УДК 621.391.18

Д.Ю. Гужва

ЭВОЛЮЦИОННЫЙ СИНТЕЗ *VPN*-СЕТЕЙ В ИНФОТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

Предложен метод поиска рациональной топологии VPN-сети с использованием подхода, называемого эволюционным синтезом. Рассмотрен генетический алгоритм оптимизации структуры сети. Приведена структура хромосомы, однозначно отображающая конкретную топологию сети. Предложена модель оценки пропускной способности VPN-сети, основанная на теории массового обслуживания и учитывающая топологию и обрабатываемую нагрузку сети.

Ключевые слова: *синтез сети, топология сети, генетический алгоритм оптимизации, пропускная способность, система массового обслуживания.*

Введение. Внедрение в практику построения инфотелекоммуникационных систем (ИТКС) сетей общего пользования (СОП), в частности сети Интернет, предопределяет значительный интерес к решению задачи построения на основе открытых каналов связи защищенных компьютерных сетей. Основной технологией, призванной решать данную задачу, является технология виртуальных частных сетей (*virtual private net – VPN*), позволяющая организовывать между парами абонентов защищенные «туннели», проходящие через СОП. На концах «туннеля» находятся криптомаршрутизаторы (КМ), осуществляющие шифрацию (дешифрацию) пакетов, на которые разбиваются передаваемые по *VPN*-каналу сообщения. Между КМ организуется обычный *IP*-канал, проходящий по СОП и предоставляемый, как правило, провайдером Интернет. *VPN*-каналы могут быть *простыми* либо *составными* (*транзитными*), состоящими из последовательно соединенных простых *VPN*-каналов (рисунок 1).

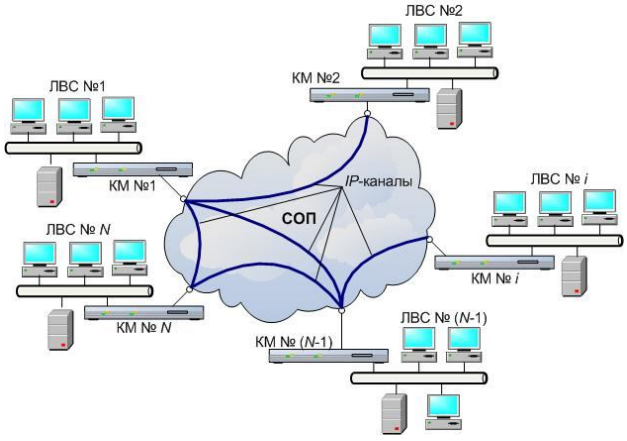

Рисунок 1 – Структура *VPN*-сети (вариант)

Постановка задачи. Базовыми типами топологии взаимосвязанных *VPN*-каналов являются: полносвязный, радиальный и кольцевой. Каждый из типов имеет свои достоинства и недостатки, отражающиеся на значениях основных показателей эффективности: пропускной способности, устойчивости, безопасности и стоимости сети. На практике, как показано на рисунке 1, обычно реализуется смешанная топология. В этом случае достигается компромиссное сочетание значений указанных показателей эффективности, отвечающих требованиям, поставленным в задаче синтеза *VPN*-сети.

Задача синтеза *VPN*-сети сводится в основном к поиску оптимальной топологии *VPN*-сети и относится к классу *NP*-полных. Общее количество возможных решений задачи определяется выражением $2^{N*(N-1)/2}$. Учитывая, что целевые функции данной задачи, как правило, являются нелинейными и имеют большое множество локальных экстремумов, применение для ее решения традиционных средств теории оптимизации весьма затруднительно. Поэтому предложим для решения задачи синтеза *VPN*-сети другой подход, известный как эволюционный синтез систем [1], являющийся одним из направлений технологии искусственного интеллекта.

Метод решения задачи. *Эволюционный синтез* – это методология поиска рациональных вариантов построения технических систем с использованием законов эволюционного развития живой природы. В частности, в нашем случае для решения задачи синтеза *VPN*-сети предлагается использовать *генетический алгоритм оптимизации* (ГАО), успешно применяющийся для решения многих задач в эволюционном моделировании [2].

Сущность метода ГАО заключается в следующем (рисунок 2).

На этапе инициализации (блок 1) случайным образом формируется начальное множество возможных альтернативных решений.

Рисунок 2 – Блок-схема генетического алгоритма оптимизации

Данное множество в методологии эволюционного синтеза обозначается термином *популяция*.

Каждое решение, или *особь* популяции, характеризуется строкой, изоморфно связанной с векторами и матрицами переменных, определяющими это решение. Эта строка называется *хромосомой,* а отдельный символ в ней – *геном*.

На каждом последующем этапе эволюционного синтеза в ГАО выполняются следующие действия. Из популяции случайным образом выбираются пары особей, именуемые *родителями*. Между ними происходит процесс скрещивания, или *кроссовера* (блок 2), в результате которого появляется пара новых особей-потомков. Хромосома каждого из потомков формируется из двух частей: одна часть берется от хромосомы «отца», а вторая – от хромосомы «матери». Потомки добавляются в общую популяцию.

Кроме того, на каждом этапе эволюционного синтеза часть особей подвергается *мутации*, в ходе которой случайным образом изменяются отдельные гены в хромосоме (блок 3).

Так как популяция имеет количественные ограничения на число особей, ее составляющих, особи, имеющие наименьшую функцию пригодности, удаляются из популяции («умирают»), что является содержанием операции отбора (блок 4).

Завершение работы ГАО (блок 5) осуществляется тогда, когда популяция выходит на устойчивое состояние, в котором особь с максимальным значением функции пригодности принимается за окончательное решение задачи. В противном случае (блок 6) осуществляется переход к новому поколению популяции на блок 2 и выполнение указанных выше операций ГАО повторяется.

Покажем, каким образом возможно применение метода ГАО к синтезу *VPN*-сети.

Введем матрицу **X** булевых переменных таких, что $x_{ij} = 1$, если между КМ-*i* и КМ-*j* существует *IP*-канал, и $x_{ij} = 0$ – в противном случае. Матрица **X**, таким образом, полностью определяет конкретный вариант топологии сети.

Так как **X** является симметричной матрицей, по главной диагонали которой всегда расположены нулевые элементы, предложим структуру хромосомы особи в следующем виде:

$$Xp = \{x_{12}, x_{13}, \ldots, x_{1N}; x_{23}, \ldots, x_{2N}; \ldots; x_{(N-1)N}\}. \quad (1)$$

Как легко проверить, если взять две различные хромосомы вида (1) и выполнить над ними операцию кроссовера, то в результате всегда будет получаться дочерняя хромосома, соответствующая некоторому реальному варианту построения сети *VPN*. Тем самым окончательно подтверждается, что применение ГАО для решения синтеза *VPN*-сети возможно. Необходимо только определить вид функции пригодности и ее зависимость от **X**.

Расчет показателя пригодности. Построение функции пригодности в случае многокритериального синтеза является отдельной достаточно сложной задачей. В настоящей статье ограничимся случаем с одним критерием пригодности. Выберем в качестве данного критерия показатель, характеризующий пропускную способность сети. Данный показатель, по мнению автора, обладает наибольшей сложностью расчета.

Тогда выбранный показатель будет являться функцией пригодности ГАО.

Покажем, как можно оценить пропускную способность *VPN*-сети в зависимости от значений, которые принимают элементы матрицы **X**.

Исходим из того, что пропускная способность *VPN*-сети определяется суммой пропускных способностей всех ее составляющих *VPN*-каналов.

Для оценки пропускной способности *VPN*-каналов предлагается следующая модель, основанная на теории массового обслуживания.

Представим простой *VPN*-канал сетью массового обслуживания (СеМО), состоящей из трех последовательно соединенных систем массового обслуживания (СМО): начального КМ (СМО-1), *IP*-канала (СМО-2) и конечного КМ (СМО-3), как показано на рисунке 3.

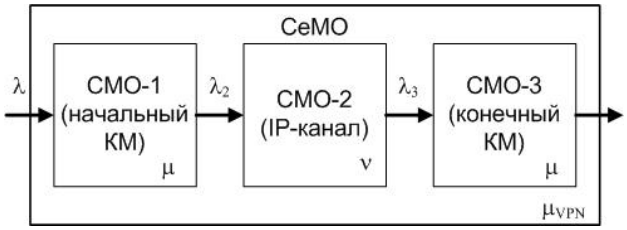

Рисунок 3 – Модель простого *VPN*-канала

Введем необходимые обозначения. На вход СМО-1 (и, следовательно, в целом СеМО) поступает нагрузка λ. Нагрузку на СМО-2 и СМО-3 обозначим соответственно λ_2 и λ_3.

В начальном и конечном КМ *VPN*-канала используются одинаковые алгоритмы шифрования. Поэтому интенсивности обслуживания СМО-1 и СМО-3 равны. Обозначим их μ. Интенсивность обслуживания СМО-2 обозначим ν. По сути, она представляет собой пропускную способность *IP*-канала.

Постановка задачи сводится к нахождению интенсивности обслуживания СеМО, обозначенной μ_{VPN}, являющейся, по сути, пропускной способностью простого *VPN*-канала.

Для нахождения μ_{VPN} необходимо рассчитать суммарное время задержки поступающих в СеМО заявок.

Для расчета времени задержки $T_{зад1}$ на СМО-1 воспользуемся известной теоремой Литтла [3], которая утверждает, что для СМО-1 справедливо следующее выражение:

$$T_{зад1} = \frac{1}{\mu - \lambda}. \qquad (2)$$

Обозначим коэффициент использования СМО-1 через $\rho = \lambda/\mu$. Тогда (2) перепишется в виде

$$T_{зад1} = \lambda^{-1}\rho(1-\rho)^{-1}. \qquad (3)$$

Используя ρ, можно следующим образом определить нагрузку λ_2:

$$\lambda_2 = \lambda(1-\rho). \qquad (4)$$

Тогда время задержки $T_{зад2}$ на СМО-2 будет равняться

$$T_{зад2} = \frac{1}{\nu - \lambda(1-\rho)}. \qquad (5)$$

Обозначим $\eta = \lambda/\nu$. Тогда (5) перепишется в виде

$$T_{зад2} = \lambda^{-1}\eta(1-\eta(1-\rho))^{-1}. \qquad (6)$$

Обозначим $\rho_2 = \lambda_2/\nu$. Используя (4), для расчета ρ_2 имеем

$$\rho_2 = (1-\rho)\eta. \qquad (7)$$

Для определения времени задержки $T_{зад3}$ на СМО-3 воспользуемся следующими выражениями:

$$\lambda_3 = (1-\rho_2)\lambda_2 = (1-\rho)(1-\eta(1-\rho))\lambda; \qquad (8)$$

$$T_{зад3} = \frac{1}{\mu - \lambda_3} = \lambda^{-1}\rho(1-\rho(1-\rho)(1-\eta(1-\rho)))^{-1}. \qquad (9)$$

Суммарное время задержки заявок в СеМО определяется путем сложения $T_{зад1}$, $T_{зад2}$ и $T_{зад3}$:

$$T_{задVPN} = T_{зад1} + T_{зад2} + T_{зад3}. \qquad (10)$$

Подставляя в (10) результаты, полученные в (3), (6) и (9), приходим к окончательному виду выражения для суммарного времени задержки:

$$T_{задVPN} = \lambda^{-1}(\rho(1-\rho)^{-1} + \eta(1-\eta(1-\rho))^{-1} + \\ + \rho(1-\rho(1-\rho)(1-\eta(1-\rho)))^{-1}). \qquad (11)$$

Запишем теперь формулу Литтла для простого *VPN*-канала как СМО, подставляя в нее $T_{задVPN}$ и μ_{VPN}:

$$T_{задVPN} = \frac{1}{\mu_{VPN} - \lambda}. \qquad (12)$$

Обозначим $\rho_{VPN} = \lambda/\mu_{VPN}$.

Выразим ρ_{VPN}, используя (12):

$$\rho_{VPN} = 1 - \frac{1}{1 + \lambda T_{задVPN}}. \qquad (13)$$

Подставляя (11) в (13) и делая необходимые преобразования, получаем искомый вид выражения, определяющего коэффициент использования простого *VPN*-канала как СМО:

$$\rho_{VPN} = 1 - \\ - \frac{1}{1 + \frac{\rho}{1-\rho} + \frac{\eta}{1-\eta(1-\rho)} + \frac{\rho}{1-\rho(1-\rho)(1-\eta(1-\rho))}}. \qquad (14)$$

Рассмотрим теперь случай с составным *VPN*-каналом (рисунок 4).

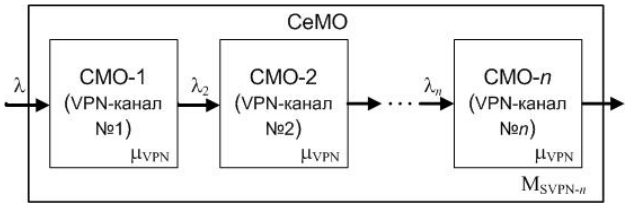

Рисунок 4 – Модель составного VPN-канала

Без потери общности можно полагать, что все СМО, входящие в состав СеМО,

отображающей составной *VPN*-канал, являются однородными. Обозначим $\rho = \lambda/\mu_{VPN}$.

Теперь воспользуемся выражением (14). Подставим в него ρ вместо η и выполним после этого над ним необходимые преобразования.

В результате получим следующее выражение для коэффициента использования $P_{SVPN\text{-}3} = \lambda/\mathbf{M}_{VPN-3}$ составного *VPN*-канала с двумя транзитами:

$$\mathbf{P}_{SVPN-3} = 1 - \frac{1}{1+\frac{\rho}{1-\rho}+\frac{\rho}{1-\rho(1-\rho)}+\frac{\rho}{1-\rho(1-\rho)(1-\rho(1-\rho))}}. \quad (15)$$

Переходя от *VPN*-канала с *i* транзитами к каналу с (*i*+1) транзитом, можно показать, что для расчета пропускной способности составного *VPN*-канала с (*n*–1) транзитами можно использовать выражение

$$\mathbf{P}_{SVPN-n} = 1 - \frac{1}{1+\sum_{i=1}^{n}\frac{\rho}{\tau_i}}, \quad (16)$$

где $\tau_1 = 1-\rho$, а значение τ_i определяется рекурсивно через τ_{i-1} согласно следующему выражению:

$$\tau_i = 1-(1-\tau_{i-1})\tau_{i-1}. \quad (17)$$

Таким образом, если известно количество транзитов для каждого составного *VPN*-канала, а также нагрузка в каждом информационном направлении сети, то, используя (14)–(17), легко можно найти пропускную способность всей *VPN*-сети.

Покажем, как решается задача поиска числа транзитов составных *VPN*-каналов с учетом значений элементов матрицы **X**.

Введем для *VPN*-канала понятие *ранга r* таким образом, что простой канал имеет $r = 1$, составной канал с одним транзитом – $r = 2$, с двумя – $r = 3$ и т.д.

Перейдем от матрицы **X** к *матрице рангов* **R** следующим образом: если $x_{ij} = 1$, то $r_{ij} = 1$; если $x_{ij} = 0$, то $r_{ij} = \infty$.

Определение матрицы $\mathbf{R}_{\min}$ минимальных рангов путей, существующих между узлами сети, осуществляется путем последовательного возведения в степень исходной матрицы **R** и замены процедуры перемножения матриц на операцию Δ [4]. Операция Δ над двумя квадратными матрицами $\mathbf{A} = \|\alpha_{ij}\|$ и $\mathbf{B} = \|\beta_{ij}\|$ порядка *N* задается следующим образом: вхождение γ_{ij} новой матрицы $\mathbf{C} = \mathbf{A}\,\Delta\,\mathbf{B}$ определяется как минимум из почленных сумм *i* -й строки матрицы **A** и *j*-го столбца матрицы **B**:

$$\gamma_{ij} = \min\left(\alpha_{i1}+\beta_{1j}, \alpha_{i2}+\beta_{2j}, ..., \alpha_{iN}+\beta_{Nj}\right). \quad (18)$$

В результате матрица $\mathbf{R}^{2\Delta} = \mathbf{R}\,\Delta\,\mathbf{R}$ показывает все пути, имеющие минимальный ранг не выше двух, матрица $\mathbf{R}^{3\Delta} = \mathbf{R}^{2\Delta}\,\Delta\,\mathbf{R}$ – все пути ранга, не превышающего три, и т.д.

Возведение матрицы **R** в степень с помощью операции Δ продолжается до тех пор, пока из результирующей матрицы полностью не исчезнут элементы, равные ∞.

В частности, для примера, представленного на рисунке 1, матрицы **R** и $\mathbf{R}_{\min}$ имеют следующий вид:

$$R = \begin{bmatrix} 0 & 1 & \infty & 1 & 1 \\ 1 & 0 & \infty & \infty & \infty \\ \infty & \infty & 0 & 1 & \infty \\ 1 & \infty & 1 & 0 & 1 \\ 1 & \infty & \infty & 1 & 0 \end{bmatrix}, \quad (19)$$

$$R_{\min} = \begin{bmatrix} 0 & 1 & 2 & 1 & 1 \\ 1 & 0 & 3 & 2 & 2 \\ 2 & 3 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 \end{bmatrix}. \quad (20)$$

Максимальный ранг в рассматриваемом варианте структуры *VPN*-сети равняется $r = 3$ и наблюдается в информационном направлении (КМ-2, КМ-*i*).

С учетом матрицы минимальных рангов определяется число транзитов и рассчитываются коэффициенты использования *VPN*-каналов согласно (16) – (17). Зная обрабатываемую сетью нагрузку $\mathbf{\Lambda} = \|\lambda_{ij}\|$, переходим от коэффициентов использования к пропускным способностям вначале *VPN*-каналов, а затем – всей сети. Тем самым определяется функция пригодности ГАО в зависимости от матрицы **X**.

Апробация метода. Для практической проверки эффективности предлагаемого метода синтеза *VPN*-сети были взяты две сети: одна состояла из $N_1 = 5$ узлов, а вторая – из $N_2 = 10$ узлов.

Проверка предлагаемого метода по результативности осуществлялась путем решения задачи синтеза в случае задания таких вариантов исходных данных, которые соответствуют заранее известным решениям.

Роль таких варьируемых исходных данных играет в поставленной задаче синтеза матрица $\mathbf{\Lambda} = \|\lambda_{ij}\|$, элементы которой определяют величину обрабатываемой нагрузки в соответствующих *VPN*-направлениях.

Задача синтеза решалась для трех вариантов задания исходных данных по обрабатываемой сетью нагрузке, которые априори соответствуют возможным граничным вариантам топологии *VPN*-сети:

кольцевой;

радиальной;

полносвязной.

Матрица $\Lambda = \|\lambda_{ij}\|$, задающая требования по обрабатываемой сетью нагрузке, для кольцевой топологии в случае $N_1 = 5$ имела следующий вид:

$$\Lambda_{\text{кол}} = \begin{bmatrix} 0 & 100 & 0 & 0 & 100 \\ 100 & 0 & 100 & 0 & 0 \\ 0 & 100 & 0 & 100 & 0 \\ 0 & 0 & 100 & 0 & 100 \\ 100 & 0 & 0 & 100 & 0 \end{bmatrix}. \quad (21)$$

Матрица обрабатываемой нагрузки в случае $N_1 = 5$ для радиальной топологии имела вид

$$\Lambda_{\text{рад}} = \begin{bmatrix} 0 & 100 & 100 & 100 & 100 \\ 100 & 0 & 0 & 0 & 0 \\ 100 & 0 & 0 & 0 & 0 \\ 100 & 0 & 0 & 0 & 0 \\ 100 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (22)$$

а для полносвязной топологии:

$$\Lambda_{\text{полн}} = \begin{bmatrix} 0 & 100 & 100 & 100 & 100 \\ 100 & 0 & 100 & 100 & 100 \\ 100 & 100 & 0 & 100 & 100 \\ 100 & 100 & 100 & 0 & 100 \\ 100 & 100 & 100 & 100 & 0 \end{bmatrix}. \quad (23)$$

Поиск искомого решения задачи с помощью ГАО для случаев $N_1 = 5$ и $N_2 = 10$ продолжался до получения такой популяции особей, в которой особь, имеющая наибольшее значение функции пригодности, имела хромосому, соответствующую выбранному варианту граничной топологии сети.

Вероятности мутации и скрещивания особей имели соответственно следующие значения: $P_{\text{мут}} = 0{,}01$ и $P_{\text{скр}} = 0{,}1$. Количество особей в популяции ограничивалось числом 200.

Отбор особей для выполнения над ними операций мутации и скрещивания осуществлялся методом "колеса рулетки", сущность которого заключалась в следующем.

Особи, являющиеся кандидатами на отбор, выбирались для мутации и скрещивания путем использования колеса рулетки, на котором каждая особь S_i, принадлежащая популяции, представляется в виде сектора, ширина которого пропорциональна соответствующему значению функции пригодности $f(S_i)$.

В результате особи, имеющие большую пригодность, соответствуют большему сектору на колесе, а особи с меньшим значением функции пригодности получают меньший сектор.

Процедура отбора сводилась к вращению колеса рулетки M раз, где M – количество особей в популяции.

Доля на колесе Δf_i, выделенная для i-й особи S_i, вычислялась по формуле

$$\Delta f_i = \frac{f(S_i)}{\sum_{j=1}^{M} f(S_j)}. \quad (24)$$

В качестве кандидатов на мутацию и скрещивание из текущего поколения выбирались те особи $S_1, S_2, ..., S_m$, на которые попадал указатель рулетки после ее вращения.

Применение ГАО для решения задачи эволюционного синтеза *VPN*-сети привело к получению следующих результатов. Достижение искомых граничных топологий сети в среднем происходило:

для $N_1 = 5$ – на 15 поколении;

для $N_2 = 10$ – на 35 поколении.

Время обработки одного поколения на персональном компьютере с процессором *P*-IV и оперативной памятью 1 Гбайт составляло 0,5 минуты. Это свидетельствует о том, что предлагаемый метод эволюционного синтеза *VPN*-сети вполне позволяет получать оптимальные решения в масштабе реального либо близкого к нему времени.

Заключение. В статье рассмотрен порядок применения ГАО как метода эволюционного синтеза *VPN*-сетей. Получены необходимые выражения, которые позволяют найти топологию *VPN*-сети, отвечающую требованиям ИТКС по пропускной способности.

Библиографический список

1. *Балашов Е.П.* Эволюционный синтез систем. – М.: Радио и связь, 1985. – 328 .

2. *Емельянов В.В., Курейчик В.В., Курейчик В.М.* Теория и практика эволюционного моделирования. – М.: ФИЗМАТЛИТ, 2003. – 432 с.

3. *Бертсекас Д., Галлагер Р.* Сети передачи данных: пер. с англ. – М.: Мир, 1989. – 544 с.

4. *Давыдов Г.Б.* и др. Сети электросвязи. – М.: Связь, 1977.