует применять в условиях, когда величины γ_{jk} известны проектировщику и постоянны, что дает возможность, используя соотношение (4), определить линейные потоки в сети; после чего оптимально распределить ресурс пропускной способности сети в соответствии с внутренним трафиком, на основе равенства типа (20).

Если γ_{jk} либо неизвестны, либо изменяются во времени, что приводит к недопустимому рассогласованию с расчетными данными, то следует использовать допускающую альтернативы маршрутизацию, позволяющую приспособить трафик к неудачному набору пропускных способностей линий связи, путем поиска альтернативных путей передачи сообщений, содержащих линии с недогруженной пропускной способностью.

9. АЛЬТЕРНАТИВНЫЕ СТРАТЕГИИ РАСПРЕДЕЛЕНИЯ ПРОПУСКНОЙ СПОСОБНОСТИ В РЕШЕНИИ ЗАДАЧИ ВПС

Кроме стратегии распределения связного ресурса сети - пропускной способности, именуемой правилом квадратного корня, рассмотрим еще две альтернативы.

1. Стратегия равномерного распределения.

В этом случае общая пропускная способность сети распределяется поровну между всеми линиями, независимо от интенсивности трафика, проходящего по линии.

2. Стратегия пропорционального распределения.

Здесь величина C_i выбирается пропорционально значению λ_i :

$$C_{i\ prop} = C\frac{\lambda_i}{\lambda} \quad .$$

После этого, значение среднего времени задержки для линии i определяется по формуле (7), а значение среднесетевой задержки - по формуле (8).

Сравнительный анализ стратегий распределения пропускной способности приводит к следующим выводам.

1. При минимизации среднесетевой задержки по правилу квадратного корня слабо загруженным линиям соответствует меньшее значение пропускной способности, чем сильно загруженным. Поэтому время задержки слабо загруженных линий намного выше, чем сильно загруженных. Следовательно,

при использовании стратегии распределения пропускных способностей по правилу квадратного корня, пользователь, генерирующий слабый поток, дискриминируется в пользу пользователей со значительным трафиком.

2. Правило пропорционального выбора пропускной способности еще больше усиливает различие между слабо и сильно загруженными линиями. Неравенство между пропускными способностями больше, чем при выборе по правилу квадратного корня, при одновременном увеличении значения среднесетевой задержки.

3. Правило равномерного распределения хотя и ведет к увеличению среднесетевой задержки, но уменьшает разницу во времени задержки для слабо и сильно загруженных линий.

Завершая рассмотрение метода решения задачи ВПС, отметим, что изложенный выше подход позволяет оценить, какими должны быть пропускные способности линий связи, чтобы среднее время задержки передачи сообщений через сеть (реактивность сети) находилось в заданном диапазоне значений.

10. ПРИМЕР РЕШЕНИЯ ЗАДАЧИ ВПС ДЛЯ РАСПРЕДЕЛЕННОЙ СЕТИ

Для сети, конфигурация которой приведена на рис.6, проиллюстрируем основные моменты рассмотренного метода выбора пропускных способностей линий связи на основе стратегии квадратного корня с линейной функцией стоимости. Рассмотрим случай единичной, равной для всех линий, стоимости единицы пропускной способности:

$d_i = d = 1$ [ед.стоим.]

$S_i(C_i) = d_i C_i = d\,C_i = C_i$ [ед.стоим.] .

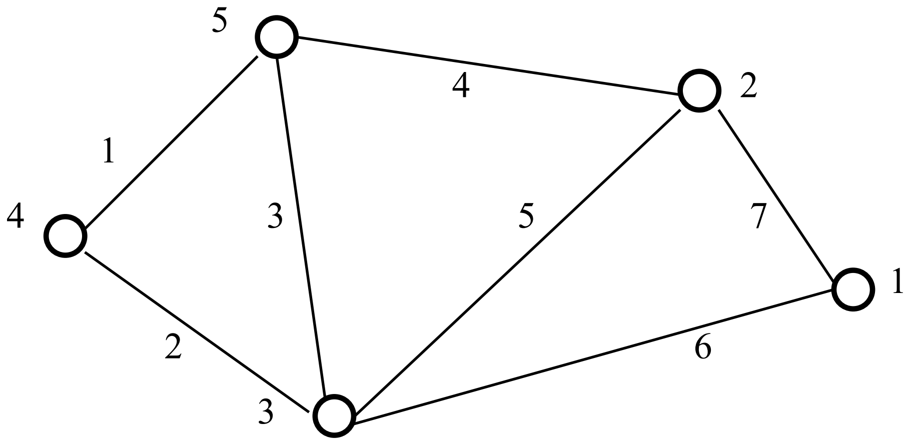


Рис.6.

Для выбранной топологии сети необходимо задать интенсивность (j-k)-трафика, передаваемого из узла j в узел k, а также процедуру маршрутизации сообщений, поступающих в узел-источник и направляемых в узел-адресат. Интенсивности γ_{jk} потоков внешнего трафика задаются потоковой матрицей Γ:

$$\Gamma = \left\|\gamma_{jk}\right\| = \begin{bmatrix} 0.000 & 9.340 & 0.935 & 2.940 & 0.610 \\ 9.340 & 0.000 & 0.820 & 2.400 & 0.628 \\ 0.935 & 0.820 & 0.000 & 0.608 & 0.131 \\ 2.940 & 2.400 & 0.608 & 0.000 & 0.753 \\ 0.610 & 0.628 & 0.131 & 0.753 & 0.000 \end{bmatrix}$$

Элемент γ_{jk} , $j,k = \overline{1,5}$, определяет интенсивность потока передаваемых сообщений от j-ого к k-ому узлу.

Все семь линий предполагаются дуплексными, с одинаковыми пропускными способностями для противоположных каналов одной и той же линии связи. Кроме того, трафик сети имеет симметричный характер - его средние

характеристики в каждом направлении на определенной линии одинаковы. Это позволяет рассматривать трафик, идущий только в одном направлении и, тем самым уменьшить размерность задачи: вместо вычисления времени задержки для 14 однонаправленных линий расчет достаточно провести для 7 линий.

Общая интенсивность трафика, выходящего из одного любого узла сети, есть сумма элементов соответствующей строки матрицы Г. Общая интенсивность трафика, входящего в один любой узел сети, есть сумма элементов соответствующего столбца матрицы Г. Общая интенсивность трафика, поступающего в сеть, есть сумма всех элементов матрицы Г.

Следовательно, для рассматриваемого примера интенсивность полного внешнего потока, поступающего в сеть, будет вдвое больше общей интенсивности однонаправленного трафика.

Значение пропускной способности дуплексной линии связи будет вдвое больше ее найденного значения для соответствующих однонаправленных (симплексных) линий.

Теперь необходимо выбрать процедуру маршрутизации. Будем считать, что используется процедура фиксированной маршрутизации, которая направляет каждое сообщение по “кратчайшему” в некотором смысле маршруту π_{jk}.

Термин “кратчайший” определяется тем, какой физический смысл имеет весовая функция $q: E \to \mathbf{R}^+$, определенная на множестве E ребер графа $G = < V,E,q >$. Пусть в данном примере “ вес ” каждого ребра графа есть географическая длина соответствующей линии связи сети. Тогда процедура маршрутизации направляет сообщения для передачи по наиболее короткому в географическом отношении маршруту.

Не вдаваясь в подробное описание используемой процедуры маршрутизации, отметим лишь, что для этой цели можно использовать эффективные алгоритмы поиска кратчайших путей на взвешенных графах такие, как алгоритм Дийкстры, алгоритм Флойда и др. Если используется алгоритм Дийкстры, то результатом работы процедуры маршрутизации является дерево кратчайших маршрутов (путей) из некоторой вершины графа (соответственно, узла сети) во все остальные.

Описание алгоритма Дийкстры (Dijkstra).

Вход: взвешенный ориентированный или неориентированный граф G с начальной вершиной v_0

$$G = < V, E, \mu, v_0 > \ .$$

Выход: помеченное дерево кратчайших путей с корнем v_0.

Метод:

Шаг 0. Пометим все вершины графа следующим образом. Положим $B(v_0) := 0$ для начальной вершины; $B(v) := \infty$ для всех остальных вершин $v \neq v_0$. Метку $B(v)$ назовем временной меткой.

Шаг1. Положим $D(v_0) := 0$; $S := \{v_0\}$, где S - множество помеченных меткой $D(v)$ вершин. Метку $D(v)$ будем называть постоянной. $D(v)$ означает расстояние от v_0 до v.

Положим $v_w := v_0$, где v_w - последняя (текущая) вершина, помеченная меткой $D(v)$.

Шаг 2. Для каждой, неотмеченной постоянно, вершины, т.е. для $v \in V \setminus S$ найдем новое значение временной метки $B(v)$, а именно

$$B(v) := \min(\, B(v), D(v_w) + \mu(v_w, v)\,),$$

где $\mu(v_w, v)$ - вес дуги $e = (v_w, v)$, если такая дуга имеется в графе G, и $\mu(v_w, v) = \infty$ - в противном случае.

Шаг 3. Если новое значение временной метки для каждой неотмеченной вершины есть ∞, то путей из отмеченных (т.е. из S) вершин в неотмеченные вершины не существует и алгоритм свою работу завершает.

В противном случае, из всех неотмеченных постоянно вершин выбираем вершину v_z с минимальной временной меткой $B(v_z)$, которая определяется как

$$B(v_z) = \min_{v \in V \setminus S} B(v).$$

Вершину v_z помечаем постоянной меткой

$$D(v_z) := B(v_z) .$$

Она становится новой последней, отмеченной постоянно, вершиной, т.е.

$$v_w := v_z \text{ и } S := S \bigcup \{v_z\}.$$

Кроме того, помечаем и вносим в список дуг дерева кратчайших путей дугу, для которой реализуется минимальное значение - $B(v_z)$.

Шаг 4. Если S = V, т.е. все вершины помечены постоянными метками $D(v)$, то алгоритм завершает работу. В противном случае перейти к шагу 2.

В результате работы алгоритма, на основе полученного списка дуг, формируется ориентированное дерево кратчайших путей, растущее из начальной вершины v_0, причем каждая вершина дерева имеет метку, означающую расстояние соответствующей вершины до начальной вершины.

Как следует из описания, алгоритм Дийкстры строит дерево кратчайших путей последовательно, присоединяя на каждом шаге одну вершину и дугу, заходящую в эту вершину, одновременно помечая присоединяемую вершину. Если последовательностью присоединяемых вершин является последовательность

$$v_0, v_1, v_2, \ldots, v_i, \ldots, v_N ,$$

то их метки - $D(v_i)$ будут располагаться в возрастающем порядке

$$0 = D(v_0) \le D(v_1) \le \ldots \le D(v_N) .$$

Присоединяемая на i -ом шаге вершина является i -ой, ближайшей к начальной, вершиной. Присвоение меток вершинам графа продолжается до тех пор, пока не будут помечены все вершины из связанного множества вершин.

Многократное (по числу узлов в сети) применение алгоритма поиска кратчайших путей дает множество списков кратчайших путей. Элемент этого множества есть список (дерево) кратчайших путей для отдельного узла во все остальные.

Пусть для рассматриваемого примера множество кратчайших путей будет таким, как показано на рис.7.

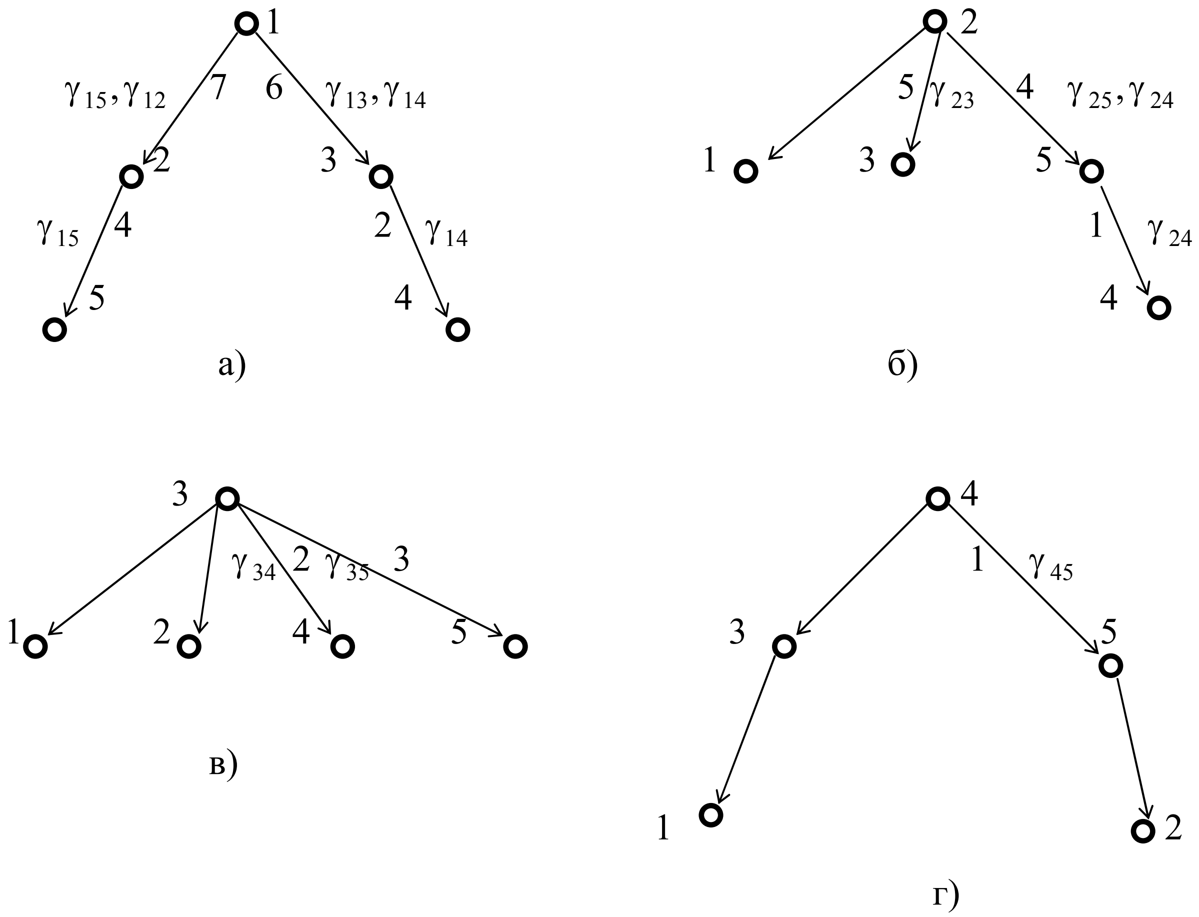


Рис.7.

Учитывая особенность, обусловленную симметрией трафика, дерево кратчайших путей для узла 5 можно не формировать, поскольку список узлов-адресатов, получаемых по одной из треугольных подматриц матрицы Г, к примеру, по верхней треугольной подматрице, будет пустым для узла 5. Если используется нижняя треугольная подматрица, то пустым будет список узлов-адресатов для узла 1.

Потоки γ_{jk}, передаваемые по путям π_{jk}, на рис.7 обозначены соответствующими значениями, а линии, входящие в состав таких путей, обозначены в соответствии с рис.6 номерами от 1 до 7.

Пользуясь исходными данными, определяемыми матрицей Г, получим значения интенсивностей внутреннего потока в сети, соответствующие занумерованным линиям.

$\lambda_1 = \gamma_{24} + \gamma_{45} = 3.153\ [c^{-1}]$

$\lambda_2 = \gamma_{14} + \gamma_{34} = 3.548$

$\lambda_3 = \gamma_{35} = 0.131$

$\lambda_4 = \gamma_{15} + \gamma_{25} + \gamma_{24} = 3.638$

$\lambda_5 = \gamma_{23} = 0.820$

$\lambda_6 = \gamma_{13} + \gamma_{14} = 3.875$

$\lambda_7 = \gamma_{12} + \gamma_{15} = 9.950$

Здесь отметим, что потоки γ_{24} и γ_{45} в линии 1 суммируются, несмотря на разнонаправленность. Фактически, при этом учитывается симметричный поток. Например, суммируя с потоком γ_{24} противоположный ему по направлению поток γ_{45}, тем самым учитывается поток γ_{54}, симметричный потоку γ_{45}.

Интенсивность полного потока в каждой линии будет вдвое выше полученных значений для однонаправленных потоков.

Интенсивность внутреннего потока в сети при однонаправленном трафике составляет величину

$$\lambda_{\frac{1}{2}} = \sum_{i=1}^{7} \lambda_i = 25.115\ [c^{-1}]\,.$$

Интенсивность внешнего трафика, порождаемого треугольной подматрицей матрицы Г, найдем суммированием:

$$\gamma_{\frac{1}{2}} = \sum_{j=1}^{4} \sum_{k=j+1}^{5} \gamma_{jk} = 19.165\ [c^{-1}]\,.$$

Средняя длина пути, по которому передается сообщение в сети, составит величину

$$n = \frac{\lambda}{\gamma} = \frac{25.115}{19.165} = 1.3 \ [\text{линии}] .$$

При экспоненциальном распределении длин сообщений средняя длина сообщения определится, как $1/\mu$. Пусть средняя длина поступающих на передачу сообщений равняется 200 [бит].

Для проведения расчетов по формулам (20а) и (21а) необходимо определить добавочную стоимость - D_e. С этой целью можно воспользоваться формулой (23а). Предварительно зададим величину нагрузки на сеть - ρ. Пусть $\rho = 0.6$.

Тогда ресурс общей пропускной способности, которую предстоит оптимально распределить по линиям сети, имеет значение

$$C = \frac{\Lambda}{\mu\rho} ,$$

что следует из формулы

$$\rho = \frac{\Lambda}{\mu C} .$$

В нашем случае это составит величину

$$C_{\frac{1}{2}} = \frac{25.115}{0.6} 200 = 8371.6 \, [\text{бит/с}] .$$

Теперь определим значение D_e:

$$D_{e\frac{1}{2}} = C_{\frac{1}{2}}(1-\rho) = 8371.6(1-0.6) = 3348.64 \, [\text{бит/с}].$$

На этом этапе завершается подготовка к расчету по формулам (20а) и (21а), и можно непосредственно вычислить значения $C_{i\,opt}$ и T_{min} для сети. Результаты расчетов сведем в таблицу (таблица 1).

Таблица1

Л и н и я			Линейный трафик	Правило выбора пропускной способности сети			
Номер	Узел			Квадратный корень		Пропорцион-е распред-е	
	нач.	кон	λ_i	C_i	T_i	C_i	T_i
1	4	5	3.153	1127.848	0.402	1050.992	0.476
2	3	4	3.548	1237.076	0.379	1182.658	0.423
3	3	5	0.131	127.555	1.972	43.666	11.494
4	2	5	3.638	1261.724	0.374	1212.658	0.412
5	2	3	0.820	417.582	0.789	273.331	1.828
6	1	3	3.875	1326.247	0.363	1291.658	0.387
7	1	2	9.950	2873.329	0.226	3316.644	0.151
				$T_{min}=$	0.444	$T_{prop}=$	0.546

Для сравнения в таблице1 представлены результаты расчета для одной из альтернативных стратегий - правила пропорционального распределения пропускных способностей.

Следующим моментом, на который необходимо обратить внимание, является аппроксимация расчетных значений пропускных способностей значениями из ряда стандартных значений.

Расчетное значение пропускной способности линии i заменяется одним или суммой стандартных значений из ряда так, чтобы погрешность аппроксимации была минимальной.

Погрешность аппроксимации пропускной способности i-ой линии - δ_{C_i} рассчитывается по формуле:

$$\delta_{C_i} = \frac{C_i - C_i^*}{C_i} \cdot 100\% \ , \tag{24}$$

здесь C_i^* - пропускная способность i-ой линии, составленная из дискретных значений стандартного ряда.

Определение C_i^* производится по любому алгоритму, который обеспечивает минимальное значение погрешности аппроксимации:

$$|\delta_{C_i}| \to \min .$$

Результатом аппроксимации расчетного значения пропускной способности является количество стандартных (в тональном диапазоне, в нашем случае) каналов связи, пучок которых составляет данную линию связи.

Данные аппроксимации сводятся в таблицу 2а для распределения пропускной способности по правилу квадратного корня и в таблицу 2б - для пропорционального распределения.

Таблица 2а

Л и н и я			Правило выбора пропускной способности сети				
Но-	У з е л		К в а д р а т н ы й к о р е н ь				
мер	нач.	кон	C_i	C_i^*	C_{stand}	Kol_{stand}	δ_{C_i}[%]
1	4	5	1127.848	1200	1200	1	-6.40
2	3	4	1237.076	1200	1200	1	3.00
3	3	5	127.555	600	600	1	-370.39
4	2	5	1261.724	1200	1200	1	4.89
5	2	3	417.582	600	600	1	-43.68
6	1	3	1326.247	1200	1200	1	9.52
7	1	2	2873.329	3000	2400,600	2	-4.41
			C = 8371.607	C^*= 9000			δ_C =-7.51

Отрицательные значения погрешности указывают на “перерасход” ресурса в сравнении с теоретически требуемым значением, что обусловлено дискретностью ряда стандартных аппроксимирующих значений пропускной способности.

Очевидно, при аппроксимации непрерывных значений дискретными необходимо сохранение условия реализуемости решения задачи ВПС, т.е. должно выполняться:

$$C_i > \lambda_i / \mu \ .$$

Для пропорционального распределения пропускной способности аппроксимирующие значения - C_i^* , в рассматриваемом примере, остаются такими же , как и для правила квадратного корня.

Таблица 2б

Л и н и я			Правило выбора пропускной способности сети				
Но-мер	У з е л		Пропорциональное распределение				
	нач.	кон	C_i	C_i^*	C_{stand}	Kol_{stand}	δ_{C_i}[%]
1	4	5	1050.992	1200	1200	1	-14.18
2	3	4	1182.658	1200	1200	1	-1.47
3	3	5	43.666	600	600	1	-1274.07
4	2	5	1212.658	1200	1200	1	1.04
5	2	3	273.331	600	600	1	-119.51
6	1	3	1291.658	1200	1200	1	7.10
7	1	2	3316.644	3000	2400,600	2	9.55
			C = 8371.607	C^*= 9000			δ_C =-7.51

Однако, пропорциональное распределение усиливает неравенство между пропускными способностями сильно и слабо загруженных линий связи, что отражается на распределении относительных погрешностей - δ_{C_i} .

Влияние аппроксимации на время передачи можно проследить, анализируя значения временных характеристик по таблицам 3а и 3б, соответственно для правила квадратного корня и пропорционального распределения.

В этих таблицах значения для T_i выбираются из соответствующих столбцов таблицы 1, а значения для T_i^* рассчитываются для каждой стратегии распределения пропускной способности по формуле (7), путем подстановки в нее соответствующих дискретных значений C_i^*, т.е.

$$T_i^* = \frac{1}{\mu C_i^* - \lambda_i} \ .$$

Для определения среднесетевой задержки T^*, полученной после аппроксимации дискретными значениями пропускных способностей, в каждом из рассматриваемых случаев стратегий распределения используется формула (6) в модификации:

$$T^* = \sum_{i=1}^{M} \frac{\lambda_i}{\gamma} T_i^*$$

Таблица 3а

Л и н и я			Правило выбора проп.спос-ти сети		
Номер	Узел		Квадратный корень		
	нач.	кон	T_i	T_i^*	δ_{T_i} [%]
1	4	5	0.402	0.351	12.69
2	3	4	0.379	0.408	-7.65
3	3	5	1.972	0.349	82.30
4	2	5	0.374	0.423	-13.10
5	2	3	0.789	0.459	41.83
6	1	3	0.363	0.471	-29.75
7	1	2	0.226	0.198	12.39
			T_{min} =0.444	T_{min}^* =0.434	δ_T =2.25

Относительная погрешность аппроксимации рассчитывается по формуле:

$$\delta_{T_i} = \frac{T_i - T_i^*}{T_i} \cdot 100\% \ .$$

Отрицательные ее значения указывают на ухудшение временных характеристик в результате аппроксимации.

Таблица 3б

Линия			Правило выбора проп.спос-ти сети		
Но-мер	Узел		Пропорцион-е распределение		
	нач.	кон	T_i	T_i^*	δ_{T_i} [%]
1	4	5	0.476	0.351	26.26
2	3	4	0.423	0.408	3.55
3	3	5	11.494	0.349	96.96
4	2	5	0.412	0.423	-2.67
5	2	3	1.828	0.459	74.89
6	1	3	0.387	0.471	-21.71
7	1	2	0.151	0.198	-31.13
			T_{prop}=0.546	T_{prop}^*=0.434	δ_T=20.51

Как следует из двух последних таблиц, наибольшую погрешность аппроксимация вносит в неоптимальную стратегию распределения. Следует отметить очень частный характер одного результата (полученного на основе рассмотрения лишь единственного примера). Для получения каких-либо обобщающих выводов, требуется определенная статистика.

Остается определить значения T_0 и γ^* для рассматриваемой сети.

Задержка передачи в условиях бесконечно малой нагрузки вычисляется в соответствии с формулой (14):

$$T_0 = n\sum_{i=1}^{7}\frac{\lambda_i}{\lambda}\cdot\frac{1}{\mu C_i} = 0.186\,[c^{-1}]\ .$$

Для определения γ^* предварительно установим, что

$$\max_i\left\{\frac{\lambda_i}{\mu C_i}\right\} = \frac{\lambda_7}{\mu C_7} = 0.692\ .$$

Нагрузочный коэффициент h находится из соотношения:

$$h\cdot\lambda_7 = \mu\cdot C_7\ ,$$

Откуда

$$h = \frac{\mu C_7}{\lambda_7} = \frac{2873.329}{200 \cdot 9.950} = 1.444\ .$$

Окончательно,

$$\gamma^* = h \cdot \gamma = 1.444 \cdot 19.165 = 27.674\ [c^{-1}]\ .$$

11. ТЕХНИЧЕСКОЕ ЗАДАНИЕ НА ВЫПОЛНЕНИЕ КУРСОВОЙ РАБОТЫ

ТЕХНИЧЕСКОЕ ЗАДАНИЕ

на выполнение курсовой работы по дисциплине
"ММ и АУИС" в VIII-ом семестре
кафедра "Прикладная математика" спец. 230401

Ф.И.О.____

Группа______________

Тема курсовой работы:

“Моделирование системы управления потоками данных в распределенной информационно-вычислительной сети с оптимизацией пропускных способностей линий связи коммуникационной подсистемы”.

Цель работы:
Получить расчетные данные проектирования, основываясь на общесистемной математической модели процесса передачи данных в информационно вычислительной сети и **индивидуальных** исходных данных.
В рамках этой темы студент выполняет задание с **индивидуальными** исходными данными, выбор которых должен удовлетворять приводимым ниже требованиям и условиям.

1. Выбрать топологию распределенной СПД. Число узлов - не меньше 10. Число линий связи - не меньше 15. Связность сети - не меньше 2.

Чертеж топологической структуры

2. Трафик (j-k) задается матрицей γ - требований на передачу, элемент которой γ_{jk} определяет интенсивность потока, требующего передачи из узла j в узел k.

Изображение матрицы требований

3. Линии связи - дуплексные. Если выбран симметричный трафик (что совсем необязательно), линии связи также можно считать симметричными по пропускной способности (т.е. в графе сети вместо дуг будут везде ребра).
4. В качестве “веса” линии (j-k) использовать значение трафика из узла j в узел k.
5. Маршрутизация в сети - фиксированная, задается списком путей (j-k) для передачи соответствующего трафика. Список путей (j-k) определить по дереву кратчайших путей.
6. Для построения дерева кратчайших путей передачи воспользоваться методом Дийкстры. Разработать схему алгоритма процедуры и программу.
7. Сформировать множество (по числу узлов в сети) списков кратчайших путей, многократно используя программу Dijkstra для каждого узла сети.
8. Разработать схему алгоритма процедуры и программу расчета значений линейных потоков в СПД. В качестве основы использовать соотношение (4).
Примечание: если выбран симметричный (j-k)-трафик в сети, то, полагая, что линии связи также симметричны по пропускной способности, линейные потоки рассчитываются на основе треугольной матрицы, т.е. от узла с меньшим номером к узлу с большим номером. При этом следует иметь ввиду, что полученный однонаправленный трафик составляет половину полного внутреннего трафика.
9. Выполнить расчет параметров T_0, γ^* и построение пороговой модели нагрузочной кривой – $T(\gamma)$.
10. Произвести расчет оптимальных значений пропускных способностей линий связи в случае линейной стоимостной функции, исходя из величины общей пропускной способности сети, определенной для заданной загрузки сети -.$\boldsymbol{\rho}$.
11. Выполнить сравнительный анализ полученных результатов решения задачи ВПС для **трех** стратегий распределения пропускных способностей:
A. стратегия распределения **по правилу квадратного корня**;
B. стратегия **пропорционального** распределения;
C. стратегия **равномерного** распределения.

12. Выполненную работу оформить по стандарту в виде расчетно-пояснительной записки (файл **MMKurs.doc**), в содержание которой наряду с иными рабочими материалами, должны быть включены:

- **топологическая структура** СПД (графическое изображение);
- **матрица требований -** γ на передачу потоков γ_{jk} через сеть;
- **схема алгоритма** маршрутизации (на основе метода Дийкстры);
- **схема алгоритма** процедуры расчета линейных потоков в сети;
- **тексты исходных модулей** программы с комментариями;
- **спецификации** переменных программы в форме таблицы соответствия имени и типа переменной в программе - физической величине;
- **"ручной"** расчет по методу Дийкстры **для одного** из узлов сети, используя табличное представление работы алгоритма; сопоставление с результатом работы программы для этого же узла сети с целью подтверждения правильности работы процедуры в программе;
- **деревья** кратчайших путей передачи данных с указанием маршрутных потоков в линиях (выполнить графически);
- **таблицы** с результатами решения задачи ВПС (таблицы по форме №№ 1, 2а, 2б, 3а, 3б - см. метод.пособие); точность представления данных в таблицах - **3** знака после запятой.
- **графические зависимости** T(γ) для каждой из сравниваемых стратегий выбора пропускных способностей (разместить на одной и той же (!) координатной плоскости) с изображением рабочей точки сети на соответствующих нагрузочных кривых.
- **сводная таблица** исходных данных и полученных в процессе выполнения курсовой работы результатов - значений параметров сети (в форме спецификации).
- аналитические **выводы** по результатам решения оптимизационной задачи.

Рекомендуемая литература.

1. Клейнрок Л. Вычислительные системы с очередями. М.: Мир, 1979.

2. Шварц М. Сети ЭВМ: анализ и проектирование. М.: Радио и связь, 1981.

3. Свами М., Тхуласираман К. Графы, сети и алгоритмы. М.: Мир, 1984.

4. Бертсекас Д., Галагер Р. Сети передачи данных. М.: Мир, 1989.