

УДК 321.3

## ФОРМИРОВАНИЕ ИНДИВИДУАЛЬНОЙ ОБРАЗОВАТЕЛЬНОЙ ТРАЕКТОРИИ В СИСТЕМАХ ДИСТАНЦИОННОГО ОБУЧЕНИЯ

Купневич Ю.Н., Фототов А.М.

Донецкий национальный технический университет, Украина.

*Разработана модель структуризации учебного контента на основе нечетких множеств и нечетких отношений, реализован алгоритм формирования индивидуальной образовательной траектории на основе формализованных моделей связности учебного материала и результатов входного тестового контроля.*

### Общая постановка проблемы

Дистанционное обучение все больше пользуется популярностью. В нашей стране и на территории СНГ дистанционным обучением пользуются в основном для повышения квалификации, получения второго или заочного образования. Чаще всего слушателями дистанционных курсов становятся люди ограниченные во времени и дорожащие им.

Целью подсистемы формирования индивидуальной образовательной траектории является предоставление обучающимся наиболее удобного для них индивидуального плана. В настоящее время системы дистанционного обучения (СДО) в полной мере используют современные программно-инструментальные средства управления учебным процессом. Однако проблематика динамического управления образовательной траекторией с учетом формализованных моделей взаимосвязи образовательных ресурсов остается открытой. В связи с этим становится актуальной задача по автоматизации процесса формирования индивидуальных учебных планов и программ на основе начального уровня подготовки каждого обучаемого, а так же объединения моделей структуризации учебного материала и тестового контроля.

В результате анализа предметной области мы имеем: набор обучаемых ( $Q_i$ ); разный уровень подготовки обучаемых ( $L_{ij}$ ) по дисциплинам ( $W_j$ ); различный объем свободного времени каждого обучающегося; разный уровень усвоения материала; взаимосвязь между изучаемыми дисциплинами. На основе имеющейся информации необходимо построить индивидуальную образовательную траекторию для каждого обучающегося.

Индивидуальная образовательная траектория представляет собой индивидуальный учебный план, который составлен на основе первичной информации.

Для решения поставленной задачи необходимо:

- представить образовательный контент в виде множества взаимосвязанных элементов;
- определить уровень знаний обучаемого;
- определить набор дисциплин и тем, которые должны быть изучены;
- составить план обучения с учетом количества доступного времени у обучаемого, уровня его знаний, связности дисциплин и тем.

### Представление образовательного контента

В работе [1] была описана модель представления образовательного контента в виде иерархической сетевой модели данных Tree-Net. Структура разрабатываемой модели с одной стороны является иерархической, а с другой имеет черты семантической сети (рисунок 1).

В статье [2] была предложена модификация модели представления контента с помощью нечетких множеств. Рассмотрим черты нечеткой модели, описанные в [2].

Множество  $G$  указывает на тематические группы, в которых могут участвовать элементы контента.

$$G = \{g_1, \dots, g_{n_G}\}, \quad (1)$$

где  $nG$  — количество тематических групп контента.

Простейшим элементом контента является представление  $(v_i)$  — подтема, которое соответствует одной Web-странице сайта. Множество всех элементов контента:

$$V = \{v_i\}, \quad (2)$$

где  $i = 1, \dots, nV$ .

Каждый элемент контента может принимать участие в произвольном количестве тематических групп с вероятностью попадания элемента в ту или иную тематическую группу. Применим теорию нечетких множеств для представления множества  $G$  тематических групп:

$$G = \{(g_i, v_j, \mu_{g_i v_j}(g_i, v_j))\}, \quad (3)$$

где  $\mu_{g_i v_j}(g_i, v_j)$  — функция принадлежности, которая ставит в соответствие каждому элементу контента  $v_j$ , принимающему участие в тематической группе  $g_i$ , некоторое действительное число из  $[0, 1]$ .

### Определение начального уровня знаний обучаемого

Для определения уровня знаний обучаемого по каждой дисциплине введем и опишем входные лингвистические переменные «оценка» и «уровень сложности».

Определим лингвистическую переменную «оценка», заданную кортежем  $\langle \text{оценка}, T, U, G, M \rangle$ , где  $T = \{\text{плохо, удовлетворительно, хорошо, отлично}\}$  – множество первичных термов;  $U = [0, 5]$  – универсальное множество, т.е. область определения значений лингвистической переменной оценка;  $G$  – синтаксическая процедура образования новых термов с помощью нечетких связей «и», «или» и модификаторов «не», «очень»;  $M$  – семантическая процедура, которая ставит в соответствие каждому значению лингвистической переменной (каждой нечеткой переменной) определенное нечеткое множество.

Определим лингвистическую переменную «уровень сложности», заданную кортежем  $\langle \text{уровень сложности}, T, U, G, M \rangle$ , где  $T = \{\text{начальный, средний, сложный}\}$  – множество первичных термов;  $U = [0, 2]$  – универсальное множество.

Определим выходную лингвистическую переменную «уровень прохождения темы», заданную кортежем  $\langle \text{уровень прохождения темы}, T, U, G, M \rangle$ , где  $T = \{\text{все, от среднего, высокий}\}$  – множество первичных термов;  $U = [0, 1]$  – универсальное множество.

Опишем алгоритм определения индивидуального уровня знаний обучаемого после прохождения входного тестирования:

1. Проходим первоначальный тест. Получаем множество результатов по каждому из вопросов (4):

$$R = \{(q_i, r(q_i))\}, \quad (4)$$

где  $q_i$  – вопрос,  $r(q_i)$  – результат ответа на вопрос.

Множество принадлежности вопроса к тематической группе (5):

$$Pg = \{(q_i, pg(q_i))\}, \quad (5)$$

где  $q_i$  – вопрос,  $pg(q_i)$  – группа, к которой принадлежит вопрос.

Множество «сложность вопроса»:

$$S = \{(q_i, s(q_i))\}, \quad (6)$$

где  $q_i$  – вопрос,  $s(q_i)$  – сложность вопроса.

2. На следующем шаге берем вопросы из первой тематической группы и переносим нужные данные в множество  $R_n$  из множества  $R$ , и в множество  $Pg_n$  из множества  $Pg$ , и в множество  $S_n$  из множества  $S$ .

3. Получаем результаты прохождения теста по выбранной тематической группе (оценку и уровень сложности прохождения).

Для определения оценки используем формулу (7):

$$O = \frac{\sum_{j=1}^n r(q_j)}{n} \quad (7)$$

где  $n$  – количество вопросов,  $r(q_j)$  – результат ответа на вопрос.

Для определения сложности прохождения используем формулу (8):

$$SI = \frac{\begin{cases} \sum_{j=1}^n s(q_j), r(q_j) \geq 3 \\ 0, r(q_j) \leq 2 \end{cases}}{n}, \quad (8)$$

где  $s(q_j)$  – сложность вопроса,  $r(q_j)$  – результат ответа на вопрос.

4. По алгоритму нечеткого вывода Сугена:

4.1. При формировании базы правил используются правила вида:

ЕСЛИ «оценка» = «плохо» И «уровень сложности» = «начальный» ТО «сдаваемые темы» = «все»;

ЕСЛИ «оценка» = «хорошо» И «уровень сложности» = «сложный» ТО «сдаваемые темы» = «от среднего» и т.д.

4.2. Фазификация входных переменных. До начала фазификации известны конкретные значения всех входных переменных системы нечеткого вывода: оценка, уровень сложности. Далее рассматривается каждое правило, где есть «оценка» и «уровень сложности», и находится количественное значение  $b'_i$  – это значение является результатом фазификации.

4.3. Агрегирование. До начала агрегирования известны значения истинности всех правил системы нечеткого вывода. Далее рассматривается каждое из условий правила. При этом для определения результата ( $b'$ ) используется функция минимум.

Результатом агрегирования является степень истинности всего условия правила.

4.4. Активизация. На данном этапе сначала находится значение степени истинности всех заключений. Затем находится значение  $w$ , где  $w$  = «сдаваемые темы».

Таким образом, получаем множество значений с координатами  $w$  и весом  $C_i$ , где  $C_i = b'$

Результатом активизации заключения является функция принадлежности полученного из исходной функции принадлежности заключения.

4.5. Аккумуляция фактически отсутствует, т.к. расчеты осуществляются с обычными действительными числами, а не с функциями принадлежности.

4.6. На этапе дефазификации используется модифицированный вариант в форме метода центра тяжести для одноточечных множеств (9):

$$y = \frac{\sum_i C_i W_i}{\sum_i C_i}. \quad (9)$$

Итог дефазификации: процент прохождения подтем в теме.

В результате получаем множество с номерами страниц, которые будут изучаться в выбранной тематической группе:

$$N_k = \{st_j\}, \quad (10)$$

где  $st_j$  – номер страницы.

5. На следующем шаге берем следующую тематическую группу (делаем тоже самое с множествами, что и на шаге 2) и возвращаемся к шагу 3, если группы закончились, переходим к шагу 11.

6. В результате получаем индивидуальный план обучения, множество  $I = \{N_k\}$ .

После прохождения теста получаем индивидуальный план который записывается в множество  $I = \{N_k\}$ . Фактическое представление (матрица) имеет следующий вид (11):

$$\begin{array}{cccc}
 I_{11} & I_{12} & I_{1i} & I_{1n_1} \\
 I_{21} & I_{22} & I_{2i} & I_{2n_2} \\
 I_{j1} & I_{j2} & I_{ji} & I_{jn_j} \\
 I_{m1} & I_{m2} & I_{mi} & I_{mn_m}
 \end{array} \tag{11}$$

где  $m$  – количество тематических групп;  $n_j$  – количество элементов контента для каждой тематической группы;  $I_{ji}$  – элемент контента.

### Формирование индивидуальной образовательной траектории

Для формирования индивидуальной образовательной траектории воспользуемся классическим генетическим алгоритмом.

Индивидуальный план обучаемого характеризуется порядком изучения дисциплин и количеством часов в семестре. Есть связанные и несвязанные дисциплины. Связанные дисциплины желательно изучать последовательно. Уровень связности определяется на основе построенной ранее иерархии дисциплин.

Для начала определим фитнес-функцию:

$$f = k_1 * \sum x_i + k_2 * \sum y_i, \tag{12}$$

где  $k_1$  и  $k_2$  – коэффициент значимости;  $x_i$  – отклонение от заданного числа часов в  $i$ -м семестре, вычисляется по формуле (13);  $y_i$  – уровень связности дисциплин в семестрах. При правильной связности

$$x_i = 1 - \frac{t_z - t_{ip}}{t_z}, \tag{13}$$

где  $t_{ip}$  – полученное число часов,  $t_z$  – число часов в семестре.

В качестве способа задания особи выберем классическую хромосому, в которую закодируем перечень кодов изучаемых дисциплин (каждая дисциплина состоит из набора тематических групп  $I_{kl}$ ). При декодировании хромосомы будем последовательно извлекать из нее коды изучаемых дисциплин, набирая для текущего семестра множество дисциплин, так чтобы выполнялось неравенство (14).

$$\begin{cases} \sum_{1 < n_j}^i t_{ji} \leq t_z, & p_j \leq P_3 \\ \sum_{1 < n_j}^i t_{ji} > t_z, & p_j > P_3 \end{cases}, \tag{14}$$

где  $P_3$  – вероятность что в  $j$ -ом семестре будет вычитано больше часов чем заданное  $t_z$ .

Создадим начальную популяцию особей, где особь – номер предмета. Начальная популяция имеет вид: 1, 3, | 5, 3, 2, | 4, 6, | 6, | 7, 6, 8.

Применяя стандартные операции кроссинговера и мутации, получаем разнообразные образовательные траектории. Алгоритм заканчивает работу если в течении пяти популяций не получено лучшего решения.

### Выводы

Описанный подход позволил на основе связности и построенной иерархии тематических групп в учебном плане построить индивидуальную образовательную траекторию, учитывая пожелания пользователя относительно отводимого объема времени в семестр, а так же его начальную подготовку.

Предложенная модель и генетический алгоритм формирования индивидуальной образовательной траектории позволит комплексно подойти к решению задач обучения в организациях

и учреждениях различного характера. Семантическая составляющая и дерево тематических групп дает возможность построения иерархии предметных областей.

### Литература

- [1] Титенко С.В., Гагарин О.О. Формування навчального контенту на основі моделі даних Tree-Net. Матеріали Всеукраїнської науково-технічної конференції «Комп'ютерна математика в інженерії, науці та освіті» (CMSEE-2007),— Полтава: ПолНТУ, 2007. – 42с.
- [2] Купневич Ю.Н., Фонотов А.М. Представление модели учебного контента в системах дистанционного обучения. Сборник материалов второй всеукраинской научно-технической конференции студентов, аспирантов и молодых ученых — Донецк: ДонНТУ, 2011. – с.79-83
- [3] Скобцов Ю.А. Основы эволюционных вычислений: учебное пособие для вузов/ Ю.А. Скобцов. – Донецкий нац. техн. ун-т. – Донецк: ДонНТУ, 2008.
- [4] Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский; пер. с польск. И.Д. Рудинский – М.: Горячая линия-Телеком, 2006. – 452 с.
- [1] Кофман А. Введение в теорию нечетких множеств: Пер с франц. – М.: Радио и связь, 1982. – 432 с.