

УДК 004.8(075.8)

ГЕНЕТИЧЕСКАЯ ОПТИМИЗАЦИЯ ПРИ АВТОМАТИЗИРОВАННОМ ПРОЕКТИРОВАНИИ ВЫЧИСЛИТЕЛЬНЫХ СЕТЕЙ

Н.Г. Ярушкина¹, А.А. Стецко²

В статье рассмотрены возможности генетической оптимизации для различных этапов проектирования вычислительных сетей (ВС). Для основных задач проектирования: поиска топологии ВС, выбора активного коммуникационного оборудования, размещения коммутационных узлов, размещения сервисов по серверам ВС описываются схемы генетической оптимизации. Для каждой задачи сформулированы особенности формирования исходных популяций, кодирования хромосом и целевых функций. Каждая целевая функция выражается алгоритмически, как процедура моделирования трафика ВС. На основе разработанного подхода к генетической оптимизации создан программный продукт – САПР ВС. В статье приведены конкретные результаты генетической оптимизации на основе моделирования трафика ВС.

Введение

Анализ современного состояния автоматизированного проектирования ВС позволил сформулировать базовые задачи данного исследования [Ярушкина и др., 2007]:

1. Необходимо построить математическую модель автоматизированного проектирования ВС, адаптировать генетические алгоритмы к задаче оптимизации ресурсов в вычислительных сетях, то есть предложить способы кодирования структуры сети, определения трафика и целевой функции. Необходимо исследовать разнообразные типы генетических алгоритмов с точки зрения применимости к ВС.
2. Необходимо разработать модель трафика ВС, позволяющую использовать при моделировании, как качественные оценки, так и результаты статистических измерений; необходимо описать

¹ 432027, Ульяновск, ул. Северный Венец, 32, УлГТУ, jng@ulstu.ru

² 432022, Ульяновск, ул. Солнечная 20, ФНПЦ ОАО «НПО МАРС», mars@mv.ru

процедуры моделирования трафика, как алгоритмы вычисления целевой функции генетического алгоритма.

1. Кодирование варианта топологии ВС как хромосомы при генетической оптимизации

Рассмотрим модель корпоративной сети [Малашенко и др., 1998], которая состоит из узлов маршрутизации R и каналов связи между ними K . Пропускная способность i -го канала равна C_i (бит в секунду). Предположим, что каналы связи K являются абсолютно помехозащищенными, а узлы R абсолютно надежны и выполняют операции по маршрутизации данных, включая распаковку данных, выбор маршрутов, хранение в буфере, уведомление. Считаем, что времена обработки в узлах равны и являются постоянными. Кроме того, в модели имеются очереди к каналам и задержки при передаче. Трафик, поступающий в узел из внешних источников, образует пуассоновский процесс со средним значением g_{jk} сообщений в секунду, для тех сообщений, которые возникают в узле j и предназначаются для узла k . Полный внешний трафик, поступающий в узел и покидающий его для сети с N узлов и M каналов можно описать как

$$g = \sum_{j=1}^N \sum_{k=1}^N g_{jk} .$$

При этом величина g_i имеет вид

$$g_i = \sum_k g_{ik} .$$

Каждый канал в сети рассматривается как отдельное обслуживающее устройство. Обозначим через I_i среднее число сообщений в секунду, проходящих по i -му каналу. Как и для внешнего трафика узла, определим полный трафик в сети следующим образом:

$$I = \sum_{i=1}^M I_i .$$

Решения для генетического алгоритма (ГА) кодируются в виде хромосомы. В данной работе предлагается хромосому представлять в виде вектора целых чисел. Каждый ген (позиция вектора) представляет собой один узел, поэтому размер хромосомы соответствует числу узлов. Так как

узлы в таблице статистики имеют произвольные идентификаторы, строится таблица соотношения идентификатора узла с порядковым номером гена. По этой таблице определяется соответствие элемента вектора конкретному узлу.

В соответствие каждому концентратору или коммутатору также ставится целое число, соответствующее номеру записи в таблице коммуникационных устройств. Значение, которое хранится в позиции вектора, представляет собой номер записи в таблице концентраторов или коммутаторов, содержащей информацию об устройстве, к которому подключен соответствующий позиции вектора узел, то есть порядковый номер устройства на главном канале (бэббоне).

Общая схема ГА может уточняться при использовании эволюционных алгоритмов разных типов. Для реализации необходимо уточнять следующие параметры: способ кодировки решения (хромосомы), целевую функцию (оценки) каждой хромосомы, содержание операторов отбора (селекции), рекомбинации и мутации, вероятностные параметры управления сходимостью эволюции, условие завершения эволюции.

2. Алгоритм моделирования трафика ВС – базовая процедура вычисления целевой функции ГА

Модель трафика ВС на транспортном уровне представляется на основе дискретной нечеткой случайной величины (НСВ) [Ярушкина, 2004], [Батыршин и др., 2007], которая имеет вид:

$$T_r = \{A_1 / P_1, A_2 / P_2, \dots, A_n / P_n\},$$

где A_1, A_2, \dots, A_n – нечеткие значения, которые величина T_r принимает с вероятностями P_1, P_2, \dots, P_n . Значения НСВ являются взаимоисключающими, поэтому сумма вероятностей этих значений должна быть равна 1, то есть должно выполняться условие нормирования НСВ: $P_1 + P_2 + \dots + P_n = 1$.

Поскольку при проектировании ВС специалист выдает прогноз словом, набор вероятностей P_1, P_2, \dots, P_n можно также перевести в лингвистическую форму. Для этого вводится понятие «степень уверенности НСВ». Каждый вектор вероятностей кодирует степень уверенности НСВ, которая представляет собой лингвистическую оценку вида «точно», «скорее всего» («скорее»), «наверное» («возможно»). Каждую такую оценку можно представить в виде функции распределения НСВ.

Каждый процесс в функциональной модели рассматривается как генератор сетевого трафика с величиной, заданной словесной оценкой. На рассматриваемом интервале времени значение трафика распределяется в соответствии с мерой возможности его оценки. Например, оценка «скорее

высокий» расшифровывается как НСВ {«низкий»/0, «средний»/0.25, «высокий»/0.75}. На рис. 1 показано распределение значения трафика в интервале времени.

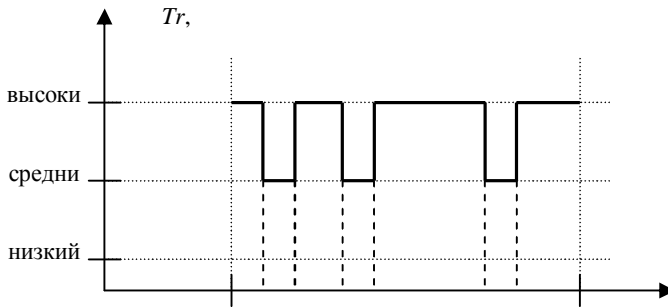


Рис. 1. Распределение значения трафика по времени

Критерием оценки качества хромосомы является значение трафика на бэббоне при таком подключении узлов, которое закодировано в данной хромосоме. Более приспособленной считается хромосома, значение целевой функции для которой минимально.

Основным источником сетевого трафика являются процессы, происходящие на узле сети. При составлении функциональной модели сети следует обозначить все эти процессы и потоки данных, генерируемые ими. Кроме того, каждый из процессов имеет свое расписание выполнения в течение дня, и его необходимо учитывать при моделировании.

В связи с этим к правилам языка Data Flow Diagram (DFD), описывающим сущность «процесс», были сделаны дополнения для более корректного описания процессов, происходящих в ВС. В свойства сущности «процесс» помимо названия и номера включены еще нечеткие прогнозные оценки трафика и вычислительной загрузки, а также расписание выполнения процесса (рис. 2).

При моделировании работы системы расписание имеет решающее значение, так как в определенные моменты времени процесс не функционирует и генерации трафика в сеть не происходит. Кроме того, для дальнейшей взаимосвязи с физической структурой сети в свойства процесса добавлен IP-адрес того компьютера, на котором этот процесс выполняется.

Таким образом, целевая функция при генетической оптимизации для различных задач автоматизированного проектирования ВС представляет собой алгоритм имитационного моделирования трафика на каждом канале ВС.

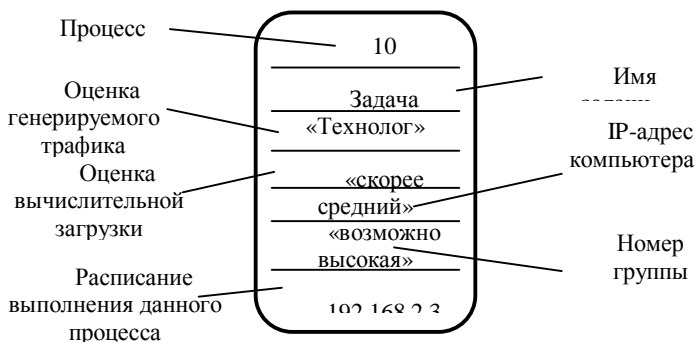


Рис. 2. Изображение процесса на потоковой диаграмме

3. САПР вычислительных сетей

В данной секции описана реализация САПР ВС [Стецко, 2007], функционально состоящая из следующих модулей: DFD-редактор, мастер перехода, редактор ВС, модуль имитационного моделирования на основе сетей Петри, модуль моделирования роутинга, модуль генетической оптимизации и модуль байесовской оптимизации. Связи между модулями изображены на рис. 3.

4. Анализ вычислительного эксперимента моделирования ВС

Рассмотрим результаты вычислительного эксперимента по размещению сервисов по топологической схеме [Стецко и др., 2007]. В сеть рассматриваемого предприятия объединены отделы, связанные, прежде всего, с разработкой программного обеспечения, конструкторские, технологические отделы, отделы программистов и разработчиков, а также менеджеры по управлению этими подразделениями.

Основой для ВС на глобальном уровне служит движение электронной инженерно-технической документации. На уровне подразделений сеть служит для распределенного ведения комплексных проектов различного оборудования, а также проектов программного обеспечения. Кроме этого,

на уровне подразделений сеть позволяет вести централизованную распечатку документов.



Рис. 3. Архитектура САПР ВС

На рис. 4 показан сегмент диаграммы, описанный на дополненном языке DFD, отражающий процессы одного научно-исследовательского отдела (НИО).

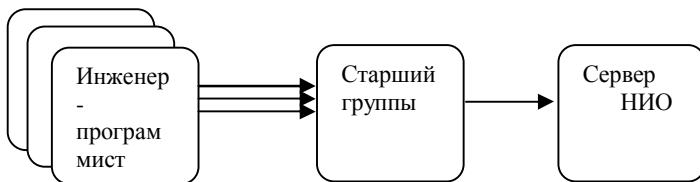


Рис. 4. Прикладные процессы одного НИО

В магистральной сети содержатся все серверы подразделений. В большинстве подсетей для коммуникации использованы коммутаторы. В

подсетях крупных отделов применено одноуровневое каскадирование коммутаторов.

В целях изучения эффективности процесса генетической оптимизации описанная структура прикладных процессов была сопоставлена со структурой действующей локальной сети. Для «исторически» сложившегося варианта размещения прикладных процессов на предприятии было рассчитано значение целевой функции, которое рассматривалось как база сравнения в вычислительных экспериментах. Затем неоднократно выполнялась оптимизация заложенного в систему проекта сети. Эксперимент проводился несколько раз с различными параметрами канонического ГА. Изменялся размер популяции, уровень сходимости и вероятность мутации.

Графики сходимости стандартного ГА для каждого набора параметров приведены на рис. 5. Некоторые данные о значениях целевой функции в процессе эволюционного поиска приведены в табл. 1. Значение целевой функции для существующего варианта расположения прикладных процессов по сети – 0,47 единиц, на рис. 5 данное значение отмечено пунктирной линией. Из графика видно, что генетическая оптимизация позволяет найти более эффективное решение со значением целевой функции 0,45 единиц. Эксперименты по эволюционному поиску проводились на компьютере класса Pentium 4, с частотой процессора 2200 МГц и объемом оперативной памяти 256 Мбайт. Длительность расчетов приведена в табл. 1.

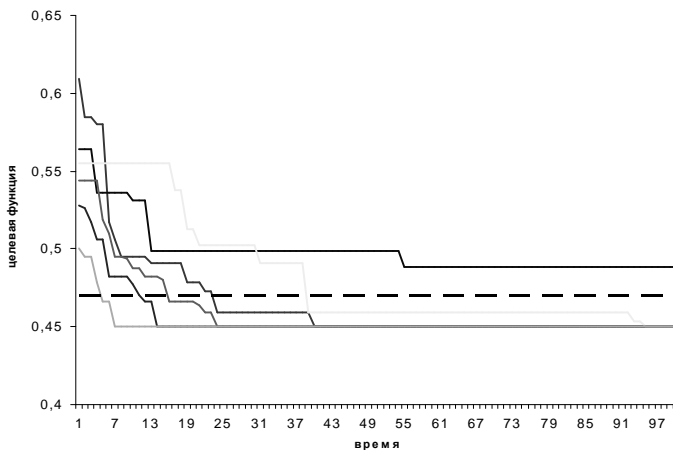


Рис. 5. Сходимость ГА в вычислительных экспериментах

Табл. 1

Эволюционное время	P=100, Элита=10, Мутация 0,9	P=500, Элита=100, Мутация 0,4	P=5000, Элита =2000, Мутация 0,4	P=100, Элита=10, Мутация 0,4	P=10000, Элита=3000, Мутация 0,4
1	0,609375	0,5640625	0,54375	0,555	0,528125
2 ...	0,584375	0,5640625	0,54375	0,555	0,5265625
99	0,45	0,48875	0,45	0,45	0,45
100	0,45	0,48875	0,45	0,45	0,45

5. Опыт практического использования генетической оптимизации

Для предприятия ОАО «Ульяновское конструкторское бюро приборостроения» при помощи предлагаемой системы было произведено проектирование новой корпоративной сети на 100 рабочих станций. Проведем оценочный расчет количества времени, требуемого на обработку вариантов распределения бизнес-процессов ОАО «УКБП» по транспортной схеме сети. Согласно этому расчету на обработку одного варианта такого распределения средний «специалист-сетевик» затратит около 8 часов. В результате отработки ГА с параметрами 5000/2000/0,4 за 100 циклов алгоритма обрабатывается $5000 \cdot 100 = 500000$ вариантов. На это современный компьютер затрачивает 2 ч 20 мин. В случае выполнения проектирования сети вручную при обработке 500000 вариантов распределения человек затратит $500000 \cdot 8 = 4000000$ часов, т.е. около полумиллиона рабочих дней. В реальности человек обрабатывает 2-3 варианта и выбирает один из них. Очевидно, что качество полученного таким образом результата будет существенно ниже, чем при использовании автоматизированного проектирования.

Заключение

Таким образом, разработан подход к решению оптимизационных задач при автоматизированном проектировании на основе схемы генетической оптимизации. Разработана система проектирования корпоративных сетей с учетом моделирования трафика корпоративных сетей с использованием четких и нечетких параметров. Проведены вычислительные эксперименты по исследованию эффективности разработанной САПР ВС. Разработанная на основе данного подхода САПР ВС была отмечена бронзовой медалью на Международном салоне изобретений в Женеве в 2006 г.

Список литературы

[Ярушкина и др., 2007] Ярушкина Н.Г., Макеев А.С., Стецко А.А., Тронин В.Г. Автоматизированное проектирование корпоративных сетей на основе нечетких

гиперграфов. В кн. тр. Международной конференции «Интеллектуальные системы» AIS'07, М.: Физматлит, 2007.

[Ярушкина, 2004] Ярушкина Н.Г. Основы теории нечетких и гибридных систем: – М.: Финансы и статистика, 2004. – 320 с.

[Батыршин и др., 2007] Батыршин И.З., Недосекин А.О., Стецко А.А, Тарасов В.Б., Язенин А.В., Ярушкина Н.Г. Нечеткие гибридные системы. Теория и практика/Под ред. Н.Г. Ярушкиной.- М.:Физматлит, 2007.

[Стецко и др., 2007] Стецко А.А., Ярушкина Н.Г. Автоматизированное проектирование вычислительной сети крупной проектной организации // Программные продукты и системы, № 4 (80), 2007

[Стецко, 2007] Стецко А.А. Архитектура программной системы автоматизированного проектирования и моделирования вычислительных сетей // Программные продукты и системы, № 4 (80), 2007

[Малашенко и др., 1998] Малашенко Ю.Е., Новикова Н.М., Смирнов М.М. Анализ многопользовательских сетевых систем с учетом неопределенности. Многокритериальная и параметрическая постановка для неизвестных требований . // Известия РАН. Теория и системы управления, N 3, 1998.