

СЕТИ СВЯЗИ

УДК 621.391

КЛАСТЕРИЗАЦИЯ ИНФОРМАЦИОННЫХ ПОТОКОВ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ЗАНЯТИЯ КАНАЛЬНОГО РЕСУРСА В МУЛЬТИСЕРВИСНЫХ СЕТЯХ

С. Н. Степанов, директор департамента ОАО «Интеллект Телеком», д. т. н.; stepanov@i-tc.ru
Ю. Ю. Савенков, аспирант МТУСИ

Ключевые слова: мультисервисная сеть, каналный ресурс, резервирование, кластеризация.

Введение. При планировании пропускной способности звеньев и узлов мультисервисных сетей связи традиционно используются методики и общие положения, вытекающие из опыта проектирования и оптимизации сетей коммутации каналов. Один из постулатов гласит: при объединении передаточных возможностей линий связи возрастает коэффициент использования единицы каналного ресурса. Для ряда классических моделей теории телетрафика при одинаковых требованиях к ресурсу передачи для поступающих заявок данное утверждение может быть доказано теоретически.

Возьмем для примера хорошо известную модель Эрланга, представляющую собой полнодоступную систему из v каналов с пуассоновским входным потоком заявок интенсивности a и экспоненциально распределенным временем обслуживания со средним значением, равным 1. Для нее доля потерянных заявок π и среднее число занятых единиц каналного ресурса m оцениваются из выражений

$$\pi = E(v, a) = \frac{\frac{a^v}{v!}}{1 + a + \frac{a^2}{2!} + \dots + \frac{a^v}{v!}}, \quad m = a(1 - E(v, a)).$$

Пусть $\rho(v, a) = \frac{a}{v}(1 - E(v, a))$ — коэффициент загрузки

единицы каналного ресурса линии. Рассмотрим две аналогичные системы с параметрами каналного ресурса и поступающих заявок — v_1, v_2 и a_1, a_2 соответственно. Объединим ресурс передачи и предоставим его суммарному потоку заявок. Можно показать, что в интегрированной системе коэффициент использования единицы ресурса возрастает, т. е. выполняется соотношение

$$\rho(v_1 + v_2, a_1 + a_2) \geq \frac{v_1}{v_1 + v_2} \rho(v_1, a_1) + \frac{v_2}{v_1 + v_2} \rho(v_2, a_2).$$

Мультисервисные сети, где требования к ресурсу передачи зависят от типа сервиса, этим свойством не обладают. Более того, обслуживание подобных заявок общей транспортной средой может вызвать неконтролируемое перераспределение ресурса передачи в пользу отдельных групп заявок. Возникающие проблемы осложняют реализацию процедур, направленных на повышение эффективности работы сети и предоставление пользователям гарантированного качества обслуживания. В статье рассматриваются пути решения этих проблем, основанные на кластеризации потоков заявок по уровню требований к каналному ресурсу и на применении схем резервирования. Исследование ведется на при-

мере модели мультисервисной сети [1—7] с возможностью резервирования каналного ресурса, которую зададим функцией внутренней блокировки [7, 8].

Описание модели. Пусть n — число потоков заявок на выделение каналного ресурса, J — общее число линий, а v_j — скорость передачи j -й линии ($j = 1, 2, \dots, J$), выраженная в единицах ресурса, требуемого для обслуживания поступающих заявок на передачу трафика. Обычным путем вводятся параметры k -го потока заявок: λ_k — интенсивность поступления, $1/\mu_k$ — среднее время обслуживания и b_k — объем каналного ресурса, необходимый для обслуживания одной заявки. Будем считать, что поступление заявок для каждого анализируемого потока подчиняется закону Пуассона, а время обслуживания одной заявки имеет экспоненциальное распределение. Возьмем $a_k = \frac{\lambda_k}{\mu_k}$. Для формализованного описания топологии сети и маршрутов следования информационных потоков введем маршрутную матрицу R , в записи которой для удобства укажем номера потоков и линий:

$$R = \|r_{j,k}\| = \begin{vmatrix} & 1 & 2 & \dots & n \\ 1 & r_{1,1} & r_{1,2} & \dots & r_{1,n} \\ 2 & r_{2,1} & r_{2,2} & \dots & r_{2,n} \\ \dots & \dots & \dots & \dots & \dots \\ J & r_{J,1} & r_{J,2} & \dots & r_{J,n} \end{vmatrix},$$

$$r_{j,k} = \begin{cases} b_k, & j\text{-я линия используется при обслуживании} \\ & \text{заявки } k\text{-го потока,} \\ 0, & \text{в остальных случаях.} \end{cases}$$

Маршрутная матрица R определяет множества N_j — номеров потоков заявок, использующих j -ю линию ($j = 1, 2, \dots, J$) и R_k — номеров линий, составляющих маршрут движения трафика, относящегося к обслуживанию k -й заявки ($k = 1, 2, \dots, n$).

Обозначим через i_k число находящихся на обслуживании заявок k -го потока и введем n -мерное пространство состояний Ω , куда включены векторы (i_1, i_2, \dots, i_n) с компонентами, удовлетворяющими ограничениям $\sum_{k \in N_j} b_k i_k \leq v_j$, $j = 1, 2, \dots, J$. Данное множество задает все теоретически возможные состояния исследуемой модели сети. Реально используемое пространство состояний S является подмножеством Ω и определяется выбором функции внутренней блокировки $\varphi_k(i_1, i_2, \dots, i_n)$, которая для каждого состояния $(i_1, i_2, \dots, i_n) \in \Omega$ определяет вероятность доступа заявок k -го потока к используемому каналному ресурсу. Заявка k -го потока принимается к обслу-

живанию с вероятностью $1 - \varphi_k(i_1, i_2, \dots, i_n)$, а с дополнительной вероятностью $\varphi_k(i_1, i_2, \dots, i_n)$ получает отказ и не возобновляется ни в какой форме.

Выбор функции внутренней блокировки позволяет использовать разные модели резервирования канального ресурса. Они необходимы оператору для осуществления функций контроля за его распределением. Приведем пример реализации традиционной схемы. Пусть θ_k — предполагаемый уровень резервирования канального ресурса для заявок k -го потока. Тогда реализация традиционной схемы обеспечивается следующим выбором значений $\varphi_k(i_1, i_2, \dots, i_n)$:

$$\varphi_k(i_1, i_2, \dots, i_n) = \begin{cases} 0, & \text{если } \forall j \in R_k, \sum_{k \in N_j} b_k i_k \leq \theta_k, \\ 1 & \text{в остальных случаях.} \end{cases}$$

Если в сети имеется только одно звено, функция внутренней блокировки приобретает вид

$$\varphi_k(i_1, i_2, \dots, i_n) = \varphi_k(i) = \begin{cases} 0, & \text{если } i = i_1 b_1 + \dots + i_n b_n \leq \theta_k, \\ 1 & \text{в остальных случаях.} \end{cases} \quad (1)$$

Варьируя значением θ_k , оператор получает возможность резервировать ресурс сети для обслуживания тех заявок, которым он необходим по условиям принятого соглашения об обслуживании. Понятно, что соответствующая цель достигается за счет уменьшения коэффициента использования канального ресурса.

Динамика изменения состояний построенной модели сети описывается марковским процессом:

$$r(t) = (i_1(t), i_2(t), \dots, i_n(t)),$$

где $i_k(t)$ — число заявок k -го потока, находящихся на обслуживании в момент времени t . Обозначим через $p(i_1, i_2, \dots, i_n)$ значения вероятностей стационарных состояний модели, определяемых решением системы уравнений равновесия.

При планировании сети используются показатели приема новых заявок к обслуживанию. Важнейшим из них является π_k — доля потерянных заявок k -го потока. Ее значение определяется из выражения

$$\pi_k = \sum_{(i_1, i_2, \dots, i_n) \in S} p(i_1, i_2, \dots, i_n) \varphi_k(i_1, i_2, \dots, i_n).$$

Для определения среднего числа заявок k -го потока y_k , одновременно находящихся на обслуживании, получаем

$$y_k = \sum_{(i_1, i_2, \dots, i_n) \in S} p(i_1, i_2, \dots, i_n) i_k.$$

Расчет значений введенных показателей обслуживания заявок может быть выполнен с помощью численных методов решения системы уравнений равновесия [3, 7]. Реализация соответствующих алгоритмов ограничена значениями структурных параметров модели, и их можно применять в основном для однозвенных сетей. Для оценки характеристик систем, состоящих из нескольких звеньев, можно использовать имитационное моделирование и приближенные методы [1–7]. Алгоритм приближенного расчета значений π_k основан на реализации метода просеянного трафика [1–3, 7].

Оценка показателей обслуживания заявок. В соответствии с основными положениями метода просеянного трафика для оценки доли потерянных заявок k -го потока используется соотношение

$$\pi_k \approx 1 - \prod_{j \in R_k} (1 - B_{k,j}), \quad k = 1, 2, \dots, n,$$

где $B_{k,j}$ — оценка доли заявок k -го потока, потерянных на j -м звене сети.

Значение $B_{k,j}$ находится с помощью модели отдельного звена с обобщенной схемой резервирования. Она является частным случаем исследуемой модели сети и может быть получена при $J = 1$. Для оценки показателей обслуживания заявок в [7] построен приближенный алгоритм. Приведем соответствующие расчетные формулы, опустив индекс j , показывающий зависимость вероятностных характеристик от номера звена.

Пусть $P(i)$ — ненормированное значение вероятности занятости i единиц канального ресурса звена. Выразим $P(i)$ через ненормированное значение вероятности $P(0)$, используя для этого следующий алгоритм:

1. Положим $P(0) = 1$.
2. Выразим значения вероятностей $P(i)$, $i = 1, 2, \dots, v$ через $P(0)$ при помощи соотношения

$$P(i) = \frac{1}{i} \sum_{k=1}^n a_k b_k P(i - b_k) (1 - \varphi_k(i - b_k)) I(i - b_k \geq 0) \quad (2)$$

и последовательно увеличивая i от 1 до v . Нетрудно проверить, что при каждом фиксированном i отдельные слагаемые, входящие в правую часть соотношения (2), либо уже представлены через $P(0)$, либо равны 0; $I(\bullet)$ — индикаторная функция.

3. Находим значение нормировочной константы:

$$N = \sum_{i=0}^v P(i).$$

4. Определяем нормированные значения вероятностей $p(i)$:

$$p(i) = \frac{P(i)}{N}, \quad i = 0, 1, \dots, v.$$

5. Находим показатели качества совместного обслуживания заявок для каждого из n анализируемых потоков. Получаем

$$B_k = \sum_{i=0}^v p(i) \varphi_k(i), \quad y_k = a_k (1 - B_k), \quad k = 1, 2, \dots, n. \quad (3)$$

Приведем основные расчетные формулы, относящиеся к определению показателей обслуживания заявок на j -м звене. Исходные параметры модели: v_j — число единиц канального ресурса; $\sigma_j(j)$ — интенсивность поступления заявок (в эрлангах); b_l — число единиц канального ресурса, необходимых для обслуживания одной заявки; $l \in N_j$ — номер потока заявок, использующих j -е звено. Количество входных потоков заявок определяется числом элементов в множестве N_j . В соответствии с (3)

$$B_{k,j} = \sum_{i=0}^{v_j} p(i) \varphi_k(i), \quad k \in N_j, \quad (4)$$

где ненормированные значения вероятностей $p(i)$, $i = 0, 1, \dots, v$ определяются из приведенного выше алгоритма, а функция внутренней блокировки $\varphi_k(i)$ находится из уравнения (1).

Характеристики $B_{k,j}$ являются функциями параметров v_j и $\sigma_j(j)$, $l \in N_j$. Обозначим эту зависимость в виде $F_{k,j}(v_j, \sigma_j(j))$, $l \in N_j$. Для применения (4) необходимо найти $\sigma_j(j)$. Из предположения о независимости отказов на отдельных звеньях сети и формулы Литтла получаем

$$\sigma_j(j) = \frac{y_l}{1 - B_{l,j}} = a_l \prod_{u \in R_l \setminus \{j\}} (1 - B_{l,u}).$$

Подставив найденное для $\sigma_j(j)$ выражение в $F_{k,j}(v_j, \sigma_j(j))$, $l \in N_j$, получаем систему неявных уравнений для определения $B_{k,j}$:

$$B_{k,j} = F_{k,j} \left(v_j, a_l \prod_{u \in R_l \setminus \{j\}} (1 - B_{l,u}), l \in N_j \right), j = 1, 2, \dots, J, k \in N_j.$$

Для оценки $B_{k,j}$ можно использовать метод последовательных подстановок в виде

$$B_{k,j}^{(s)} = F_{k,j} \left(v_j, a_l \prod_{u \in R_l \setminus \{j\}} (1 - B_{l,u}^{(s-1)}), l \in N_j \right), j = 1, 2, \dots, J, k \in N_j \quad (5)$$

с начальным приближением $B_{k,j}^{(0)} = 0, j = 1, 2, \dots, J, k \in N_j.$

Запись $B_{l,u}^{(s,s-1)}$ в (5) означает, что на s -м шаге рекурсии при вычислении правой части соотношения используются найденные на s -м шаге значения $B_{l,u}^{(s)}$. Если таковых нет, то нужно взять значения $B_{l,u}^{(s-1)}$, найденные на предыдущем, $(s-1)$ -м, шаге. Итерационный процесс завершается, когда нормированная разница между последовательными приближениями к решению становится меньше некоторой заранее выбранной малой величины. Теоретически сходимость итерационной процедуры (5) не доказана, но для большинства интересных случаев сходимость имеет место. Более подробно реализация рекуррентных соотношений вида (5) рассмотрена в [7].

Численные результаты. Проведем численный анализ проблем, возникающих при совместном использовании транспортной среды при обслуживании потоков мультисервисного трафика. Выполним соответствующее исследование на примере сети кольцевой структуры, состоящей из пяти узлов и пяти линий (рис. 1).

В модели анализируется процесс обслуживания 14 потоков заявок, различающихся маршрутами следования трафика и объемом канального ресурса, необходимого для обслуживания одной заявки. В сети выделено семь маршрутов, по каждому из которых следует два потока трафика. Один поток относится к сервисам, не требующим значительного ресурса (передача речевых сообщений, некоторых видов данных и т. д.), и для обслуживания его заявки используется одна канальная единица. Второй поток (передача видеоданных, обмен файлами и т. д.) нуждается в существенно большем ресурсе. Предполагается, что при обслуживании заявки такого потока необходимо уже 60 единиц канального ресурса. В дальнейшем трафик первого типа будем называть «легким», а второго, соответственно, «тяжелым».

В сети рассматриваются маршруты: (1, 2) с потоками 1, 2; (1, 2, 3) с потоками 3, 4; (2, 3, 4) с потоками 5, 6; (3, 4) с по-

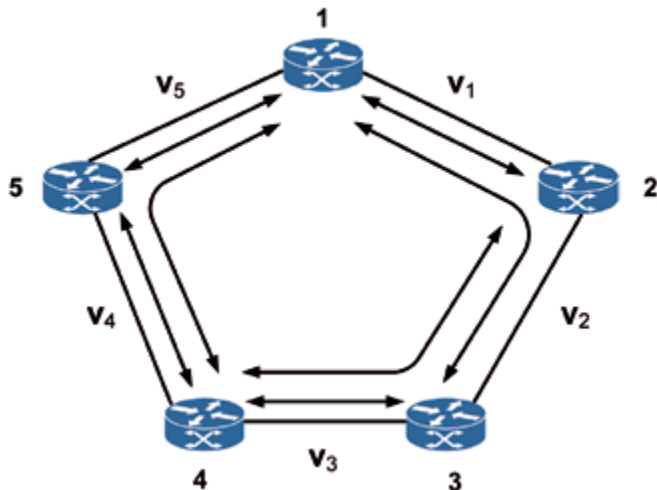


Рис. 1

токами 7, 8; (4, 5) с потоками 9, 10; (4, 5, 1) с потоками 11, 12; (5, 1) с потоками 13, 14. При этом $b_{2i-1} = 1$ и $b_{2i} = 60, i = 1, 2, \dots, 7.$ Маршрутная матрица исследуемой модели сети состоит из пяти строк и 14 столбцов:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	b_1	b_2	b_3	b_4	0	0	0	0	0	0	0	0	0	0
2	0	0	b_3	b_4	b_5	b_6	0	0	0	0	0	0	0	0
3	0	0	0	0	b_5	b_6	b_7	b_8	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	b_9	b_{10}	b_{11}	b_{12}	0	0
5	0	0	0	0	0	0	0	0	0	0	b_{11}	b_{12}	b_{13}	b_{14}

Покажем, как при таком выборе входных параметров с увеличением интенсивности поступления заявок с малым использованием ресурса передачи происходит вытеснение с обслуживания ресурсоемких заявок. Сформулируем дальнейшие предположения о значениях входных параметров. Объем ресурса линий, выраженный в единицах канального ресурса, определялся из равенств $v_1 = 180, v_2 = 240, v_3 = 180, v_4 = 180, v_5 = 300.$ Среднее время обслуживания одной заявки взято равным 1. Базовые значения интенсивностей поступления заявок выражены в эрлангах и определяются из соотношений: $a_1 = 3; a_2 = 0,5; a_3 = 60; a_4 = 1; a_5 = 60; a_6 = 1; a_7 = 30; a_8 = 0,5; a_9 = 20; a_{10} = 1/3; a_{11} = 70; a_{12} = 7/6; a_{13} = 80; a_{14} = 4/3.$ Базовые значения трафика выбраны так, чтобы суммарный объем трафика, предложенный на j -ю линию, будучи выраженным в эрлангах, равнялся бы скорости передачи линии, выраженной в единицах канального ресурса, и при этом объемы «легкого» и «тяжелого» трафика совпадали.

Далее рассматривалось изменение показателей обслуживания заявок при умножении базовых значений интенсивностей поступления заявок для первого и второго потоков на величину загрузки ресурса передачи ρ , которая менялась от 0,01 до 2,5. Перегрузка первого звена сети ведет к изменению показателей обслуживания заявок, использующих для передачи соответствующих информационных сообщений это звено. Характер изменений проиллюстрирован на рис. 2, где показана зависимость $\pi_1, \pi_2, \pi_3, \pi_4,$ от увеличения $\rho.$

Локальное уменьшение вероятности потерь для заявок первого потока объясняется тем, что с ростом ρ заявки первого потока, занимая канальный ресурс, перекрывают доступ к нему заявок второго потока, поскольку последним для передачи соответствующего трафика одновременно требуется довольно большой ресурс линии. В результате заявки первого потока получают дополнительные возможности для использования канального ресурса, которые могли быть и не предусмотрены в принятом соглашении об обслужива-

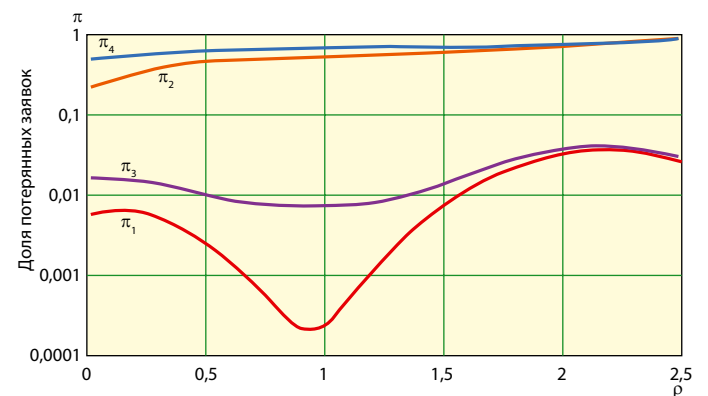


Рис. 2

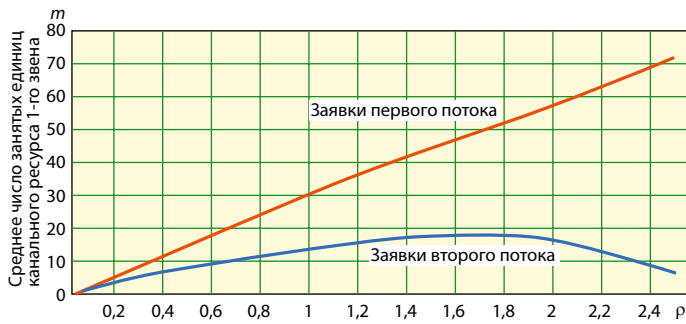


Рис. 3

нии. Аналогичные свойства имеет распределение ресурса для заявок третьего и четвертого потоков.

В критической ситуации, подобной той, что показана на рис. 2, заявки с малой потребностью в канальном ресурсе могут вытеснить из обслуживания большую часть ресурсоемких заявок. Подтверждением тому служит среднее число единиц канального ресурса (m), используемых при обслуживании заявок первого и второго потоков (рис. 3), для модели сети с теми же значениями входных параметров, что были задействованы при расчете данных, представленных на рис. 2.

Рассмотренные примеры неконтролируемого перераспределения канального ресурса нуждаются в более детальном исследовании. Интуитивно понятно, что с увеличением скорости передачи линии отрицательных последствий становится меньше. Это положение иллюстрируют численные примеры. Для устранения отмеченных недостатков можно также использовать механизмы контроля за распределением канального ресурса. В первую очередь это резервирование, ограничение доступа, назначение приоритетов. Особенности реализации соответствующих подходов активно исследуются в литературе по теории телетрафика [1—8].

Рассмотрим еще один подход, который основан на кластеризации информационных потоков по величине используемого канального ресурса и последующем отдельном обслуживании выделенных кластеров. Покажем, как предлагаемая схема в ряде случаев обеспечивает более эффективное, чем традиционные подходы, использование канального ресурса. Для сравнения вернемся к модели сети, представленной на рис. 1. Объем ресурса линий, выраженный в единицах канального ресурса, определялся из равенств $v_1 = 360, v_2 = 360, v_3 = 360, v_4 = 360, v_5 = 360$. Среднее время обслуживания одной заявки взято равным 1. Базовые значения интенсивностей поступления заявок выражены в эрлангах и определяются из соотношений $a_1 = a_2 = a_5 = a_6 = a_9 = a_{10} = a_{13} = a_{14} = 120$ и $a_3 = a_4 = a_7 = a_8 = a_{11} = a_{12} = 60$. Базовые значения трафика выбраны по тому же принципу, что и в предыдущем случае.

Сделаем дополнительные предположения о значениях входных параметров. Переменным параметром будет величина канального ресурса b , требуемая для обслуживания одной заявки ресурсоемкого сервиса. Выполняется соотношение $b_{2k} = b; k = 1, 2, \dots, 7$, где b принимает значения от 1 до 200. Будем предполагать, что при изменении потребностей в ресурсе передачи интенсивность предложенного трафика соответствующего потока не менялась, т. е. $a_{2k} b_{2k}$ остается равным константе, определяемой базовыми значениями.

Обозначим через ρ_j коэффициент использования канальной единицы j -го звена сети. На рис. 4 показано, как с ростом b меняется ρ_j . Рассмотрены три схемы распределения канального ресурса: совместное использование ресурса передачи; совместное использование ресурса передачи с резервированием (уровень резервирования для всех потоков заявок выбирался

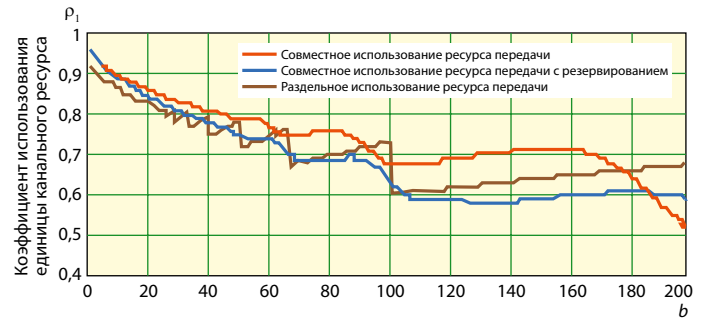


Рис. 4

из равенства $\theta_k = v_k - b; k = 1, 2, \dots, 14$); раздельное использование ресурса (канальный ресурс исходной сети делится в пропорции 4 : 5 и реализуется для обслуживания «легкого» и «тяжелого» трафика соответственно).

Результаты расчетов (см. рис. 4) показывают, что для рассмотренной модели при определенных соотношениях значений входных параметров более предпочтительным оказывается раздельное использование транспортных возможностей сети, если речь идет об обслуживании «легкого» и «тяжелого» трафика. Среди других преимуществ данного способа — простота реализации. Таким образом, выбор схемы распределения канального ресурса в каждом конкретном случае определяется характером и условиями передачи и требует изучения. Соответствующее исследование следует проводить с помощью методов и моделей теории телетрафика [1—9].

Заключение. Исследование схем распределения канального ресурса мультисервисного трафика показало, что при значительной дифференциации потребностей заявок в скорости передачи может происходить неконтролируемое перераспределение канального ресурса в пользу отдельных групп заявок. Для устранения этого недостатка рекомендуется применять механизмы контроля за распределением канального ресурса. К ним относятся резервирование, ограничение доступа. В ряде случаев для повышения эффективности использования канального ресурса целесообразно прибегнуть к кластеризации поступающих заявок по величине уровня требований к канальному ресурсу и последующее раздельное обслуживание заявок.

ЛИТЕРАТУРА

1. **Ross K. W.** Multiservice loss models for broadband telecommunication networks. — London: Springer, 1995.
2. Broadband network traffic. Performance evaluation and design of broadband multiservice networks. Final report of action COST 242/James Roberts, eds. Lecture notes in computer sciences. Springer, 1996.
3. **Лагутин В. С., Степанов С. Н.** Телетрафик мультисервисных сетей связи. — М.: Радио и связь, 2000.
4. **Iversen V. B.** Teletraffic Engineering Handbook. — ITU-D, Nov. 2002.
5. **Башарин Г. П.** Лекции по математической теории телетрафика. — М.: Изд-во РУДН, 2009.
6. **Наумов В. А., Самуйлов К. Е., Яркина Н. В.** Теория телетрафика мультисервисных сетей. — М.: Изд-во РУДН, 2007.
7. **Степанов С. Н.** Основы телетрафика мультисервисных сетей. — М.: Эко-Трендз, 2010.
8. **Степанов С. Н., Костров А. В.** Моделирование мультисервисной сети с обобщенной схемой резервирования канального ресурса//Т-Comm. — 2008. — № 4.
9. **Васькин Ю. Л.** Анализ влияния трафика на пропускную способность сети доступа к информационным ресурсам//Электросвязь. — 2010. — № 4.