УДК 681.518

И.Б. Саенко, И.В. Котенко

ПРИМЕНЕНИЕ РОЛЕВОГО ПОДХОДА И ГЕНЕТИЧЕСКОЙ ОПТИМИЗАЦИИ ДЛЯ ПРОЕКТИРОВАНИЯ VLAN В БОЛЬШИХ КРИТИЧЕСКИХ ИНФРАСТРУКТУРАХ[1]

Саенко Игорь Борисович, *доктор технических наук, профессор, окончил Военную академию связи (ВАС) им. С.М. Буденного. Ведущий научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Имеет статьи, монографии и патенты в области проблем компьютерной безопасности, методов искусственного интеллекта и информационно-телекоммуникационных систем. [e-mail: ibsaen@comsec.spb.ru].*

Котенко Игорь Витальевич, *доктор технических наук, профессор, окончил Военно-космическую академию им. А.Ф. Можайского и ВАС им. С.М. Буденного. Заведующий лабораторией проблем компьютерной безопасности СПИИРАН. Имеет статьи, монографии и патенты в области проблем компьютерной безопасности, методов искусственного интеллекта и информационно-телекоммуникационных систем. [e-mail: ivkote@comsec.spb.ru].*

Аннотация

В статье рассматривается новый подход к формированию схем разграничения доступа при проектировании виртуальных локальных вычислительных сетей (VLAN) в больших критических инфраструктурах, основанный на использовании усовершенствованного генетического алгоритма, введении множества ролей для узлов сети и учете отображения между множеством узлов сети и множеством ролей. Обосновывается постановка задачи оптимизации схем разграничения доступа в VLAN по критерию минимального количества виртуальных подсетей с учетом особенностей ролевого подхода. Показано, что решаемая задача относится к классу задач булевой матричной факторизации, в которой исходная матрица декомпозируется на четыре, образуя две независимые пары из прямых и транспонированных матриц. Усовершенствования генетического алгоритма связаны с мультихромосомным представлением особей, новым видом функции пригодности и двумерным видом операции скрещивания. Экспериментальная оценка показала выигрыш в скорости работы ролевого генетического алгоритма при большой размерности задачи до 5 раз при достаточно высокой точности поиска оптимального решения.

Ключевые слова: VLAN, разграничение доступа, генетический алгоритм, критическая инфраструктура.

doi: 10.35752/1991-2927-2020-2-60-81-88

1 Работа выполнена при частичной финансовой поддержке РФФИ (проекты 18-07-01369, 18-07-01488) и бюджетной темы 0073-2019-0002.

USING A ROLE-BASED APPROACH AND GENETIC OPTIMIZATION FOR VLAN DESIGN IN LARGE CRITICAL INFRASTRUCTURES

Igor Borisovich Saenko, *Doctor of Sciences in Engineering, Professor; graduated from the Marshal Budjonny Military Academy of Signal Corps; Leading Researcher of the Laboratory of Computer Security Problems of St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS); an author of articles, monographs, and patents in the field of computer security problems, artificial intelligence methods, information and telecommunication systems. e-mail: ibsaen@comsec.spb.ru.*

Igor Vitalevich Kotenko, *Doctor of Sciences in Engineering, Professor; graduated from the A.F. Mozhaisky Military Space Academy and the Marshal Budjonny Military Academy of Signal Corps; Head of the Laboratory of Computer Security Problems of SPIIRAS; an author of articles, monographs, and patents in the field of computer security problems, artificial intelligence methods, information and telecommunication systems. e-mail: ivkote@comsec.spb.ru.*

Abstract

The article considers a new approach to the formation of access control schemes in the design of virtual local area networks (VLANs) in large critical infrastructures, based on the use of an improved genetic algorithm, the introduction of multiple roles for network nodes and the accounting of mapping between a set of network nodes and a set of roles. The problem statement is justified as the optimization of access control schemes in VLAN according to the criterion of minimum number of virtual subnets taking into account the peculiarities of role approach. The problem being solved is shown to belong to the Boolean matrix factorization problem class, in which the original matrix is decomposed into four ones forming two independent pairs of straight and transposed matrices. Improvements of the genetic algorithm involve multi-chromosome representation of individuals, a new kind of fitness function, and a two-dimensional kind of crossing operation. The experimental evaluation showed a benefit in the speed of the role genetic algorithm at a large task dimension of up to 5 times at a sufficiently high accuracy of finding the optimal solution.

Key words: VLAN, access control, genetic algorithm, critical infrastructure.

Введение

В настоящее время резко возрастает множество угроз безопасности, которые воздействуют на критические инфраструктуры, а также значимость успешной защиты сетевых ресурсов в таких системах [1]. Так как взаимодействие с внешними сетями (Интернетом) в критических инфраструктурах, как правило, запрещено или находится под надежным контролем, основными нарушителями безопасности в таких системах являются внутренние пользователи (инсайдеры). Инсайдер может случайно или преднамеренно попытаться сделать несанкционированный доступ к другому компьютеру, к которому этот доступ для него запрещен. Применение технологии виртуальных локальных вычислительных сетей (VLAN, от англ. Virtual Local Area Networks) может обеспечить гарантированную защиту от такой попытки [2]. Технология VLAN позволяет разграничить доступ к сетевым элементам за счет формирования виртуальных подсетей в локальной вычислительной сети (ЛВС).

Суть проектирования VLAN заключается в рациональном распределении компьютеров по создаваемым подсетям. Это распределение описывается матрицей «компьютеры – подсети», которую будем называть схемой разграничения доступа в VLAN (СРД-VLAN). Формирование этой схемы, отвечающей заданным критериям безопасности, является NP-полной задачей. Для эффективного решения этой задачи необходимы эвристические методы, в частности, генетические алгоритмы (ГА), широко применяющиеся для решения похожих задач оптимизации. В предыдущих работах нами было показано, как можно применять ГА для оптимизации различных схем разграничения доступа, включая VLAN [3–6]. Однако для больших критических инфраструктур решение этой задачи известными способами [5, 6] является весьма трудоемким ввиду ее большой размерности.

В статье предлагается использовать ролевой подход для оптимизации СРД-VLAN. Это объясняется тем, что на практике в матрице «компьютеры – подсети» большой размерности всегда имеются группы компьютеров, обладающих одинаковыми связями с виртуальными подсетями. Здесь просматривается аналогия с ролевой моделью разграничения доступа RBAC (от англ. Role-Based Access Control), которая применяется для разграничения доступа в базах данных. В результате при поиске оптимальной СРД-VLAN каждая особь в ГА будет иметь две короткие хромосомы вместо одной большой. Этот подход, по аналогии с оптимизацией схемы RBAC [3, 4], должен приводить к более высокой сходимости ГА.

Сравнительная оценка результатов решения задачи оптимизации ролевой СРД-VLAN (РСРД-VLAN) является одной из основных целей настоящей работы. Другими целями являются: постановка задачи оптимизации РСРД-VLAN и выработка метода ее решения; разработка

и усовершенствование ГА для решения задачи в предложенной постановке; экспериментальная оценка предложенного ГА на разработанном инструментальном стенде.

1 Анализ известных работ

Задача оптимизации СРД-VLAN относится к классу задач булевой матричной факторизации (БМФ). В работах [7, 8] доказано, что все задачи вида БМФ являются NP-полными и требуют разработки эвристических методов для своего решения. Так в [9] предлагается решать задачи БМФ с помощью биоинспирированных алгоритмов, основанных на популяциях. При этом показано, что для решения задач БМФ наибольшей эффективностью обладают ГА. В [10, 11] для решения задачи БМФ предложен ГА, в котором функция пригодности основана на евклидовом расстоянии между начальной и результирующей матрицами. В [12] для решения задачи неотрицательной матричной факторизации (НМФ) предложен ГА со специфическим оператором мутации. Однако, несмотря на то, что эти алгоритмы показали высокую эффективность при решении различных задач БМФ и НМФ, они не могут применяться для оптимизации СРД-VLAN, так как не обеспечивают требуемое совпадение результирующих матриц.

Анализ работ по проектированию VLAN показывает, что, как правило, формирование виртуальных подсетей происходит эмпирически без решения оптимизационных задач [13]. Во многом это объясняется высокой сложностью таких задач. Однако в некоторых работах делаются попытки формального решения таких задач. Так в [14] предложено применять кластерный анализ для формирования СРД-VLAN. Однако этот подход ориентирован в основном на мобильные сети. В [15] предложено использовать ГА для оптимизации схемы доступа в ЛВС, в которой применяется технология VLAN. Однако этот алгоритм находит матрицу связности компьютеров, а не распределение компьютеров по виртуальным подсетям. В [16] предложена постановка задачи оптимизации СРД-VLAN по критерию минимума энергетических затрат. Предложен алгоритм, основанный на кластеризации методом k-средних. Показано, что он является более эффективным, чем метод нелинейного целочисленного программирования. Однако для проектирования VLAN этот подход не приемлем, так как используемый критерий не ориентирован на безопасность. В [17] рассмотрена многокритериальная задача проектирования VLAN по показателям стоимости трафика в сети и стоимости поддержки покрывающих деревьев для каждой виртуальной подсети. Однако эта задача также не учитывает критерий безопасности. В [18] задача проектирования VLAN ориентирована на безопасность. Однако ее решение связано с реализацией прикладной архитектуры безопасности VLAN, основанной, в частности, на протоколе IPSec, что ограничивает область применения этих решений.

В наших предыдущих работах [5, 6] были разработаны некоторые постановки задачи оптимизации СРД-VLAN, основанные на критериях безопасности, и предложен для их решения ГА. Однако для задачи большой размерности, свойственной большим критическим инфраструктурам, скорость работы этого ГА резко уменьшается. По этой причине возникла идея применить для задачи проектирования VLAN ролевой подход, который применяется в модели RBAC. В [19] показано, что задача оптимизации схемы RBAC так же, как и задача проектирования VLAN, относится к задачам БМФ. В [3, 4] показано, как для оптимизации схемы RBAC могут применяться ГА. Таким образом, объединение ролевого подхода и ГА должно привести к разработке более эффективного метода решения задачи проектирования VLAN большой размерности.

2 Постановка задачи оптимизации РСРД-VLAN

Пусть в компьютерной сети находится множество компьютеров $C=\{C_i\}$, $i=1, ..., n$, и между этими компьютерами задана схема разрешенных информационных потоков, определяемая булевой матрицей $\mathbf{A}[n, n]$. Если $a_{ij}=1\,(i, j=1, ..., n)$, то обмен между компьютерами i и j разрешен. В противном случае этот обмен невозможен. Далее, пусть в компьютерной сети создано множество виртуальных подсетей $V=\{V_j\}$, $j=1, ..., k$. Каждая из этих подсетей объединяет два и более компьютера. Зададим распределение компьютеров по подсетям с помощью матрицы $\mathbf{Z}[n, k]$. Если $z_{ij}=1$, то компьютер i принадлежит подсети j. В противном случае подсеть j не охватывает компьютер i. Как было показано в [5, 6], матрица $\mathbf{Z}$ связана с матрицей $\mathbf{A}$ следующим выражением:

$$\mathbf{A}=\mathbf{Z}\otimes\mathbf{Z}^T, \quad (1)$$

где $\mathbf{Z}^T$ – транспонированная матрица $\mathbf{Z}$, символ $\otimes$ обозначает булево матричное умножение, которое является формой матричного умножения, основанной на правилах булевой алгебры. Булево матричное умножение позволяет получать элементы матрицы $\mathbf{A}$ согласно следующему выражению: $a_{ij}=\vee_{j=1}^{n}\left(z\wedge z_{ij}\right)$.

Если рассматривать матрицу $\mathbf{Z}$ как переменную величину в (1), а матрицу $\mathbf{A}$ – как заданную, то (1) является задачей БМФ. Легко заметить, что (1) имеет большое множество возможных решений. Однако в задаче БМФ требуется среди возможных решений выбрать такую матрицу $\mathbf{Z}[n, k]$, для которой размерность k, определяющая количество виртуальных подсетей, является минимальной. Выбор критерия минимального количества подсетей имеет большое значение для повышения эффективности проектирования VLAN в целом. Во-первых, меньшее количество виртуальных подсетей обеспечивает меньший суммарный объем трафика в сети. Во-вторых, для VLAN с меньшим количеством подсетей свойственно более удобное администрирование. Это, в свою очередь, повышает достоверность реализации политик безопасности и, в конечном итоге, безопасность всей сети.

Если задача (1) имеет большую размерность, то ее решение будет занимать очень много времени. В то же время, при большой размерности (1) элементы множества $\mathbf{C}$ можно объединить в группы, которые мы будем называть *ролями*. В этих группах каждый элемент имеет равную принадлежность к некоторому подмножеству виртуальных подсетей $\mathbf{V}$. Обозначим множество ролей как $R = \{R_s\}$, $s = 1, ..., r$. Как и в модели RBAC, между элементами множеств $\mathbf{C}$ и $\mathbf{R}$ существует отношение «многие-ко-многим». Опишем его с помощью матрицы $\mathbf{X}[n, r] = \{x_{is}\}$, в которой $x_{is} = 1$, если компьютер i принадлежит роли s, и $x_{is} = 0$ – в противном случае. Отношение «многие-ко-многим» между множествами $\mathbf{R}$ и $\mathbf{V}$ зададим с помощью матрицы $\mathbf{Y}[r, k]$. Если $y_{sj} = 1$, то роль s связана с подсетью j. В противном случае роль s не связана с подсетью j.

Матрицы $\mathbf{A}$, $\mathbf{X}$ и $\mathbf{Y}$ связаны друг с другом с помощью следующего выражения:

$$\mathbf{A} = \mathbf{X} \otimes \mathbf{Y} \otimes \mathbf{Y}^T \otimes \mathbf{X}^T, \quad (2)$$

где $\mathbf{Y}^T$ – транспонированная матрица $\mathbf{Y}$. Пример отображений между множествами компьютеров, ролей и подсетей приведен на рисунке 1. На нем видно, что всем компьютерам присвоены три роли. К роли R_1 относятся компьютеры C_1 и C_4, к роли R_2 – C_2 и C_4, к роли R_3 – C_3 и C_5. При этом роли R_1 и R_2 образуют подсеть V_1, а роли R_2 и R_3 – подсеть V_2.

Будем рассматривать матрицы $\mathbf{X}$ и $\mathbf{Y}$ как переменные величины в (2), а матрицу $\mathbf{A}$ – как заданную. Тогда (2) является постановкой задачи БМФ, в которой исходная матрица разбивается на четыре сомножителя. Как и (1), задача (2) имеет большое множество возможных решений. Будем искать такие решения, чтобы размерности r и k были минимальными. Тогда постановка задачи оптимизации РСРД-VLAN имеет следующий вид. Исходными данными в этой задаче является заданная матрица $\mathbf{A}[n, n]$. Требуется найти матрицы $\mathbf{X}[n, r]$ и $\mathbf{Y}[r, k]$ такие, чтобы было справедливо (2), а размерности r и k были минимальны. Формально эта постановка может быть записана следующим образом:

$$\begin{cases} \mathbf{A} = \mathbf{X} \otimes \mathbf{Y} \otimes \mathbf{Y}^T \otimes \mathbf{X}^T, \\ \mathrm{Dim}(\mathbf{X}, 2) \to \min, \\ \mathrm{Dim}(\mathbf{Y}, 2) \to \min, \end{cases} \quad (3)$$

где $\mathrm{Dim}(\mathbf{X}, q)$ – функция, определяющая q-ю размерность в матрице $\mathbf{X}$.

3 Метод генетической оптимизации РСРД-VLAN

Предложим для решения задачи (3) использование ГА. Выбор этого метода для решения традиционной задачи оптимизации структуры VLAN, задаваемой выражением (1), был обоснован нами в работах [5, 6]. Для решения задачи (3) предлагается усовершенствовать применявшийся ранее ГА за счет следующих аспектов: использования не одной, а двух хромосом у особей; нового вида функции пригодности, учитывающим критерии задачи (3); мультихромосомного механизма операции скрещивания. Рассмотрим эти особенности подробнее.

Так как решение задачи (3) полностью определяется двумя матрицами $\mathbf{X}$ и $\mathbf{Y}$, а эти матрицы не зависят друг от друга, то вполне логично для кодирования решений в ГА использовать две хромосомы: хромосома $Chr\ (\mathbf{X})$ будет кодировать матрицу $\mathbf{X}$, а хромосома $Chr\ (\mathbf{Y})$ – матрицу $\mathbf{Y}$.

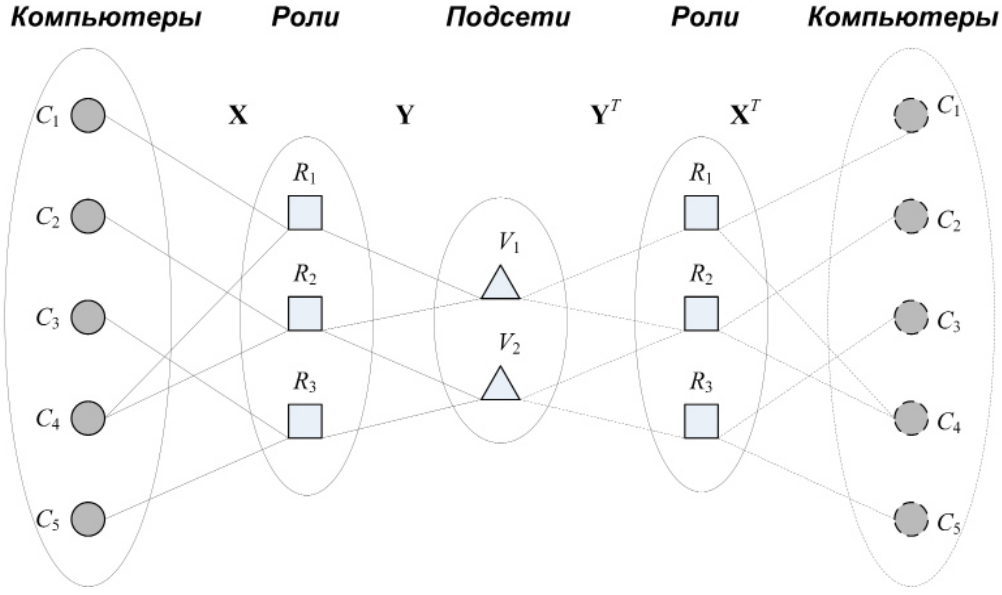

Рис. 1. Пример отображений между множествами компьютеров, ролей и подсетей

Матрица $\mathbf{X}$ имеет размерность $n \times r$ (n – количество строк, r – количество столбцов). Она представляется в виде $\mathbf{X} = \{X_s\}$, где $X_s = (x_{s1}, x_{s2}, ..., x_{sn})$ – s-й столбец матрицы ($s = 1, ..., r$). Хромосому $Chr\ (\mathbf{X})$ определим как строку, состоящую из генов. В качестве гена используется столбец X_s матрицы $\mathbf{X}$. Формально это записывается следующим образом:

$$Chr(\mathbf{X}) = \{gene_s(\mathbf{X})\},\ gene_s(\mathbf{X}) = X_s,\ s = 1, ..., r. \quad (4)$$

Аналогичным образом формируется вторая хромосома $Chr\ (\mathbf{Y})$, предназначенная для кодирования матрицы $\mathbf{Y}$. Размерность матрицы $\mathbf{Y}$ равна $r \times k$. Столбец j матрицы $\mathbf{Y}$ имеет вид $Y_j = (y_{j1}, y_{j2}, ..., y_{jr})$, $j = 1, ..., k$. Формально эта хромосома имеет следующий вид:

$$Chr(\mathbf{Y}) = \{gene_j(\mathbf{Y})\},\ gene_j(\mathbf{Y}) = Y_j,\ j = 1, ..., k. \quad (5)$$

Из (5) и (6) видно, что обе хромосомы имеют переменную длину, которая в ходе работы алгоритма постепенно уменьшается, достигая к концу его работы минимума. Этот аспект учитывается при реализации механизма выполнения оператора скрещивания.

Анализируя постановку задачи (3), можно сделать вывод, что промежуточные решения в ГА должны сначала обеспечить равенство матричного произведения матрице $\mathbf{A}$, затем обеспечить минимум значения k, после чего обеспечить минимум значения r. Следовательно, функцию пригодности можно представить в следующем виде:

$$F = \alpha F_1 + \beta F_2 + \gamma F_3, \quad (6)$$

где F_1 – функция, которая отражает полное совпадение матричного произведения и матрицы $\mathbf{A}$; F_2 – функция, отвечающая за минимизацию k; F_3 – функция, отвечающая за минимизацию r; α, β и γ – весовые коэффициенты, управляющие направлением поиска решения. Положим, что наилучшими решениями являются те, у которых значение функции пригодности минимально. Тогда между весовыми коэффициентами справедливо следующее соотношение: $\alpha \gg \beta \gg \gamma$. Это соотношение гарантирует, что в первую очередь происходит поиск решений, у которых $F_1 = 0$, затем поиск решений с минимальным значением F_2 и, наконец, поиск решений с минимальным значением F_3.

Функция F_1 будет иметь следующий вид:

$$F_1 = \sum_{i=2}^{n} \sum_{j=1}^{n-1} \left(a_{ij} - \sum_{s=1}^{k} z_{is} z_{sj} \right), \quad z_{ij} = \sum_{s=1}^{r} x_{is} x_{sj}. \quad (7)$$

Из (7) видно, что вычисление F_1 происходит через предварительное вычисление матрицы $\mathbf{Z} = \mathbf{X} \otimes \mathbf{Y}$, определяющей отношение между множествами $\mathbf{C}$ и $\mathbf{V}$. Кроме того, так как матрица $\mathbf{A}$ является симметричной, то вычисления в (7) выполняются только по элементам матрицы $\mathbf{A}$, лежащим выше главной диагонали. Заметим также, что в (7) все операции суммирования и умножения являются логическими операциями AND и OR.

Функции F_2 и F_3 имеют более простой вид. Они определяют, соответственно, значения k и r. Поэтому их можно задать с помощью следующих выражений:

$$F_2 = k; F_3 = r. \quad (8)$$

Операция скрещивания позволяет получить из пары особей-родителей, которые выбираются из текущей популяции с вероятностью W_{cross}, новые особи-потомки путем обмена частями родительских хромосом. В традиционном ГА, в котором для кодирования решений используется одна хромосома, в результате скрещивания образуется две особи-потомка. В нашем случае, когда решение кодируется двумя хромосомами, в результате скрещивания образуется четыре особи-потомка, как показано на рисунке 2.

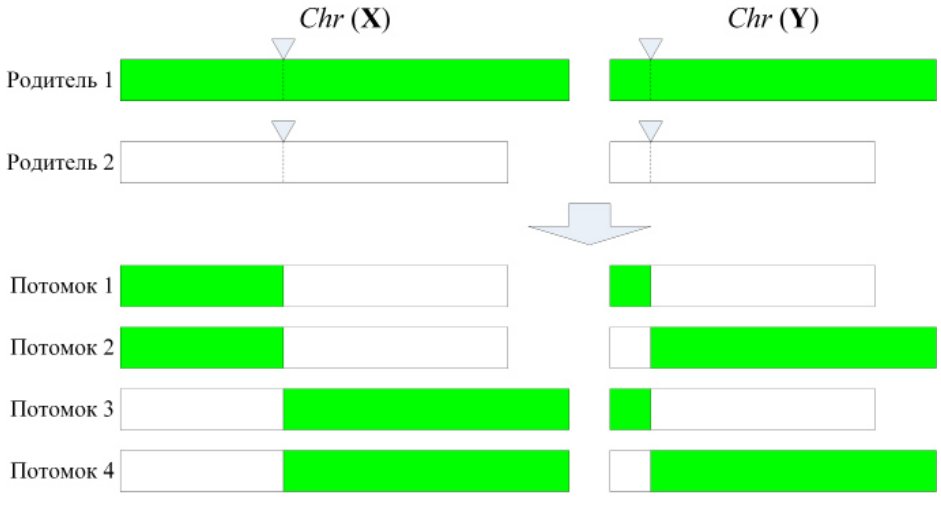

Рис. 2. Образование потомков в операции скрещивания при двух хромосомах

Для мутации выбирается особь из текущей популяции с вероятностью W_{mut}. Операцию мутации предлагается выполнять в два этапа. На первом этапе, в соответствии с традиционным подходом, с заданной вероятностью W_{gen} выбираются для мутации гены хромосом – столбцы матриц $\mathbf{X}$ и $\mathbf{Y}$. На втором этапе с вероятностью W_{el} производится инвертирование элементов выбранных столбцов. Вероятности W_{mut}, W_{gen} и W_{el} являются параметрами ГА.

4 Реализация и экспериментальная оценка метода

Эффективность работы ГА определялась двумя показателями: скоростью работы и точностью. Первый показатель – количество итераций ГА, потребовавшихся для решения задачи. Второй показатель показывал, насколько результирующие решения задачи VLAN отличаются от эталонных. Он вычислялся в соответствии со следующим выражением:

$$\theta = 1 - \frac{|k_0 - k|}{k_0}, \quad (9)$$

где k_0 – количество виртуальных подсетей в эталонной СРД-VLAN; k – количество виртуальных подсетей в результирующей СРД-VLAN.

Экспериментальная оценка ГА, предложенного для решения задачи проектирования VLAN с использованием ролевого подхода, осуществлялась в двух режимах: для конкретного фрагмента ЛВС критической инфраструктуры и для сгенерированных с помощью инструментального стенда примеров. Во всех режимах использовались следующие значения основных параметров ГА: $N_{\text{pop}} = 200$; $W_{\text{cross}} = 0{,}1$; $W_{\text{mut}} = 0{,}01$; $W_{\text{gen}} = W_{\text{el}} = 0{,}5$. Эти значения были выбраны на основании эмпирических рекомендаций, взятых из различных источников по разработке ГА.

Фрагмент ЛВС содержал 44 рабочие станции, сгруппированные в три группы. Маршрутизаторы групп были связаны с маршрутизатором серверного кластера, состоящего из четырех серверов, предоставляющих различные информационные услуги (Video, mail, FTP, VoIP, GIS, приложения). В ходе проектирования VLAN для этого фрагмента получены следующие результаты. Традиционный ГА дал оптимальное решение, содержащее 22 виртуальные подсети. При ролевом подходе достигнуто оптимальное решение, содержащее меньшее количество подсетей – 20. При этом было сформировано 29 ролей. Таким образом, выигрыш в точности проектирования VLAN при ролевом ГА составил 9 процентов. Кроме того, на поиск решения традиционным ГА потребовалось 650 итераций, а ролевым – 445. На компьютере с процессором Intel i7-8550U 1.8GHz, 16RAM длительность одной итерации в традиционном ГА составляла в среднем 225 мс, а в ролевом – 105 мс. Сокращение длительности итерации обусловлено использованием в ролевом ГА не одной большой, а двух коротких хромосом. В результате среднее время проектирования VLAN традиционным ГА было равно 146,25 с, а ролевым – 46,725 с, или в 3,1 раза быстрее.

В режиме оценки на сгенерированных примерах размерность n принимала следующие значения: 100, 200, 500 и 1000. Значение k находилось в диапазоне от 20 до 50 процентов от значения n. Эксперименты проводились следующим образом. Вначале выбирался размер сети n. Затем формировалась эталонная РСРД-VLAN, определяемая матрицами $\mathbf{X}_0$ и $\mathbf{Y}_0$, и требуемая матрица логической связности компьютеров $\mathbf{A}$. Далее с помощью ГА находились результирующие матрицы $\mathbf{X}$ и $\mathbf{Y}$. Путем сравнения результирующих матриц с эталонными рассчитывался показатель точности согласно (9). Количество итераций T, затраченных на поиск матриц $\mathbf{X}$ и $\mathbf{Y}$, являлось показателем скорости работы ГА. Параллельно решалась традиционная задача оптимизации СРД-VLAN. В ней использовалась та же матрица $\mathbf{A}$. Для ГА поиска матрицы $\mathbf{X}$ в этой задаче также рассчитывались точность и скорость работы. Результаты экспериментов приведены в таблице.

В таблице используются следующие обозначения: T_0 – скорость работы традиционного ГА; T – скорость работы ролевого ГА; δ – выигрыш в скорости работы ролевого ГА, который определяется по формуле $\delta = T_0/T$; θ – точность ролевого ГА. Случайная выборка формировалась по 10 испытаниям.

Анализируя данные, представленные в таблице, можно сделать следующие выводы. Прежде всего, во всех экспериментах зафиксирован выигрыш ролевого ГА по скорости работы. Это объясняется использованием в ролевом ГА двух более коротких хромосом. Кроме того, ролевой ГА показал достаточно высокую точность (от 0,82 до 0,92) поиска оптимального решения. Однако если размерность задачи не очень высокая (не более

Таблица

Экспериментальные результаты

n	k	T_0	T	δ	θ
100	20	1500	1050	1,42	0,89
100	30	1250	930	1,34	0,90
100	50	1100	850	1,29	0,92
200	40	1800	1100	1,63	0,85
200	60	1500	1030	1,46	0,88
200	100	1350	970	1,39	0,90
500	100	3500	1350	2,59	0,84
500	200	2500	1230	2,03	0,87
500	250	2000	1150	1,74	0,89
1000	200	9400	1620	5,80	0,82
1000	300	6500	1550	4,19	0,84
1000	500	5000	1450	3,44	0,85

100 рабочих станций в сети), то разница между применением традиционного и ролевого ГА небольшая. Можно использовать традиционный ГА. Предлагаемый ролевой ГА обладает более высокой эффективностью при более высоких размерностях задачи.

Заключение

Предложенный ролевой подход к решению задачи проектирования VLAN в больших критических инфраструктурах заключается в оптимизации РСРД-VLAN. Сформулированная постановка задачи ролевого проектирования VLAN является особой разновидностью БМФ, в которой исходная матрица декомпозируется на две пары прямых и транспонированных более коротких булевых матриц. Эта задача является более сложной, чем известные задачи БМФ, и использование для ее решения известных математических методов является неэффективным.

Предложено использовать для решения этой задачи усовершенствованный ГА. Основными усовершенствованиями ГА являются: кодирование решений двумя хромосомами; учет в функции пригодности критериев, отвечающих за минимизацию количества виртуальных подсетей и ролей; реализация мультихромосомного скрещивания особей. Экспериментальная оценка предложенного ролевого ГА показала его достаточно высокую эффективность. По скорости работы он дает выигрыш от 1,5 до 5 раз. При этом обеспечивается достаточно высокая точность решения задачи.

Дальнейшие исследования ориентированы на применения ролевого ГА к другим областям разграничения доступа.

СПИСОК ЛИТЕРАТУРЫ

1. Концепция автоматизации управления информационной безопасностью в защищенных мультисервисных сетях специального назначения / С.А. Агеев, А.С. Бушуев, Ю.П. Егоров, И.Б. Саенко // Автоматизация процессов управления. 2011. № 1 (23). С. 50–57.

2. Design and implementation of a VLAN / M.E. Gomez-Romero, M. Reyes-Ayala, E.A. Andrade-González, J.A. Tirado-Mendez // Proceedings of the 2010 international conference on Applied computing conference (ACC'10). 2010. pp. 87–90.

3. Saenko I., Kotenko I. Design and Performance Evaluation of Improved Genetic Algorithm for Role Mining Problem // Proceedings of the 2012 20th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP '12). 2012. pp. 269–274.

4. Kotenko I., Saenko I. Improved genetic algorithms for solving the optimisation tasks for design of access control schemes in computer networks // Int. J. Bio-Inspired Comput. 2015. Vol. 7, no. 2. pp. 98–110.

5. Saenko I., Kotenko I. A genetic approach for virtual computer network design // Proceedings of the 8th International Symposium on Intelligent Distributed Computing (IDC'14). Intelligent Distributed Computing VIII. Studies in Computational Intelligence. 2015. Vol. 570. pp. 95–105.

6. Саенко И.Б., Котенко И.В. Генетическая оптимизация и визуальный анализ при формировании схем доступа в ВЛВС // Информационные технологии и вычислительные системы. 2015. № 1. С. 33–46.

7. Çergani E., Miettinen P. Discovering Relations using Matrix Factorization Methods // Proceedings of the 22nd ACM International Conference on Information & Knowledge Management (CIKM'13). 2013. pp. 1549–1552.

8. Miettinen P., Vreeken J. MDL4BMF: Minimum Description Length for Boolean Matrix Factorization // ACM Trans. Knowl. Discov. Data. 2014. Vol. 8, no. 4, article 18. pp. 1–31.

9. Janecek A., Tan Y. Using population based algorithms for initializing nonnegative matrix factorization // Proceedings of the Second international conference on Advances in swarm intelligence. Volume Part II (ICSI'11). LNCS. 2011. Vol. 6729. pp. 307–316.

10. Investigating Boolean Matrix Factorization / V. Snasel, J. Platos, P. Krömer, D. Husek, R. Neruda, A. Frolov // Proceedings of the Workshop on Data Mining using Matrices and Tensors (DMMT'08). 2008. pp. 1–8.

11. Snasel V., Platos J., Krömer P. On Genetic Algorithms for Boolean Matrix Factorization // Eighth Intern. Conference on Intelligent Systems Design and Applications (ISDA 2008). 2008. Vol. 2. pp. 170–175.

12. Rezaei M., Boostani R. Using the genetic algorithm to enhance nonnegative matrix factorization initialization // Expert Sys: J. Knowl. Eng. 2014. Vol. 31, no. 3. pp. 213–219.

13. Tongkaw S., Tongkaw A. Multi-Vlan Design Over IPSec VPN for Campus Network // Proceedings of the 2018 IEEE Conference on Wireless Sensors (ICWiSe). 2018. pp. 66–71.

14. Tai Ch.-F., Chiang T.-Ch., Hou T.-W. A virtual subnet scheme on clustering algorithms for mobile ad hoc networks // Expert Syst. Appl. 2011. Vol. 38, no. 3. pp. 2099–2109.

15. Saenko I., Kotenko I. Genetic optimization of access control schemes in virtual local area networks // Proceedings of the 5th international conference on Mathematical methods, models and architectures for computer network security (MMM-ACNS'10). LNCS. 2010. Vol. 1494. pp. 209–216.

16. GreenVLAN: An energy-efficient approach for VLAN design / K. He, Y. Wang, X. Wang, W. Meng, B. Liu // Proceedings of the 2012 International Conference on Computing, Networking and Communications (ICNC). 2012. pp. 522–526.

17. A systematic approach for evolving VLAN designs / X. Sun, Y.-W.E. Sung, S.D. Krothapalli, S.G. Rao // Proceedings of the 29th conference on Information communications (INFOCOM'10). 2010. pp. 1451–1459.

18. Zhu M., Molle M., Brahmam B. Design and Implementation of Application-Based Secure VLAN // Proceedings of the 29th Annual IEEE International Conference on Local Computer Networks (LCN '04). 2004. pp. 407–408.

19. Lu H., Vaidya J., Atluri V. Optimal Boolean Matrix Decomposition: Application to Role Engineering // Proceedings of the 2008 IEEE 24th International Conference on Data Engineering (ICDE '08). 2008. pp. 297–306.

REFERENCES

1. Ageev S.A., Bushuev A.S., Egorov Iu.P., Saenko I.B. Kontseptsiia avtomatizatsii upravleniia informatsionnoi bezopasnostiu v zashchishchennykh multiservisnykh setiakh spetsialnogo naznacheniia [The Concept of Automation of Information Security Control in Protected Special-Purpose Multi-Service Networks]. *Avtomatizatsiia protsessov upravleniia* [Automation of Control Processes], 2011, no. 1 (23), pp. 50–57.

2. Gomez-Romero M.E., M. Reyes-Ayala, E.A. Andrade-González, J.A. Tirado-Mendez. Design and Implementation of a VLAN. *Proceedings of the International Conference on Applied Computing Conference (ACC'10)*. 2010, pp. 87–90.

3. Saenko I., Kotenko I. Design and Performance Evaluation of Improved Genetic Algorithm for Role Mining Problem. *Proceedings of the 20th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP '12)*. 2012, pp. 269–274.

4. Kotenko I., Saenko I. Improved Genetic algorithms for Solving the Optimisation Tasks for Design of Access Control Schemes in Computer Networks. *Int. J. Bio-Inspired Comput.*, 2015, vol. 7, no. 2, pp. 98–110.

5. Saenko I., Kotenko I. A Genetic Approach for Virtual Computer Network Design. *Proceedings of the 8th International Symposium on Intelligent Distributed Computing (IDC'14). Intelligent Distributed Computing VIII. Studies in Computational Intelligence*. 2015, vol. 570, pp. 95–105.

6. Saenko I.B., Kotenko I.V. Geneticheskaia optimizatsiia i vizualinyi analiz pri formirovanii skhem dostupa v VLVS [Genetic Optimization and Visual Analysis in the Formation of Access Schemes in the VLAN]. *Informatsionnye tekhnologii i vychislitelnye sistemy* [Information Technologies and Computing Systems], 2015, no. 1, pp. 33–46.

7. Çergani E., Miettinen P. Discovering Relations using Matrix Factorization Methods. *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management (CIKM'13)*. 2013, pp. 1549–1552.

8. Miettinen P., Vreeken J. MDL4BMF: Minimum Description Length for Boolean Matrix Factorization. *ACM Trans. Knowl. Discov. Data*, 2014, vol. 8, no. 4, article 18, pp. 1–31.

9. Janecek A., Tan Y. Using Population Based Algorithms for Initializing Nonnegative Matrix Factorization. *Proceedings of the Second International Conference on Advances in Swarm Intelligence. Volume Part II (ICSI'11). LNCS*. 2011, vol. 6729, pp. 307–316.

10. Snasel V., J. Platos, P. Krömer, D. Husek, R. Neruda, A. Frolov. Investigating Boolean Matrix Factorization. *Proceedings of the Workshop on Data Mining using Matrices and Tensors (DMMT'08)*. 2008, pp. 1–8.

11. Snasel V., Platos J., Krömer P. On Genetic Algorithms for Boolean Matrix Factorization. *Eighth Intern. Conference on Intelligent Systems Design and Applications (ISDA 2008)*. 2008, vol. 2, pp. 170–175.

12. Rezaei M., Boostani R. Using the Genetic Algorithm to Enhance Nonnegative Matrix Factorization Initialization. *Expert Sys: J. Knowl. Eng.*, 2014, vol. 31, no. 3, pp. 213–219.

13. Tongkaw S., Tongkaw A. Multi-Vlan Design Over IPSec VPN for Campus Network. *Proceedings of the 2018 IEEE Conference on Wireless Sensors (ICWiSe)*. 2018, pp. 66–71.

14. Tai Ch.-F., Chiang T.-Ch., Hou T.-W. A Virtual Subnet Scheme on Clustering Algorithms for Mobile Ad Hoc Networks. *Expert Syst. Appl.*, 2011, vol. 38, no. 3, pp. 2099–2109.

15. Saenko I., Kotenko I. Genetic Optimization of Access Control Schemes in Virtual Local Area Networks. *Proceedings of the 5th International Conference on Mathematical Methods, Models and Architectures for Computer Network Security (MMM-ACNS'10). LNCS.* 2010, vol. 1494, pp. 209–216.

16. He, K., Y. Wang, X. Wang, W. Meng, B. Liu GreenVLAN: An Energy-Efficient Approach for VLAN Design. *Proceedings of the 2012 International Conference on Computing, Networking and Communications (ICNC).* 2012, pp. 522–526.

17. Sun, X., Y.-W.E. Sung, S.D. Krothapalli, S.G. Rao. A Systematic Approach for Evolving VLAN Designs. *Proceedings of the 29th Conference on Information Communications (INFOCOM'10).* 2010, pp. 1451–1459.

18. Zhu M., Molle M., Brahmam B. Design and Implementation of Application-Based Secure VLAN. *Proceedings of the 29th Annual IEEE International Conference on Local Computer Networks (LCN '04).* 2004, pp. 407–408.

19. Lu H., Vaidya J., Atluri V. Optimal Boolean Matrix Decomposition: Application to Role Engineering. *Proceedings of the IEEE 24th International Conference on Data Engineering (ICDE '08).* 2008, pp. 297–306.