

М.М. Митев, Е.В. Рачева

Алгоритм первоначального группирования понятий в дистанционном обучении

Описано создание алгоритма, предназначенного для упорядочивания по оси времени понятий и их группирования в учебные единицы в рамках дисциплины, предложенной для дистанционного обучения. Алгоритм учитывает связи между понятиями, участвующими в учебных единицах из предыдущих и последующих изучаемых дисциплин.

The creation of an algorithm is described designed to organize the concepts along the time axis and their clustering in the training units within the discipline suggested for the distance-learning. The algorithm takes into account the connection between the concepts involved in the training units of the previous and subsequent courses.

Описано створення алгоритму, призначеного для упорядкування по осі часу понять та їх групування в навчальні одиниці в межах дисципліни, запропонованої для дистанційного навчання. Алгоритм враховує зв'язки між поняттями, що беруть участь у навчальних одиницях із попередніх та наступних дисциплін, що вивчаються.

Введение. Дистанционное обучение как форма и организация учебного процесса бесспорно получила широкое признание. В основном подготовка учебного материала носит субъективный характер и является результатом технических познаний и методических умений преподавателя. Опыт авторов в использовании генетического алгоритма для структурирования учебного материала представлен в [1]. В тех же целях для формализации использована теория графов, а также метод сортировки в графах с надстройкой для упорядочивания понятий по оси времени [2]. Разработанный алгоритм базируется на относительно самостоятельной единице *понятие* и на причинно-следственных связях между понятиями. В результате применения алгоритма понятия, определяющие данную дисциплину, упорядочиваются на основе их семантической близости. Варианты алгоритма упорядочивания понятий из так называемого *стартового пакета*, учитывающего связи с понятиями из предыдущих дисциплин согласно учебному плану, предложены в [3]. Аналогично разработаны варианты решения по формированию и упорядочиванию понятий в *финальном пакете*, учитывающем связи с понятиями из следующих дисциплин. В публикации [4] представлено обобщение полученных результатов и предложена единая система алгоритмов. Полная формализация, теоретическая постановка, разработанные алгоритмы и их программная реализация приведены в [5, 6].

Цель настоящей статьи – дальнейшее развитие идей и алгоритмов упорядочивания поня-

тий по оси времени, а также представление решения об их группировании в учебные единицы в соответствии с принятой организацией и методикой дистанционного обучения. Процесс группирования (*сегментирования*) понятий может быть реализован в два этапа:

- первоначальное (*статическое*) группирование без учета индивидуальных способностей обучаемого и
- адаптивное (*динамическое*) группирование в процессе собственно обучения в зависимости от полученных индивидуальных результатов.

Результаты первого этапа направлены на планирование и организацию учебного процесса при групповом и индивидуальном обучении. Второй этап полностью ориентирован на условия и цели дистанционного обучения.

Формализация

Учебный материал по конкретной дисциплине может быть представлен в качестве множества взаимно связанных понятий. Формально, каждое понятие можно интерпретировать как вершину графа, а связи между понятиями – как ориентированные ребра. Следовательно, представление дисциплины можно свести к конечному ориентированному графу $G(X, U, P, W, F)$, где согласно [2, 3] X – множество вершин (понятий) мощностью N , U – множество ребер (причинно-следственных связей), $P(x_i u_{ij} x_j)$ – трехместный предикат, инцидентор, W – множество весов вершин и F – множество статических весовых коэффициентов ребер. Существуют следующие одно-однозначные соответствия: $X \leftrightarrow W$ и $U \leftrightarrow F$.

Теоретическая постановка

Пусть задано решение $r \in R$, упорядочивающее понятия последовательно по оси времени. Решение определяет динамический вес каждого ребра в зависимости от местоположения инцидентных ему вершин и вершин, расположенных между ними [2].

Согласно [4] вес каждой учебной единицы ограничен W_0 . Следовательно, количество учебных единиц в определенной дисциплине задается формулой:

$$M = \frac{W}{W_0} \quad (1)$$

с точностью до первого большего целого числа, где

$$W = \sum_{i=1}^N w_i \quad (2)$$

представляет собой общий вес всех понятий в дисциплине.

Для каждой m -й учебной единицы, определяемой n -м количеством понятий с общим весом W_m , возникает разница Δw_m :

$$\Delta w_m = W_0 - W_m = W_0 - \sum_{k=1}^n w_k \quad (3)$$

при ограничительном условии:

$$\sum_{k=1}^n w_k \leq W_0. \quad (4)$$

Общее количество учебных единиц и включенных в них понятий представляет граф $G = \{g_m\}$, $m \in 1, M$ как множество непересекающихся подграфов мощностью M .

Ребра, инцидентные только вершинам данного подграфа, называются *внутренними*, а связывающие два подграфа – *внешними*.

В общем случае решение $r_0 \in R$, можно оценить по следующим показателям:

- минимальная сумма возникших разниц;
- минимальное количество внешних ребер и, соответственно, максимальное количество внутренних ребер;
- минимальная сумма статических весов внешних ребер;
- минимальная сумма динамических весов внешних ребер или
- с учетом весовых коэффициентов, отражающих разницу в динамических весах внутренних и внешних ребер.

Не нарушая общности рассуждений, задаем два соседних подграфа $g_m, g_{m+1} \in G$, определенных подмножествами X_m и X_{m+1} , и две произвольных вершины из этих подмножеств: $x_i \in X_m$ и, соответственно, $x_j \in X_{m+1}$.

Если принять, что $u_{ij} \in \{0, 1\}$, то общее количество (суммарный вес) ребер, инцидентных двум вершинам и вершинам определенных подграфов, можно определить по формуле:

$$V_{x_i} = \sum_{\forall x_k \in X_m} (u_{ik} + u_{ki}) - \sum_{\forall x_k \in X_{m+1}} (u_{ik} + u_{ki}), \quad (5)$$

$$V_{x_j} = \sum_{\forall x_k \in X_{m+1}} (u_{jk} + u_{kj}) - \sum_{\forall x_k \in X_m} (u_{jk} + u_{kj}), \quad (6)$$

$$V_{x_i x_j} = V_{x_i} + V_{x_j} - (u_{ij} + u_{ji}), \quad (7)$$

$$P(x_i u_{ij} x_j) = \text{true} \rightarrow (u_{ij} = 1) \&$$

$$\& P(x_i u_{ij} x_j) = \text{false} \rightarrow (u_{ij} = 0). \quad (8)$$

Аналогично определяются суммы статических весовых коэффициентов ребер подграфов $x_i \in X_m$ и $x_j \in X_{m+1}$:

$$V_{x_i} = \sum_{\forall x_k \in X_m} (f_{ik} + f_{ki}) - \sum_{\forall x_k \in X_{m+1}} (f_{ik} + f_{ki}), \quad (9)$$

$$V_{x_j} = \sum_{\forall x_k \in X_{m+1}} (f_{jk} + f_{kj}) - \sum_{\forall x_k \in X_m} (f_{jk} + f_{kj}), \quad (10)$$

$$V_{x_i x_j} = V_{x_i} + V_{x_j} - (f_{ij} + f_{ji}), \quad (11)$$

$$(\forall u_{ij} \in U)(\exists f_{ij} \in F) \& (\forall f_{ij} \in F)(\exists u_{ij} \in U). \quad (12)$$

Если статические коэффициенты ребер константны, то формулы (9), (10) и (11) идентичны формулам (5), (6) и (7). В противном случае результаты будут иными.

Пусть задана операция $\alpha(x_i x_j)$, меняющая местами две вершины. Это определяет новое решение $r_\alpha \in R$ и, соответственно, новую сумму $V_{x_j x_i}$. Если

$$V_{x_j x_i} > V_{x_i x_j}, \quad (13)$$

то это означает, что новое решение определяет более благоприятное соотношение внешних ребер с внутренними относительно количества или сумм статических коэффициентов, т.е. численность внутренних ребер увеличивается

по отношению к внешним. Операция $\alpha(x_i x_j)$, примененная к каждой паре вершин соседних подграфов и всех пар соседних подграфов, определит общую сумму *внешних ребер* или общую сумму статических коэффициентов ребер, которая будет меньше первоначальной.

Данная постановка не учитывает местоположение каждой вершины в упорядоченной последовательности вершин.

Пусть при решении $r_0 \in R$, определяющем $G = \{g_m\}$ вычисляются динамические коэффициенты ребер:

$$V_{x_i} = \sum_{\forall x_k \in X_m} (v_{ik} + v_{ki}) + \sum_{\forall x_k \in \forall X_{s>m}} (v_{ik} + v_{ki}) - \sum_{\forall x_k \in \forall X_{s<m}} (v_{ik} + v_{ki}), \quad (14)$$

$$V_{x_j} = \sum_{\forall x_k \in X_{m+1}} (v_{jk} + v_{kj}) + \sum_{\forall x_k \in \forall X_{s>m}} (v_{jk} + v_{kj}) - \sum_{\forall x_k \in \forall X_{s<m}} (v_{jk} + v_{kj}), \quad (15)$$

$$V_{x_i x_j} = V_{x_i} + V_{x_j}, \quad (16)$$

$$(\forall x_i x_k \subset X_m) (i < k) P(x_i u_{ij} x_j) \rightarrow \left(v_{ik} = f_{ik} \sum_{q=i}^{k-1} w_q \right) + \left(v_{ki} = A f_{ki} \sum_{q=k}^{i-1} w_q \right), \quad (17)$$

$$(\forall x_i \in X_m) (\forall x_k \in X_s) (m < s) P(x_i u_{ij} x_j) \rightarrow \left(v_{ik} = f_{ik} \sum_{q=i}^{k-1} w_q \right) + B \left(v_{ki} = f_{ki} \sum_{q=k}^{i-1} w_q \right), \quad (18)$$

где по формуле (17) вычисляются динамические веса *внутренних* ребер с коэффициентом A для ребер в обратном направлении по оси времени. Формула (18), соответственно, учитывает динамические веса *внешних* ребер с коэффициентом B для ребер с направлением, также обратном оси времени.

Проблема оптимизации логической связанности отдельных учебных единиц может рассматриваться и решаться, когда необходимо:

- внутреннее перераспределение понятий по оси времени в границах определенной учебной единицы;
- снижение суммы динамических коэффициентов внешних ребер;
- проектное решение, учитывающее оба случая одновременно.

Первый случай сводится к применению алгоритма из [2] для структурирования понятий для каждого подграфа отдельно. Заложенный в алгоритме критерий связан с переупорядочиванием понятий при снижении суммы динамических коэффициентов ребер по оси времени в прямом и обратном направлении или при выполнении частных критериев: только в прямом или только в обратном направлении. Применение данного алгоритма *сближает* сильно связанные понятия, сокращая время перехода от одного базового понятия к следующему. При этом не учитываются изменения в динамических коэффициентах внешних ребер, что, в общем случае, может привести к определению решения с неудовлетворительными показателями связи между отдельными учебными единицами.

Во *втором* случае решение может быть сведено к выполнению алгоритма по определению и упорядочиванию стартового [3] и финального пакетов, примененного к каждому подграфу отдельно, с или без учета интердисциплинарных связей. В данном случае *стартовый пакет* включает в себя понятия, связанные с предыдущими учебными единицами или дисциплинами по учебному плану. Аналогично, *финальный пакет* содержит понятия, связанные с понятиями из следующих учебных единиц или дисциплин. Использование этих алгоритмов *сближает* во времени понятия из различных учебных единиц, соответственно из предыдущих или последующих дисциплин, но может определить решения, которые не будут рациональными в отношении понятий, включенных в отдельные учебные единицы.

Целостное решение в отношении упорядочивания понятий по оси времени и их группирования в учебные единицы базируется на минимизации весовой разницы Δw_m и, соответственно, нахождении локального минимума вычисленной по формуле (16) суммы. В общем случае интердисциплинарные связи интерпретируются как внешние ребра и им может быть присвоен коэффициент C , чтобы отличать их от связей в самой дисциплине.

Алгоритм предварительного группирования понятий

Основными этапами алгоритма предварительного группирования понятий в учебных единицах заданной дисциплины являются:

- задание списка понятий – объекта изучения в рассматриваемой дисциплине, установление существующих причинно-следственных связей между ними, экспертное определение относительного веса каждого понятия и статических коэффициентов связей между ними;

- формальное представление данных с помощью конечного ориентированного графа и определение первого решения $r_0 \in R$, путем произвольного индексирования вершин и применения алгоритма упорядочивания понятий по оси времени [2];

- группирование понятий в учебные единицы при соблюдении ограничительного условия (4) и вычисление возникших разниц Δw_m ;

- выбор варианта решения, минимизирующего суммы:

- внешних или, соответственно, внутренних ребер;
- статических коэффициентов внутренних и внешних ребер или
- динамических коэффициентов ребер с учетом или без учета стартового и/или финального пакета понятий или интердисциплинарных связей;

- организация цикла для обхода вершин всех пар последовательно расположенных подграфов;

- применение операции $\alpha(x_i, x_j)$ при ограниченном условии;

- если новое решение удаляется от локального экстремума, восстанавливается первоначальное расположение вершин;

- исполнение алгоритма прекращается по исчерпанию всех возможных комбинаций пар вершин.

Заключение. Предложенный алгоритм первоначального группирования понятий в комбинации с остальными алгоритмами из цитированных литературных источников реализован программно в единой системе для разработки, моделирования и исследования курсов дистанционного обучения.

Разработанную программную систему или ее отдельные части можно использовать самостоятельно как инстру-

ментальное средство при подготовке курсов или включать в различные платформы [7], созданные для организации и проведения дистанционного обучения, если программная среда позволяет расширение.

Результаты предварительных исследований генерированных случайных структур понятий показывают следующее:

- алгоритм значительно сокращает сумму динамических коэффициентов ребер, интерпретирующих причинно-следственные связи между группами понятий;

- алгоритм показывает значительное повышение эффективности при анализе, исследовании и группировании учебных единиц, состоящих из сильно связанных понятий;

- в предложенном варианте алгоритм принадлежит к классу комбинаторных. Обычно он применяется однократно при первоначальной подготовке курса дистанционного обучения, и поэтому время получения проектного решения не имеет существенного значения.

На данной основе, как следующий этап, может быть разработан адаптивный алгоритм динамического группирования понятий в целях индивидуального обучения.

1. Рачева Е., Митев М., Калчева Е. Генетичен алгоритъм за структуриране на учебното съдържание за целите на дистанционно обучение // Списание «Компютърни науки и технологии», бр. 1/2008 г., ТУ–Варна.
2. Митев М., Рачева Е. Разработка на алгоритъм и програма за структуриране на учебен материал // Там же, бр. 2/2008 г., ТУ–Варна.
3. Mitev M., Racheva E., Marinova G. Algorithm for Defining of Initial Package of Notions for Purpose of E-Learning, CompSysTech'2009, Ruse, Bulgaria, 2009.
4. Митев М., Рачева Е., Маринова Г. Графоаналитический подход для разработки курсов дистанционного обучения // Четвертая Междунар. конф. «Новые информационные технологии в образовании для всех: инновационные методы и модели», Киев, Украина, 24–26 нояб. 2009 г. – С. 127–136.
5. Проект ФНИ №483. Разработване на учебен курс по дисциплината «Синтез и анализ на алгоритми» за целите на дистанционно обучение с изследване на начините за представяне на учебния материал и създаване на специализирана база от данни с текста и графична информация», ТУ–Варна, 2008.
6. Проект НП № 25. Теоретична разработка и реализация на програмна платформа за компютърно проектиране и изследване на курсове за дистанционно обучение, ТУ–Варна, 2009.
7. Learning Technology Standards Committee, Draft Standard for Learning Object Metadata // IEEE Standard 1484.12.1, New York: Inst. of Electr. and Electron. Eng.

E-mail: mit.evmm@abv.bg
© М.М. Митев, Е.В. Рачева, 2010