

УДК 004.413

А.И. ШЕВЧЕНКО, А.И. ОЛЬШЕВСКИЙ

**МНОГОУРОВНЕВАЯ СИСТЕМА ПОСТРОЕНИЯ МАРШРУТА В СЕТИ  
ДИСТАНЦИОННОГО ОБУЧЕНИЯ**

**Abstract.** Methods of synthesis networks, ways and algorithms applied for group dispatch of data have been considered at remote training. Combination of two types of approaches to the decision of a problem is offered: division into subnetworks supporting on the regional centers of trees-synthesis as D-structures and methods of synthesis of networks routing in a class of treelike structures. Thus, the developed system is two-level and allows to apply heuristics.

**Key words:** multilevel system, remote training, treelike structures.

**Анотація.** Розглянуті методи синтезу мереж і застосовувані при цьому способи та алгоритми для групового розсилання даних при дистанційному навчанні. Пропонується сполучити два типи підходів до рішення завдання: поділ на мережі з опорою на регіональні центри синтезу дерев у вигляді D-структур і методи синтезу маршруту мереж у класі деревоподібних структур. Таким чином, розроблена система є дворівневою й дозволяє застосовувати евристичні методи.

**Ключові слова:** багаторівнева система, дистанційне навчання, деревоподібні структури.

**Аннотация.** Рассмотрены методы синтеза сетей и применяемые при этом способы и алгоритмы для групповой рассылки данных при дистанционном обучении. Предлагается совместить два типа подходов к решению задачи: разделение на подсети с опорой на региональные центры синтеза деревьев в виде D-структур и методы синтеза маршрута сетей в классе древовидных структур. Таким образом, разработанная система является двухуровневой и позволяет применять эвристические методы.

**Ключевые слова:** многоуровневая система, дистанционное обучение, древовидные структуры.

**1. Введение**

В настоящее время наиболее эффективным и перспективным методом обучения является дистанционное обучение. Широкое распространение персональных компьютеров и активное развитие глобальных сетей вывело этот процесс на принципиально новый уровень. Теперь получить образование можно независимо от места жительства и физических возможностей. Дистанционное образование позволяет получить диплом любого вуза, любой специальности.

Актуальной проблемой дистанционного обучения стал поиск оптимального маршрута пересылки данных, так как для качественного дистанционного обучения вуз должен тесно сотрудничать со студентами. При этом немаловажным фактором проектирования маршрута сети является выбор показателя эффективности (критерия оптимизации). В связи с этим одним из вариантов постановки задачи формализованного выбора есть синтез сети по критерию стоимости – при заданных задержках в передаче данных, показателях надежности сети и заданных объемах информации, подлежащих обработке в сети, сформировать маршрут сети, который будет иметь минимальную стоимость пересылки данных. Разрабатываются более дешевые и быстрые пути передачи информации.

**2. Особенности подхода построения сети обучения**

Глобальные сети не являются стабильными, т.е. изменяются их структура, количество участников, стоимость услуг и технико-экономические характеристики технических средств и т.п. В то же время система дистанционного образования предъявляет свои требования к соответствующим системам.

Эти требования должны учитывать особенности развития сети дистанционного обучения. При этом необходимо учитывать рост количества студентов, их активность относительно участия в рассылке и объемы передаваемых данных, изменения методик обучения и другие аспекты. Поэтому разработать оптимальный маршрут невозможно. Для постоянного пересчета путей разрабатываются самые различные алгоритмы их построения, модифицируются известные алгоритмы.

Практически решение такой задачи зависит от числа узлов сети, и для размерностей порядка  $10^3$ – $10^4$  обычные традиционные методы синтеза структуры не пригодны. Возникает проблема разработки эффективных приближенных методов структурного синтеза сетей большой размерности. Один из подходов – это использование иерархии. Для дистанционного обучения – это использование региональных учебных филиалов.

Построение сети становится не приемлемым при большом количестве пользователей и если пользователи расположены на обширной территории, то есть сеть покрывает различные регионы. В данном случае появляется ряд существенных проблем: постоянная большая загруженность сервера, из-за чего снижается быстродействие сети, повышается вероятность возникновения сбоев и потери информации, что не допустимо. В таком случае более удобным и правильным, с точки зрения предъявляемых к сети требований, вариантом будет установка региональных серверов, где размещаются методические материалы и задания по дисциплинам. Такая структура снижает загруженность главного сервера, что значительно повышает быстродействие сети и ее надежность. При введении региональных серверов существенно снизилась вероятность сбоев, т.е. повысилась производительность сети, уменьшилось время реакции [1].

Одной из особенностей подхода к решению задачи является двухэтапное построение маршрута. В этом случае сначала используются методы синтеза сетей, а потом происходит построение маршрута рассылки к необходимым узлам.

Разрабатываемый алгоритм делится на две части. На первом этапе сеть разбивается на подсети по числу региональных учебных филиалов. Это позволит снизить нагрузку по рассылке с центрального учебного центра. Каждая подсеть объединяется вокруг своего регионального центра. Этот процесс сродни задаче разделения объектов на классы. Пользователь приписывается к определенному центру на основании его географической близости к нему. Если возникает спорная ситуация, т.е. абонент находится ровно посередине между двумя или более центрами, то решение о принадлежности принимается на основании стоимости его связи с данными центрами. В данной работе на этом этапе используется метод синтеза на основе D-структур [2]. Этот метод позволяет осуществлять привязку и к самому вузу, если это экономически выгодно.

Второй этап алгоритма – это построение деревьев внутри каждой подсети. При этом рассматривается задача Штейнера: нужно найти такое дерево сети, чтобы полная стоимость ребер дерева была минимальна [3]. В стоимость обычно включается время передачи единицы данных по каналу, расстояние или денежный эквивалент данного соединения, пропускная способность канала или комбинация этих и других критериев. На этом этапе используются эвристические методы построения деревьев, в частности, генетические алгоритмы для решения задачи Штейнера. В

данном случае в качестве точек Штейнера выступают промежуточные узлы – центры передачи данных.

### 3. Постановка задачи

Для дальнейшего рассмотрения методов и их комбинации следует сформулировать математическую постановку задачи синтеза сети и ввести соответствующие условные обозначения.

Как уже говорилось ранее, сеть состоит из вуза, региональных центров (РЦ), промежуточных узлов и абонентов или абонентских пунктов (АП). Число АП в сети в дальнейшем будет обозначаться  $N$ , количество РЦ вуза –  $M$ . Затраты на передачу данных будем обозначать  $W$  (в общем случае это суммарные затраты, иногда могут указываться индексы для описания передачи между определенными узлами). Для оценки алгоритмов используется величина  $N$  – пропускная способность, а также высота дерева (в некоторых случаях могут отсутствовать промежуточные узлы или наоборот – они могут иметь более сложную иерархию).

В итоге задачу можно сформулировать следующим образом: необходимо построить сеть с большим (и постоянно изменяющимся) количеством абонентов  $N$  и ограниченным числом региональных центров  $M$  так, чтобы общие затраты на передачу данных  $W$  были минимальны.

### 4. Алгоритм привязки к региональным центрам

Этот этап построения маршрута служит для разделения задачи на несколько частей, которые могут выполняться параллельно. Привязка к региональным центрам является предварительной сортировкой узлов и связей, которые между ними существуют.

В самом начале загружаются данные о сети. На основе этих данных (а точнее, на основе информации о связях) производится сортировка связей по возрастанию стоимости передачи данных.

Далее связи рассматриваются именно в порядке этой сортировки, т.е. от самой дешевой связи к самой дорогой. В данной задаче не рассматривается направление связи, поэтому концы имеют одинаковое значение.

В процессе работы алгоритма для каждого конечного абонента устанавливается привязка в определенному региональному центру. Так как за основу взят алгоритм D-структур, то привязка может осуществляться и к главному центру, если это экономически выгодно [2].

Если один из концов связи является региональным центром, то привязка осуществляется при условии «свободного» второго конца связи. Так как связи рассматриваются от дешевой к дорогой, то предполагается следующее: если абонент уже привязан к РЦ, значит, эта связь была дешевле, чем текущая, если он еще не привязан, то следующая возможная связь не будет более дешевой.

Если же рассматривается связь между абонентом и промежуточным центром, то необходимо проверить (при необходимости установить) его привязку к РЦ, затем найти ближайшую связь абонента с РЦ и выбрать более выгодное решение.

Алгоритм действует до тех пор, пока все абоненты не получат привязку к какому-либо узлу: РЦ или промежуточному. Не обязательно все промежуточные узлы будут иметь привязку, только

те, которые действительно нужны для связи абонентов с РЦ. Таким образом, сеть будет связной, но не избыточной. Алгоритм представлен на рис. 1.

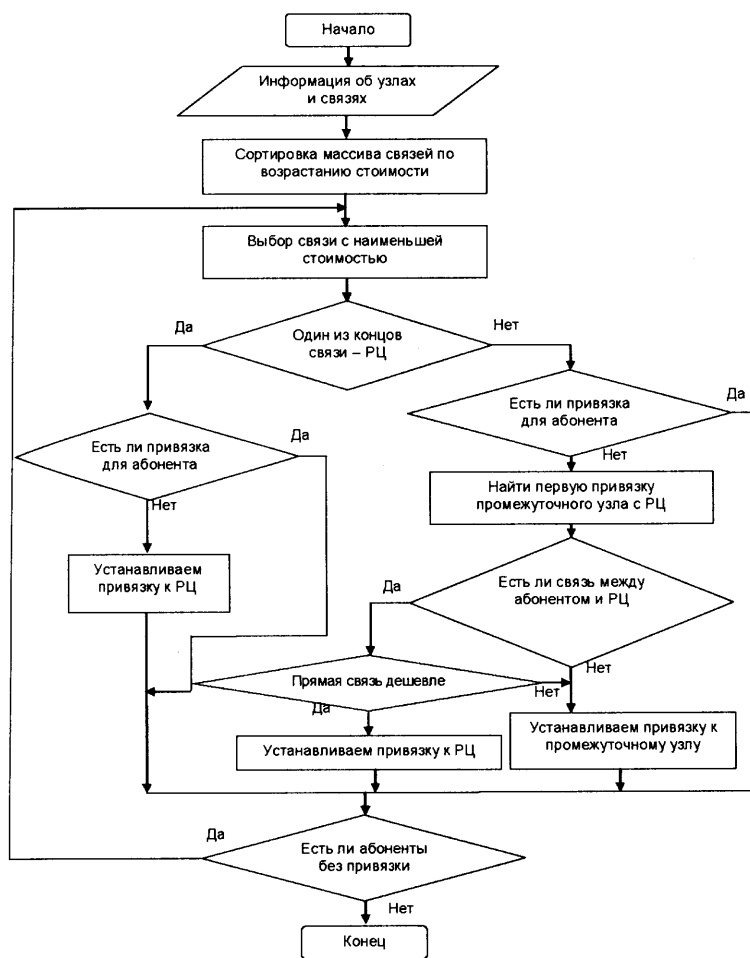


Рис. 1. Алгоритм привязки

## 5. Алгоритм построения маршрута

На втором этапе построения маршрута используются генетические алгоритмы [4]. Чтобы подробнее рассмотреть их применение к данной задаче, этап разделяется на несколько частей, описывающих отдельные операции.

Основной структурой данных для генетических алгоритмов является хромосома. Каждая хромосома формируется как набор двоичных чисел. Длина хромосомы равна количеству рассматриваемых связей. Единица или ноль на соответствующей позиции в хромосоме говорит об участии или отсутствии данной связи в этом варианте маршрута.

Популяция представляет собой набор хромосом. Их количество в популяции постоянно. Создание каждой новой популяции (кроме начальной) происходит в результате отбора элиты и генерации новых хромосом. Общий алгоритм этапа приведен на рис. 2.

Исходными данными для алгоритма являются информация о сети, данные об узлах и связях. На основе этих данных выполняются следующие три действия: выбор промежуточных узлов, связей и активных узлов.

Генерация популяции – это процесс задания начального набора хромосом. Подробнее о них было сказано в предыдущем разделе.

Далее в алгоритме идут генетические операторы и действия, такие как отбор элиты, Определение пригодности сочетает в себе вычисление фитнес-функции и штрафной функции одновременно. Действия генетических операторов повторяются до достижения критерия окончания работы алгоритма. В данном случае таким критерием является заданное число популяций. При этом по мере вычисления пригодности хромосомы будет запоминаться наилучший вариант для данной популяции. В соответствии с механизмом отбора элиты он может переходить в следующую

популяцию. Таким образом, на момент окончания работы алгоритма будет определен наилучший из рассмотренных вариантов построения маршрута.

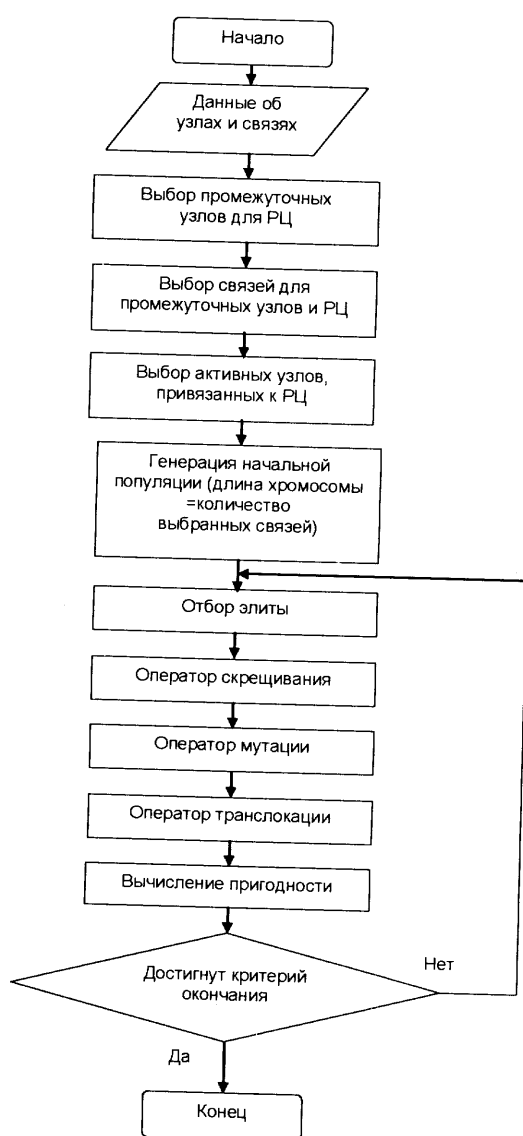


Рис. 2. Алгоритм построения деревьев

Чем меньше стоимость предлагаемого маршрута для конкретной хромосомы, тем больше она имеет шансов стать родителем.

Далее, в соответствии с алгоритмом, производится мутация определенного процента хромосом, который задается как параметр алгоритма. Мутация состоит в обмене местами генов, соседних с выбранным. Выбор этого гена происходит случайным образом.

После мутации применяется оператор транслокации. Как и в предыдущем случае, процент хромосом, подвергаемых транслокации, определяется параметром алгоритма.

После вышеописанных генетических операторов вычисляется пригодность хромосомы. Следует определить действия: «забраковать хромосому» и «забраковать ген». В первом случае считается, что хромосома не пригодна для построения маршрута, так как она не осуществляет

Алгоритм представлен для отдельного регионального центра. В целом же все эти действия выполняются отдельно для каждого РЦ сети.

Процесс отбора элиты применяется для увеличения шансов на получение хорошего результата в следующей популяции. Особенно это эффективно в сочетании с пропорциональным отбором родителей, при котором учитывается фитнес-функция хромосомы.

Таким образом заполняется часть новой популяции. Размер этой части определяется процентом элитизма, который задается как параметр алгоритма. Остальную часть популяции получают с помощью оператора скрещивания. Алгоритм этого процесса учитывает такие настройки, как вид оператора скрещивания: одноточечный или двухточечный. В ходе алгоритма указан процесс выбора родителей.

В этом алгоритме учитывается такой параметр алгоритма, как способ отбора родителей: равномерный или пропорциональный. В равномерном отборе все хромосомы популяции имеют одинаковые шансы стать родителями. Сам выбор осуществляется случайно. При пропорциональном отборе весьма существенным фактором является фитнес-функция хромосомы.

соединение со всеми требуемыми абонентами, участвующими в рассылке. А значит, её не следует учитывать в дальнейшем определении самого дешевого маршрута.

Действие «забраковать ген» означает, что эта связь является избыточной. Она ведет к узлу, который не участвует в рассылке, а потому при построении маршрута эта связь не будет использоваться и её стоимость не будет учитываться при вычислении общей стоимости рассылки. Описанные выше алгоритмы представляют собой адаптацию генетических операторов для поставленной задачи. Исходные данные, результаты и приемы обработки хромосом учитывают особенности используемых структур данных.

Основной сложностью применения генетических алгоритмов для поставленной задачи являются оптимальное кодирование и выбор эффективных генетических операторов. Повышение эффективности алгоритма достигается за счет введения в процедуру поиска операторов мутации, транслокации и рекомбинации. Оператор транслокации до последнего времени не применялся при решении задач оптимизации.

Применение генетических алгоритмов для решения задачи маршрутизации дает, возможно, не самый лучший результат, но главное достоинство в этом случае – относительно небольшое время решения. За счет этого есть возможность рассматривать задачу с различными вариантами генетических операторов до получения приемлемого результата. Окончательное решение о применении предложенного маршрута остается за человеком.

## 6. Оценка эффективности системы

Для оценки эффективности алгоритмов и апробации их на конкретных данных возникает необходимость разработки системы, предоставляющей следующие возможности: возможность синтеза и редактирования сети; визуализация схемы построенного маршрута; вывод результатов

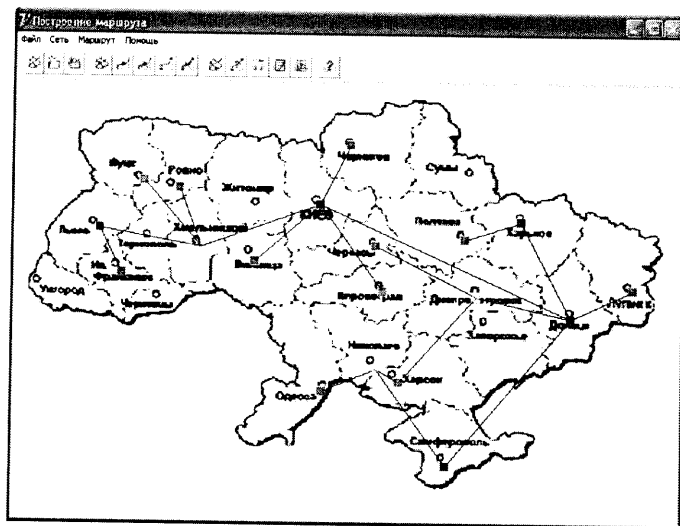


Рис. 3. Пример использования карты

расчетов для пересылки по построенному маршруту; возможность настраивать параметры алгоритма для получения наиболее приемлемых результатов и др.

В качестве среды разработки программного продукта была выбрана среда Borland Delphi 7.0. Этот выбор обусловлен несколькими причинами.

Во-первых, данная среда предоставляет широкие возможности по созданию оконных приложений для ОС Windows. Имеющиеся средства позволяют легко создавать и

редактировать элементы пользовательского интерфейса, наполнять его функциональностью, связывать различные элементы интерфейса между собой и осуществлять их взаимодействие. Это

облегчает создание системы дублирования команд различными элементами пользовательского интерфейса.

Во-вторых, в Borland Delphi 7.0 существуют удобные средства для представления структуры сети. В частности, такой элемент, как изображение (Image), который позволяет размещать объекты в соответствии с заданными координатами, применять какой-либо фон (например, с помощью функции наложения географической карты можно задать структуру реально существующей учебной сети, рис. 3). Это упрощает обработку и визуализацию данных.

В-третьих, выбранная среда позволяет создавать оконные приложения, реагирующие на самые разнообразные события и действия пользователя. Это весьма полезно при основном упоре на работу с мышью. Отображаемая структура сети позволяет определять, что именно сделал пользователь, сколько раз и какую кнопку мыши нажал, на какой части изображения это произошло (на пустом месте или на каком-либо объекте). Это дает возможность существенно расширить функциональность программы, предоставить пользователю больше свободы в работе с сетью и алгоритмом.

С целью проведения анализа эффективности разработанных алгоритмов на вход системы были поданы тестовые наборы данных. Тестирование проводилось для различного числа узлов и связей сети. Результаты представлены на рис. 4.



Рис. 4. Оценка эффективности системы

В качестве сравниваемой величины рассматривалось среднее значение стоимости связи. Для маршрута, полученного с помощью генетических алгоритмов, среднее значение вычислялось среди используемых в конечном маршруте связей. Так как точные методы не могут дать результат для большого числа узлов за приемлемое время, то в качестве альтернативы генетическим алгоритмам было решено выбрать среднее значение стоимости по всем связям сети.

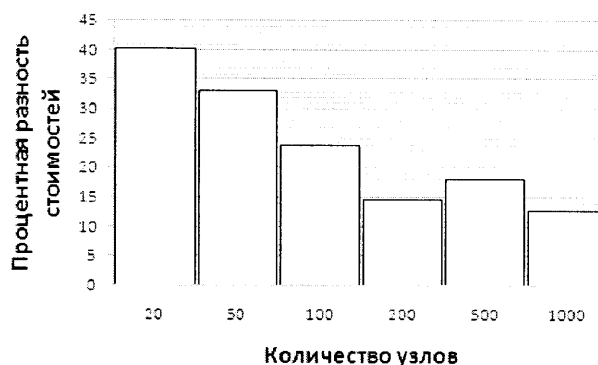


Рис. 5. Процентная разность стоимостей

На представленной диаграмме приведены результаты для групп из 20, 50, 100, 200, 500 и 1000 узлов. Количество связей в каждом случае было несколько больше числа узлов, чтобы обеспечить вариативность соединений. В каждой из групп слева показано среднее значение стоимости связи для связей, используемых в построенном маршруте, а справа – среднее значение по всей сети. На рис. 5

представлена процентная разность стоимостей.

Как видно из диаграммы, для небольшого числа узлов стоимости значительно отличаются. Но с увеличением числа узлов эта разница уменьшается и сглаживается. В среднем стоимость используемых связей всегда меньше, чем средняя стоимость связей в сети вообще.

## 7. Выводы

В данной статье приводится обоснование выбора построения маршрута сети дистанционного обучения как многоуровневой системы. На первом этапе осуществляется привязка к региональным центрам, что представляет собой некую предварительную фильтрацию. Таким образом, на втором этапе построения генетические алгоритмы ищут решение среди более перспективных вариантов.

Был проведен анализ эффективности разработанной системы на тестовых наборах данных. В среднем стоимость используемых связей была почти всегда меньше, чем средняя стоимость связей в сети вообще (за исключением сетей с небольшим набором узлов, где стоимости были почти равны). Также было замечено, что более выгодные маршруты предлагаются при использовании пропорционального оператора отбора.

Таким образом, проведенный анализ показал, что разработанная многоуровневая система представляет собой синтез точных и эвристических методов и применение предложенного подхода позволяет сочетать быстродействие с эффективностью.

В разработанной системе предусмотрены возможности для дальнейшего развития, в том числе применения новых эвристических алгоритмов.

Разработанная многоуровневая система позволяет строить маршруты за приемлемое время, выбирая подходящий вариант путем изменения настроек алгоритмов пользователем. Практическое применение работы состоит в том, что данную систему можно применить к любой территориально распределенной сети. Возможность использовать различные настройки программы поможет подобрать наиболее приемлемое решение в каждом конкретном случае рассылки.

## СПИСОК ЛИТЕРАТУРЫ

1. Ольшевский А.И. Интерактивная система проектирования топологии сетей дистанционного обучения на базе ДонГИИИ / А.И. Ольшевский // Искусственный интеллект. – 2001. – № 2. – С. 52 – 56.
2. Маршалл У. Берн Поиск кратчайших сетей / Маршалл У. Берн, Рональд Л. Грэм // Scientific American (издание на русском языке). – 1989. – № 3. – С. 64 – 70.
3. Рыженко Н.В. Алгоритм построения минимальных связывающих деревьев с дополнительными вершинами (деревьев Штейнера) для случая прямоугольной метрики / Н.В. Рыженко // Труды ИМВС РАН. – 2002. – Режим доступа: [http://www.mipt.ru/nauka/conf\\_mipt/conf2001/frtk/vychteh/ryjenko.html](http://www.mipt.ru/nauka/conf_mipt/conf2001/frtk/vychteh/ryjenko.html).
4. Калашников Р.С. Построение дерева Штейнера методом генетического поиска / Р.С. Калашников // Перспективные информационные технологии и интеллектуальные системы. – 2005. – № 2 (22). – С. 54.

*Статья надійшла до редакції 28.04.2009*