УДК 004.042

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И УПРАВЛЕНИЕ ИНТЕНСИВНЫМИ ИНФОРМАЦИОННЫМИ ПОТОКАМИ В ТУННЕЛИРУЕМЫХ ВИРТУАЛЬНЫХ СРЕДАХ

А.С. Сигов, И.О. Дементьев

ГУВПО «Московский государственный институт радиотехники, электроники и автоматики (технический университет)»
E-mail: dementiev@mirea.ru

Предлагается обновленный подход к модельному представлению виртуальных сред в части управления интенсивными информационными потоками.

Ключевые слова:
Виртуальная среда, виртуальное образовательное пространство, туннелируемая сеть, виртуальная сеть.

Задача управления туннелируемыми виртуальными средами [1], в том числе управления информационными потоками в виртуальных сетях, в настоящее время вызывает у специалистов в области информационных технологий и систем особый интерес [2, 3]. Активно развиваются подходы, методы и целые концепции построения виртуализированных образовательных пространств. Например, одним из таких подходов, стала концепция, разработанная специалистами Центра новых информационных технологий МИРЭА, заключающаяся в использовании туннелирования персональных образовательных пространств конечных пользователей из единого репозитария данных (знаний). Подробно эта концепция рассматривается в [3, 4]. Согласно предложенной концепции, равно как и множеству публикаций по теме [5, 6 и др.], основной задачей, требующей решения при реализации подобных средств портального обустройства и виртуализации образовательного пространства учебного заведения в целом, становится задача управления интенсивными информационными потоками внутри виртуальной образовательной среды.

Анализ существующих подходов и методов моделирования виртуальных сред показал недостаточную разработку методов математического моделирования туннелированных пространств (сетей) применительно к сформулированной выше задаче. Наиболее часто применяются методы аналитического, симуляционного моделирования и моделирования с использованием трендов [7], так же, как и математического аппарата Теории массового обслуживания.

Для решения задачи управления туннелируемыми виртуальными средами авторами предлагается решение, основанное на обновлении модельного представления туннелируемой виртуальной сети [8] и введении дополнительных ограничений на модель одновременно с использованием в ней регулирующего критерия взаимного искажения потоков данных. Согласно авторскому видению критерий этот основан на представлении передаточной энтропийной функции на основе определения взаимной энтропии.

Используя общие для сетевого моделирования нотации, согласно существующим подходам, виртуальная сеть, состоящая из N узлов, соединенных L связями, каждая пропускной способностью C_l может быть представлена следующим образом: s_k – генератор информационного потока; t_k – точка назначения информационного потока; $k \in K$ – запрос, проходящий по сети между s_k и t_k, а d_k – это количество таких запросов в сети. $m \in P_k$ – маршрут от s_k до t_k, по которому передается информация по запросу k; x_{km} – коэффициент использования пропускной способности канала согласно запросу k по маршруту m. Общий объём передаваемой по маршруту информации можно определить как $x_{km}\, d_k$.

Применяя указанные нотации (в расширительном толковании авторов статьи), появляется возможность ввести в модель туннелируемой сети представленные ниже ограничения.

Ограничение превышения суммарного объема данных, передаваемых по всем туннелям над пропускной способностью линии связи в модели может быть представлено следующим математическим описанием:

$$\sum_{k \in K} x_{km} \le 1, \quad k \in K. \qquad (1)$$

Для маршрутизации информационных потоков в виртуальной сети, передача данных в которой, преимущественно происходит по туннелям, видится рациональным ввести в модель переменную, позволяющую контролировать движение данных по различным маршрутам, с опорой на запрос – δ^{l}_{km}. Переменная принимает два значения:

1 – если путь m по запросу $к$ использует линию связи l;

0 – в обратном случае.

При помощи введенной переменной также определяется суммарный объем потока данных на линии связи l:

$$F_l = \sum_{k \in K} \sum_{m \in P_k} \delta^{l}_{km} x_{km} d_k\,. \qquad (2)$$

Для того, чтобы убедиться, что подобный объем сможет передаваться по заданному маршруту, достаточно удостовериться в том, что он не превышает пропускной способности линии связи. Авторами предлагается на модели использовать следующую границу, определяемую с опорой на (2):

$$\sum_{k\in K}\sum_{m\in P_k}\delta^l_{km}x_{km}d_k \le C_l,\; l\in L. \qquad (3)$$

Значение x_{km} определяет статус туннеля m для запроса k.

При x_{km}=0 туннель не существует, в обратном случае, он задействован в маршруте, отсюда, определяемое авторами модельное ограничение как указатель активности пути (туннеля), должен быть бинарным со значением 1 или 0. Соответствующую переменную предлагается определить следующим образом:

$$\begin{gathered} w_{km} \le x_{km}d_k, \\ x_{km} \le w_{km}, \\ m\in P_k, k\in K. \end{gathered} \qquad (4)$$

Здесь же можно учесть и ситуацию, когда пропускная способность связи настолько мала, что её нужно расценивать как несуществующую. Для этого авторами в модельное представление (4) вводится уточняющий коэффициент ε:

$$\begin{gathered} \varepsilon w_{km} \le x_{km}d_k, \\ x_{km} \le w_{km}, \\ m\in P_k, k\in K. \end{gathered} \qquad (5)$$

Значение ε определяется в каждом отдельном случае согласно топологии конкретной сети, требованиям к объемам передаваемых потоков и их интенсивности, значение ε может лежать в пределах [0,1].

Вышеописанных дополнительных ограничений на модель достаточно, чтобы определить количество активных туннелей на линии связи виртуальной сети:

$$t_l = \sum_{k\in K}\sum_{m\in P_k}\delta^l_{km}w_{km},\; l\in L.$$

Опираясь на приведенную формулу, можно сформулировать подход к управлению интенсивными информационными потоками в виртуальной сети на основе одновременного регулирования количества туннелей на линии связи и учета взаимного искажения потоков при передаче внутри различных туннелей, уменьшая тем самым искажения при передаче данных в виртуальных сетях.

Чтобы учесть в модели ранжирование количества туннелей вводится дополнительное ограничение в модель, гарантирующее, что количество туннелей всегда меньше, чем T_l:

$$\sum_{k\in K}\sum_{m\in P_k}\delta^l_{km}w_{km} \le T_l,\; l\in L.$$

Едино с формулировками (1, 3, 5) автором предлагается подход, позволяющий на модельном уровне управлять передаваемыми данными в сетях произвольных топологий, основывающийся на ограничении количества туннелей.

В целевой функции управления информационными потоками в сети должны отражаться степень взаимного негативного влияния мультиплексированных потоков, с целью выбора более приемлемых маршрутов передачи потоков данных от источника к получателю.

Одновременно решить указанные задачи позволила предлагаемая автором формулировка функции управления трафиком в виртуальной сети:

$$f(x) - \sum_{l\in L} f_p(t_l),$$

где $f(x)$ – целевая функция; $\sum_{l\in L} f_p(t_l)$, вычитаемая из целевой функции, – функция, отражающая взаимное негативное влияние мультиплексированных потоков, основывающаяся на передаточной функции сетевого узла.

На основании этой формулировки, можно определить единую функцию управления трафиком в туннелируемой виртуальной сети. Нотации модели выглядят следующим образом:

N: количество узлов в сети;

K: количество запросов в сети;

Q_k: количество потоков на линии связи в сети;

L: количество линий связи в сети;

ζ_k: коэффициент предпочтительности пути для запроса $k\in K$;

λ^p_k: интенсивность потока данных $p\in Q_k$ для запроса $k\in K$ (Мбит/с);

T_l: максимальное количество туннелей на линии связи $l\in L$;

C_l: пропускная способность линии связи $l\in L$ (Мбит/с);

Λ^{pq}_{km}: регулирующий коэффициент уровня искажения мультиплексированных данных.

Дополнительно к нотациям определяется критерий выбора пути для запроса:

$$\sum_{m\in P_k} x^p_{km} \le 1{,}0;\; x^p_{km}\in\{0,1\},$$

где $p\in Q_k$, $m\in P_k$, $k\in K$.

Таким образом, общую пропускную способность сети (F_l) для передачи заданного объема потоковых данных можно определить следующим образом:

$$F_l = \sum_{k\in K}\sum_{m\in P_k}\sum_{p\in Q_k}\delta^l_{km}x^p_{km}\lambda^p_k.$$

С использованием уже полученного модельного представления, ограничение на общую интенсивность вещания в сети находится как

$$\sum_{k\in K}\sum_{m\in P_k}\sum_{p\in Q_k}\delta^l_{km}x^p_{km}\lambda^p_k \le C_l,\; l\in L.$$

Ограничение на количество туннелей на линии связи для функции управления трафиком имеет вид:

$$\sum_{k\in K}\sum_{m\in P_k}\delta^l_{km}w_{km} \le T_l,$$

где $x^p_{km}\in w_{km}$, $w^p_{km}\in\{0,1\}$, $p\in Q_k$, $m\in P_k$, $k\in K$.

Исходя из постановки задачи управления информационными потоками в туннелируемой виртуальной сети, можно определить $f(x)$ функции управления трафиком как $\sum_{k\in K}\sum_{m\in P_k}\sum_{p\in Q_k}\zeta_k x_{km}^p \lambda_k^p$.

Эта часть функции отражает выход на максимальные значения функции с целью обеспечения предельной пропускной способности сети в целом.

Вторая часть функции управления трафиком, отражающая негативное влияние взаимного искажения потоков данных, которая, соответственно, вычитается в функции управления трафиком из первой части и должна быть минимизирована, авторами определена следующим образом:

$$\sum_{k\in K}\sum_{m\in P_k}\sum_{p\in Q_k}\sum_{q\in\{Q_k\setminus p\}} x_{km}^p \Lambda_{km}^{pq} x_{km}^q .$$

В итоге, формулировка принципа управления интенсивными информационными потоками, основывающегося на оптимизации функции управления в туннелируемой виртуальной сети, принимает следующий вид:

$$\max_{\{x,w\}} f = \sum_{k\in K}\sum_{m\in P_k}\sum_{p\in Q_k}[\zeta_k x_{km}^p \lambda_k^p - \varphi_k \sum_{q\in\{Q_k\setminus p\}} x_{km}^p \Lambda_{km}^{pq} x_{km}^q],$$

где

$$x_{km}^p \le w_{km}, \quad p\in Q_k, \quad m\in P_k, \quad k\in K;$$

$$\sum_{m\in P_k} x_{km}^p \le 1{,}0, \quad p\in Q_k, \quad k\in K;$$

$$\sum_{k\in K}\sum_{m\in P_k} x_{km}^p \le 1{,}0, \quad p\in Q_k, \quad k\in K;$$

$$\sum_{k\in K}\sum_{m\in P_k}\sum_{p\in Q_k}\delta_{km}^l x_{km}^p \lambda_k^p \le C_l, \quad l\in L;$$

$$\sum_{k\in K}\sum_{m\in P_k}\delta_{km}^l w_{km} \le T_l, \quad l\in L;$$

$$x_{km}^p \in \{0,1\}, \quad p\in Q_k, \quad m\in P_k, \quad k\in K;$$

$$w_{km} \in \{0,1\}, \quad m\in P_k, \quad k\in K.$$

Для определения взаимного негативного влияния мультиплексированных данных (Λ_{km}^{pq}), передаваемых в виртуальной сети авторами предложен подход, основанный на использовании взаимной энтропии [9] с разработкой критерия – передаточной функции в оценке взаимных искажений туннелями друг друга.

Взаимная энтропия, или энтропия объединения, предназначена для расчёта энтропии взаимосвязанных систем (энтропии совместного появления статистически зависимых сообщений) и обозначается $H(AB)$, где A, характеризует передатчик, а B – приёмник.

Взаимосвязь переданных и полученных сигналов описывается вероятностями совместных событий $p(a_ib_j)$, и для полного описания характеристик канала требуется только одна матрица, таблица.

Для более общего случая, когда описывается не канал, а просто взаимодействующие системы, матрица необязательно должна быть квадратной. Сумма всех элементов столбца с номером j даст $p(b_j)$, сумма строки с номером i есть $p(a_i)$, а сумма всех элементов матрицы равна 1. Совместная вероятность $p(a_ib_j)$ событий a_i и b_j вычисляется как произведение исходной и условной вероятности,

$$p(a_ib_j) = p(a_i)p(b_j \mid a_i) = p(b_j)p(a_i \mid b_j).$$

***Таблица**. Канальная матрица*

$p(a_1b_1)$	$p(a_1b_2)$	…	$p(a_1b_j)$	…	$p(a_1b_m)$
$p(a_2b_1)$	$p(a_2b_2)$	…	$p(a_2b_j)$	…	$p(a_2b_m)$
…	…	…	…	…	…
$p(a_ib_1)$	$p(a_ib_2)$	…	$p(a_ib_j)$	…	$p(a_ib_m)$
…	…	…	…	…	…
$p(a_mb_1)$	$p(a_mb_2)$	…	$p(a_mb_j)$	…	$p(a_mb_m)$

Условные вероятности производятся по формуле Байеса. Таким образом, имеются все данные для вычисления энтропий источника и приёмника:

$$H(A) = -\sum_i\left(\sum_j p(a_ib_j)\log\sum_j p(a_ib_j)\right);$$

$$H(B) = -\sum_j\left(\sum_i p(a_ib_j)\log\sum_i p(a_ib_j)\right).$$

Взаимная энтропия вычисляется последовательным суммированием по строкам (или по столбцам) всех вероятностей матрицы, умноженных на их логарифм:

$$H(AB) = -\sum_i\sum_j p(a_ib_j)\log p(a_ib_j).$$

Взаимная энтропия обладает свойством информационной полноты – из неё можно получить все рассматриваемые величины.

Используя такое определение можно определить передаточную функцию (коэффициент) искажения Λ_{km}^{pq} следующим образом:

$$\Lambda_{km}^{pq} = \sum_i \frac{\sum_j p(a_ib_j)\log\sum_j p(a_ib_j)}{\sum_j p(a_ib_j)\log p(a_ib_j)}.$$

Пределы суммирования определяются в соответствии с общим модельным представлением и зависят от размерностей виртуального пространства. Подобная формулировка позволяет определить степень взаимного негативного влияния интенсивных информационных потоков и интегрировать количественную оценку в математическое описание туннелируемой виртуальной сети.

Видны границы коэффициента, вытекающие из определения энтропии, определяемой на отрезке [0,1] (см. [9]).

При отсутствии взаимного влияния (в рассматриваемом случае негативного) взаимная энтропия стремится к нулю, сводя к минимуму вычитаемую часть целевой функции управления виртуальной сетью. Заметим, однако, что на практике едва ли достижимо граничное значение, при котором можно пренебрегать взаимным влиянием информационными потоками (экспериментально определено на границе Λ_{km}^{pq}=0,3). Достаточное значение определяется исходя из топологии и масштаба виртуальной сети и носит частный характер в каждом отдельном случае, вследствие чего полученные экспериментальные данные в настоящей статье не приводятся.

Задача оптимизации сформулированной в настоящей статье функции решена за счет декомпозиции с использованием релаксации Лагранжа, полученные решения проверены на сходимость и вычислительную сложность. Согласно проведенным экспериментам, в среднем, задача оптимизации функции решалась за 400 итераций алгоритма. Указанные результаты в статье не приводятся в силу ограниченного объема статьи.

В заключение можно отметить, что разработан обновленный подход к математическому моделированию туннелируемых виртуальных сред на основе ограничения количества туннелей в виртуальной сети, позволяющий организовать управление интенсивными информационными потоками на основе расширения нотаций в математической модели туннелируемой сети и формулировки целевой функции управления трафиком.

Указанная работа проводилась в рамках госбюджетных НИОКР, в том числе в составе Аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы 2006–2008» и позволило реализовать и внедрить в учебный процесс МИРЭА многоуровневую систему портальной поддержки технического университета.

СПИСОК ЛИТЕРАТУРЫ

1. Deutsch D. The Fabric of Reality. Allen lane the penguin press / Перевод с англ. Н.А. Зубченко / Под общей ред. акад. РАН В.А. Садовничего. – М.-Ижевск: РХД, 2001. – 375 с.
2. Система информационного поиска научной информации Google Scholar. Режим доступа: http://scholar.google.com/scholar?hl=en&lr=&scoring=r&q=virtual%2Bnetwork&as_ylo=2003&btnG=Search. – 29.12.2008.
3. Сигов А.С. Многоуровневое портальное строительство в образовательной практике технического университета // Подготовка и переподготовка специалистов по информационным технологиям: Матер. Всеросс. научно-практ. конф. – М., 2007. – С. 4–11.
4. Отчет ЦНИТ МИРЭА-МГДД(Ю)Т за 2008 год. Т. 2. (В шести томах) / Центр новых информационных технологий МИРЭА-МГДД(Ю)Т. – М.: МИРЭА, 2008. – 462 с.
5. Система информационного поиска научной информации Google Scholar. [Электронный ресурс]. — режим доступа: http://scholar.google.com/scholar?q=%D0%B2%D0%B8%D1%80%D1%82%D1%83%D0%B0%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F+%D1%81%D0%B5%D1%82%D1%8C&hl=en&lr=&btnG=Search. – 29.12.2008.
6. Образовательные порталы России. Сб. статей. Вып. 1 / Научн. ред. В.В. Радаев. – М.: Технопечать, 2004. – 148 с.
7. Семенов Ю.А. Telecommunication technologies – телекоммуникационные технологии / v3.1. – 19 марта 2008 года. [http://book.itep.ru/].
8. Srivastava S., Medhi D. Traffic Engineering of Tunnel-based Networks having Class Specific Diversity Requirements, in Journal of Combinatorial Optimization, special issue on Communication Networks and Internet Applications, V. 12. – № 1–2, September 2006. – P. 97–125.
9. Габидулин Э.М., Пилипчук Н.И. Лекции по теории информации. – М.: МФТИ, 2007. – 214 с.

Поступила 29.12.2008 г.