УДК 621.34

С.В. Савченко[1], В.А. Краснобаев[2]

[1]*Национальный технический университет «ХПИ», Харьков*
[2]*Харьковский национальный технический университет сельского хозяйства им. П. Василенко*

ОПТИМИЗАЦИЯ ЦИФРОВОЙ СЕТИ ИНТЕГРИРОВАННОГО ОБСЛУЖИВАНИЯ ПО КРИТЕРИЮ СТОИМОСТИ ПЕРЕДАЧИ ЕДИНИЦЫ ИНФОРМАЦИИ

В статье рассмотрены вопросы аналитического решения задачи оптимизации цифровых сетей интегрированного обслуживания по критерию стоимости передачи единицы информации в результате определения числа элементов буферной памяти и соответствующего оптимального значения плотности потока информации, обеспечивающего минимальную среднюю задержку передачи сообщений в сети.

оптимизация цифровых сетей интегрированного обслуживания, критерий стоимости передачи единицы информации

Анализ литературы и постановка задачи исследования

Основной причиной перегрузок сети является конечное число буферов в узлах коммутации и ограниченность канального ресурса, связанная со стоимостью аренды каналов. Анализ литературы [1 – 4] показывает, что постановка задачи с неограниченными узловыми ресурсами дает завышенные значения среднего времени задержки, а ограничение объема буферов при прочих равных условиях позволяет транспортировать по сети более значительные потоки информации [5]. При этом ограничение объема буферной памяти сокращает среднее время задержки за счет уменьшения времени пребывания пакета в очереди, однако ограничивает возможность передачи длинных сообщений методами с промежуточным накоплением, когда длина сообщения может превысить объем буферной памяти, необходимой для обеспечения оптимального решения. Выходом из этого положения является использование комбинированных методов коммутации, например, в пакетной сети длинные сообщения транспортировать методом коммутации каналов (гибридная коммутация). С другой стороны попытка решить данную задачу для ограниченного числа мест в очереди не позволяет получить строгое аналитическое решение при ограничении стоимости аренды каналов связи $D \le D_{зад}$ из-за сложности полученного функционала оптимизации для известных видов функции стоимости.

Результаты исследований

С целью упрощения функционала оптимизации и обеспечения возможности получения аналитического решения задачи предлагается использовать в качестве ограничивающего условия стоимость передачи количества информации, приходящуюся на единицу пропускной способности. Тогда суммарное количество переданной информации будет определять общий доход от использования средств связи. Общая стоимость сети может определяться на заключительном этапе проектирования, вычисляемая, например, как сумма пропускных способностей. Такой подход позволит определить сроки окупаемости сети с учетом затрат на ее реализацию и доход от ее эксплуатации.

На основании формулы Литтла, согласно клейнроковской аппроксимации [4], очереди пакетов на входе в каждый канал связи представим СМО типа M/M/1 с ожиданием. То есть на вход i-ой очереди поступает пуассоновский поток пакетов с интенсивностью λ_i пакетов в секунду и средним временем обслуживания μ_i в секунду, распределенным по экспоненциальному закону. Учитывая коэффициент использования канала (загрузки сети) $\rho_i = \lambda_i / \mu_i$, относительная пропускная способность (средняя доля поступивших заявок, обслуживаемая системой) определяется [6] как

$$\overline{g}_i = \frac{1-\rho_i^{m_i+1}}{1-\rho_i^{m_i+2}}, \qquad (1)$$

где m_i – число мест в очереди.

Среднее число заявок в очереди определяется как:

$$\overline{r}_i = \frac{\rho_i^2\,[1-\rho_i^{m_i}\,(m_i+1-m_i\cdot\rho_i)]}{(1-\rho_i^{m_i+2})\,(1-\rho_i)}. \qquad (2)$$

Среднее время задержки $\overline{T}_i$, равное времени пребывания заявки в СМО, выражается общей формулой [5]:

$$\overline{T}_i = \overline{r}_i/\lambda_i + \overline{g}_i/\mu_i. \qquad (3)$$

С учетом (1) и (2) выражение (3) после преобразования принимает вид:

$$\overline{T}_i = \frac{1}{\mu_i}\frac{1-\rho_i^{m_i+1}[(m_i+2)-\rho_i(m_i+1)]}{(1-\rho_i^{m_i+2})(1-\rho_i)}. \quad (4)$$

Обозначим

$$\sum_{k=0}^{m_i+1}\rho_i^k = \frac{1-\rho_i^{m_i+2}}{1-\rho_i} = \Sigma_{m_i},$$

где Σ_{m_i} – сумма геометрической прогрессии.

Выражение (4) представим в преобразованном виде:

$$\overline{T}_i = \frac{1}{\mu_i}\cdot\frac{\sum_{\alpha=0}^{m_i}(1+\alpha)\rho_i^{\alpha}}{\sum_{\alpha=0}^{m_i+1}\rho_i^{\alpha}} = \frac{1}{\mu_i}\cdot\frac{\left(\sum_{\alpha=0}^{m_i+1}\rho_i^{\alpha}\right)'}{\sum_{\alpha=0}^{m_i+1}\rho_i^{\alpha}} = \frac{1}{\mu_i}\frac{\Sigma'_{m_i}}{\Sigma_{m_i}}, \quad (5)$$

где Σ' обозначает производную $\frac{\partial\Sigma}{\partial\rho}$.

Среднее время задержки во всей сети находится с учетом (5):

$$\overline{T} = \frac{1}{\gamma}\sum_{i=1}^{k}\rho_i\frac{\left(\sum_{\alpha=0}^{m_i+1}\rho_i^{\alpha}\right)'}{\sum_{\alpha=0}^{m_i+1}\rho_i^{\alpha}} = \frac{1}{\gamma}\sum_{i=1}^{k}\rho_i\frac{\Sigma'_{m_i}}{\Sigma_{m_i}}. \quad (6)$$

Зависимость (6) при $m_i \to \infty$ преобразуется в известную формулу $\overline{T} = \frac{1}{\gamma}\sum_{i=1}^{k}\frac{\rho_i}{1-\rho_i}$, что соответствует модели сети в виде СМО с неограниченной очередью [5], а при $m_i = 0$ в $\overline{T} = \frac{1}{\gamma}\sum_{i=1}^{k}\frac{\rho_i}{1+\rho_i}$, что соответствует модели сети в виде СМО с отказами [7].

Функция (6) является выпуклой функцией, но не содержит экстремумов, что не позволяет найти минимум среднего времени задержки путем вычисления частных производных $\frac{\partial\overline{T}}{\partial\rho_i} = 0$. Таким образом, данная задача является задачей условной оптимизации. Аналитическое решение поставленной задачи возможно при соответствующем выборе в качестве ограничивающего условия функции стоимости. Числовые расчеты [5] показывают, что обычно нет большого различия между случаями использования стоимостных функций того или иного вида, то есть следует выбирать ту стоимостную функцию, которая наиболее полно соответствует условиям конкретной задачи. Рассмотрим стоимостную функцию [1] вида

$$D = \nu\sum_{i=1}^{k}\frac{F_i}{V_i}, \quad (7)$$

где при пакетной передаче сообщений $F_i = L\lambda_i$, $V_i = L\mu_i$; L – фиксированная длина пакета (бит). Тогда функция стоимости (7) принимает вид

$$D = \nu\sum_{i=1}^{k}\rho_i \quad (8)$$

и выражается в единицах стоимости передачи единицы количества информации, то есть плотности потока информации, что соответствует принятым принципам оплаты за использование средств связи.

Таким образом, оптимизационная задача может быть сформулирована в следующем виде: определить оптимальные значения плотности информационного потока, минимизирующего среднюю задержку

$$\overline{T} = \frac{1}{\gamma}\sum_{i=1}^{k}\rho_i\frac{\Sigma'_{m_i}}{\Sigma_{m_i}} \to \min \quad (9)$$

при ограничении на стоимость передачи суммарного количества информации, приходящейся на единицу пропускной способности линий связи

$$D = \nu\sum_{i=1}^{k}\rho_i \le D_{зад}. \quad (10)$$

Для решения данной задачи применен метод неопределенных множителей Лагранжа. Составим функционал оптимизации:

$$\Phi = \frac{1}{\gamma}\sum_{i=1}^{k}\rho_i\frac{\Sigma'_{m_i}}{\Sigma_{m_i}} + P\nu\sum_{i=1}^{k}\rho_i, \quad (11)$$

где P – неопределенный множитель Лагранжа.

Вычисляя частные производные $\frac{\partial\Phi}{\partial\rho_i} = 0$, получим систему из k уравнений вида

$$\left(\rho_i\frac{\Sigma'_{m_i}}{\Sigma_{m_i}}\right)' + \gamma P\nu = 0, \quad i = \overline{1,k}. \quad (12)$$

Анализ выражения (12) показывает, что каждое уравнение, этой системы зависит от переменной ρ_i и параметров m_i, γ, P, ν. Если положить $m_i = m$ одинаковым для всех узлов, то эти параметры не зависят от индекса i, т.е. в результате его решения относительно ρ_i, получаем $\rho_i = F(m, \gamma, P, \nu)$. Это позволяет сделать вывод, что $\rho_i^{опт} = \rho = const$, т.е. оптимальные значения плотностей потока информации одинаковы для всех ветвей и не зависят от номера ветви связи (изотропная сеть).

После дифференцирования и некоторых преобразований получим дифференциальное уравнение второго порядка для каждой ветви. Опуская индекс i, имеем

$$\frac{\Sigma'_m}{\Sigma_m} + \rho\frac{\Sigma''_m}{\Sigma_m} - \rho\frac{(\Sigma'_m)^2}{\Sigma_m^2} + \gamma P\nu = 0. \quad (13)$$

Путем замены переменной

$$\frac{\Sigma'_m}{\Sigma_m} = Z \text{ и } \Sigma''_m = Z'\Sigma_m + Z\Sigma'_m \quad (14)$$

уравнение (13) преобразуется в неоднородное линейное уравнение 1-го порядка

$$Z' + \frac{1}{\rho}Z = -\frac{1}{\rho}\gamma P\nu\,. \qquad (15)$$

Общее решение уравнения (15) находится методом вариации произвольной постоянной [8, 9]. Соответствующее однородное уравнение

$$Z' + \frac{1}{\rho}Z = 0\,, \qquad (16)$$

с разделяющимися переменными имеет общее решение в следующем виде:

$$Z = \frac{a_1}{\rho}\,. \qquad (17)$$

Положим $a_1 = a_1(\rho)$ некоторой непрерывно дифференцируемой функцией от ρ, тогда

$$Z = \frac{a_1(\rho)}{\rho}\,. \qquad (18)$$

Выберем функцию $a_1(\rho)$ так, чтобы выражение (18) удовлетворяло уравнению (15). Подставляя (18) в (15) после преобразований, получим

$$a_1(\rho) = \ \gamma P\nu\,. \qquad (19)$$

Интегрируя (19), имеем $a_1(\rho) = -\gamma P\nu\rho + a_2$ и, следовательно

$$Z = -\gamma P\nu + \frac{a_2}{\rho}\,. \qquad (20)$$

Возвращаясь к старой переменной (15), получим

$$\partial\Sigma'_m/\Sigma_m = -\gamma P\nu + (a_2/\rho)\,. \qquad (21)$$

Разделяя в выражении (21) переменные, придем к следующему уравнению $\frac{d\Sigma_m}{\Sigma_m} = -\gamma P\nu\, d\rho + a_2\frac{d\rho}{\rho}$, интегрируя которое, окончательно получаем

$$\Sigma_m = a_3\rho^{c_2}\cdot e^{\ \gamma P\nu\rho}\,, \qquad (22)$$

где a_3 – постоянная интегрирования 2-й квадратуры.

По уравнению (22) находим произвольные постоянные интегрирования a_2 и a_3 путем решения задачи Коши при заданных начальных условиях.

В дальнейших расчетах ограничимся зависимостью (22), из которой определены значения ρ для каждой ветви рассматриваемой сети

$$\rho = \frac{a_2}{\gamma P\nu + \left(\Sigma'_m/\Sigma_m\right)}\,. \qquad (23)$$

Определим значения a_2 из начального условия $\rho_o = 1$ и уравнения (6)

$$\left.\left(\Sigma'_m/\Sigma_m\right)\right|_{\rho=1} = \frac{1+2+\cdots+(m+1)}{m+2} = \frac{(m+1)(m+2)}{2(m+2)} = \frac{m+1}{2}. \qquad (24)$$

В выражении (24) учтено, что числитель $1+2+\ldots+(m_i+1) = \frac{(m+1)(m+2)}{2}$ – есть сумма арифметической прогрессии.

Из уравнения (23) при условии (24) определяем произвольную постоянную a_2

$$a_2 = \gamma P\nu + \frac{m+1}{2}\,.$$

Окончательно получаем

$$\rho = \left(\gamma P\nu + \frac{m+1}{2}\right)\Big/\left(\gamma P\nu + \left(\frac{\Sigma'_m}{\Sigma_m}\right)\right)\,. \qquad (25)$$

Для определения неопределенного множителя Лагранжа воспользуемся условием (10) для предельного значения стоимости

$$\nu\sum_{i=1}^{n}\frac{\gamma P\nu + (m+1)/2}{\gamma P\nu + \left(\Sigma'_m/\Sigma_m\right)} = \nu n\,\frac{\gamma P\nu + (m+1)/2}{\gamma P\nu + \left(\Sigma'_m/\Sigma_m\right)} = D_{зад}\,.$$

После преобразований получаем значение множителя Лагранжа P:

$$P = \frac{(m+1)/2 - \left(\Sigma'_m/\Sigma_m\right)\cdot\left(D_{зад}/\nu n\right)}{\gamma\nu\left(\left(D_{зад}/\nu n\right)-1\right)}\,. \qquad (26)$$

Подставляя (26) в (25), получим условия экстремумов $\overline{T}_i$ выражения (6):

$$\left(\rho_{опт} - D_{зад}/\nu n\right)\left((m+1)/2 - \Sigma'_m/\Sigma_m\right) = 0\,. \qquad (27)$$

Условия (27) выполняются, если любой их сомножитель равен нулю, то есть

$$\rho_{опт} - D_{зад}/k\nu = 0\,; \qquad (28)$$

$$\Sigma'_m/\Sigma_m - (m+1)/2 = 0\,. \qquad (29)$$

Условие (28) определяет оптимальное значение удельного потока в ветвях

$$\rho_{опт} = D_{зад}/\nu n\,, \qquad (30)$$

обеспечивающее минимальное значение среднего времени задержки:

$$\overline{T}^{min} = \frac{1}{\gamma}\cdot\frac{D_{зад}}{\nu}\cdot\left(\frac{\Sigma'_m}{\Sigma_m}\right)_{опт}\,,\ 0 < \rho_{опт} < 1. \qquad (31)$$

Условие (29) соответствует максимальному значению задержки в сети при $\rho = 1$, $\overline{T}^{max} = (n/\gamma)\cdot(m+1)/2$, которое достигается независимо от стоимости сети.

Анализ полученных результатов (31) показывает, что стоимость сети определяется в основном затратами на передачу данных, поэтому необходимо ресурсы сети использовать максимально эффективно. Согласно выражению (30) можно сделать вывод, что сеть связи должна быть изотропной в смысле постоянства значений плотности потока передаваемой информации во всех линиях связи ($\rho_{опт} = D_{зад}/kn < 1$ не зависит от номера ветви связи).

Если потоки в ветвях при синтезе сети заданы в виде матрицы тяготений $\|\lambda_i\|$, то при фиксированной длине пакета L пропускные способности соответствующих ветвей прямо пропорциональны значениям потоков этих ветвей, то есть

$$V_i = \left(\nu n / D_{зад}\right) \cdot F_i , \qquad (32)$$

что является необходимым условием исключения блокировок сети (то есть $V_i > F_i$), причем степень этого превышения определяется отношением числа ветвей сети к их стоимости.

Предположение о том, что m не зависит от номера узла или ветви, является справедливым, так как согласно (32) увеличение потока F_i приводит к необходимости пропорционального увеличения пропускной способности, что ведёт к более быстрому освобождению буферов, так что число требований на входе в каждый канал остаётся неизменным и необходимое число буферов оказывается постоянным.

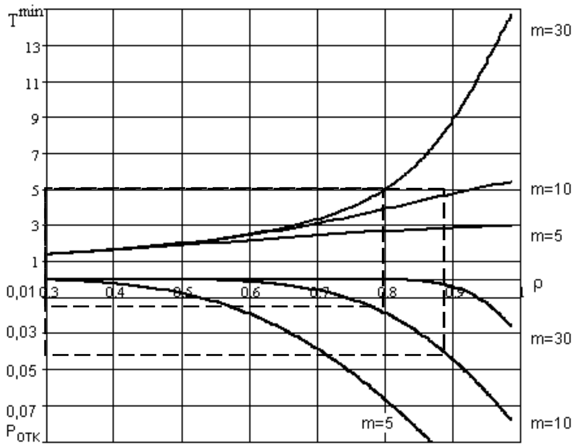

Рис. 1. Кривые зависимости средней задержки от оптимальных значений удельного потока

Кривые зависимости минимальных значений средней задержки $\overline{T}^{min}$, рассчитанные по (31) при различных значениях числа буферов в узлах коммутации от оптимальных значений удельного потока представлены в верхней части рис. 1. Предположение о бесконечном числе буферов позволяет получить верхнюю границу для задержки, которую можно достичь при конечном числе элементов буферной памяти, при этом задержка в системе без буферов дает нижнюю границу для задержки большого разнообразия систем множественного доступа с буферизацией и управляемым потоком. Однако ограничение числа буферов в узлах коммутации неизбежно приводит к тому, что часть пакетов отвергается узлом. Для учета этого обстоятельства на рис. 1, приведены совмещенные кривые зависимости вероятности отказов $\overline{T}^{min}(\rho)$ и функции $P(\rho)$ при одинаковых значениях числа буферов m, построенные в соответствии с выражением

$$P_{отк}(\rho, m) = \frac{\rho^{m+1}}{\sum_{i=0}^{m+1} \rho^i} .$$

Выводы

Анализ кривых позволяет сделать вывод о том, что минимальная средняя задержка (T^{min} = 5 с), соответствующая $\rho'_{опт} = 0{,}8$ и $m = 30$, может быть достигнута при более высокой плотности потока передаваемой по сети информации ($\rho''_{опт} = 0{,}9$ и $m = 10$) при уменьшении числа буферов в узлах коммутации в три раза. При этом вероятность отказа увеличивается с $P = 0{,}015$ до $P = 0{,}045$, что является вполне приемлемым [1].

Таким образом, полученные аналитические выражения (30, 31) позволяют при заданной стоимости передачи единицы информации осуществить выбор числа элементов буферной памяти и оптимального значения плотности потока информации, обеспечивающего минимальную среднюю задержку передачи сообщений в сети связи.

Список литературы

1. *Бакланов И.Г. Технологии измерения первичной сети. Ч.2. Системы синхронизации, B-ISDN, ATM. – М.: Эко-Трендз, 2000. – 150 с.*
2. *Блэк Ю. Сети ЭВМ: протоколы, стандарты, интерфейсы. – М.: Мир, 1990. – 506 с.*
3. *Кучук Г.А., Стасєва Я.Ю., Болюбаш О.О. Розрахунок навантаження мультисервісної мережі // Системи озброєння і військова техніка. – 2006. – № 4(8). – С. 130-134.*
4. *Кучук Г.А. Метод дослідження фрактального мережного трафіка // Системи обробки інформації. – Х.: ХУ ПС, 2005. – Вип. 5 (45). – С. 74-84.*
5. *Бертсекас Д., Галлагер Р. Сети передачи данных. – М.: Мир, 1989. – 544 с.*
6. *Новиков О.А., Петухов С.И. Прикладные вопросы теории массового обслуживания. – М.: Наука, 1989. – 400 с.*
7. *Назаров А.Н., Симонов М.В. АТМ: Технология высокоскоростных сетей. – М.: Эко-Трензд, 1999. – 252 с.*
8. *Олифер В.Г., Олифер Н.А. Новые технологии и оборудование IP- сетей. – С.-Пб.: БХВ-Санкт-Петербург, 2000. – 512 с.*
9. *Кучук Г.А., Кіріллов І.Г.. Пашнєв А.А. Моделювання трафіка мультисервісної розподіленої телекомунікаційної мережі // Системи обробки інформації. – Х.: ХУ ПС, 2006. – Вип. 9 (58). – С. 50-59.*

Поступила в редколлегию 9.03.2007

Рецензент: д-р техн. наук, проф. С.В. Козелков, Центральный НИИ навигации и управления, Киев.