

Роевой алгоритм синтеза учебных пособий на основе онтологического подхода

Рассматривается формальный метод создания индивидуальных траекторий обучения на модели образовательного пространства в онтологических базах образовательных ресурсов представляемых в виде И/ИЛИ графовой модели. Предлагается роевой алгоритм построения оптимального маршрута обучения, основанный на моделях адаптивного поведения муравьиной колонии.

Ключевые слова: электронное учебное пособие, траектория обучения, онтология, синтез учебных пособий, роевой алгоритм, муравьиная колония.

SWARM ALGORITHM OF SYNTHESIS TUTORIALS BASED ON ONTOLOGICAL APPROACH

We consider a formal method to create individual learning paths on the model of educational space in the ontological bases of educational resources presented in the form of AND / OR graph model. It is proposed swarm learning algorithm for constructing an optimal route based on models of adaptive behavior of ant colony.

Keywords: electronic textbook, a trajectory learning, ontology, the synthesis of textbooks. swarm algorithm, ant colony.

Введение

Основу образовательного процесса в открытом образовании составляет целенаправленная, контролируемая, интенсивная самостоятельная работа обучающегося, который может учиться в удобном для себя месте, по индивидуальному расписанию, имея при себе комплект специальных средств обучения. На сегодняшний день в рамках семантических технологий наиболее активно исследуется и развивается онтологический подход к представлению знаний предметной области, на основании которого разрабатываются интеллектуальные информационные системы, и в том числе порталы знаний, предназначенные для обеспечения эффективного обмена знаниями и взаимодействия в процессе совместной интеллектуальной распределенной деятельности [1]. Основным требованием к современным учебным ресурсам, используемым в обучающих системах управления

знаниями, является способность адаптироваться к конкретным задачам обучения, уровню компетентности и индивидуальным характеристикам личности обучающегося. Данная задача решается за счет модульности электронных образовательных ресурсов, что позволяет впоследствии легко выстраивать траекторию обучения [2]. Для устранения недостатка отсутствия связей между автономными модулями применяется интегрирующая среда управления знаниями на основе предметных онтологий, являющихся совокупностью понятий в области знаний и отношений между ними, включая также способы интерпретации понятий и отношений. Онтология является моделью представления знаний, которая может использоваться для описания семантики объектов предметно-ориентированных информационных систем. В этом случае межпонятийные отношения выражаются дугами-ссылками вершин-понятий (концептов) онтологии. Определе-

ние необходимого состава учебных модулей может быть основано на навигации по семантической сети [2, 3]. Навигация по семантической сети понятий основана на построении имитационной графовой модели, использующей для начала поиска некоторое ключевое слово, выражающее запрос пользователя, и список понятий, связанных с текущим понятием и формирующим его окрестность. Анализируемые понятия позволяют сформировать список учебных модулей, в которых определено текущее понятие. [4, 5]. При таком подходе настройка осуществляется за счет выбора путей навигации в семантической сети и принятия решений по выбору концепта в каждой вершине сети. В работе [4] рассматривается формальный метод создания индивидуальных траекторий обучения на модели образовательного пространства в онтологических базах образовательных ресурсов представляемых в виде И/ИЛИ графовой модели. Основным не-



Борис Константинович Лебедев,
 д.т.н., профессор кафедры
 Системы автоматизированного
 проектирования
 Тел.: (928) 289-79-33
 Эл. почта: lebedev.b.k@gmail.com
 Южный федеральный университет
 www.sfedu.ru

Boris K. Lebedev,
 Doctorate of Engineering Science,
 Department of Computer Aided Design
 Tel.: (928) 289-79-33
 E-mail: lebedev.b.k@gmail.com
 Southern Federal University
 www.sfedu.ru



Олег Борисович Лебедев,
 к.т.н., доцент кафедры
 Системы автоматизированного
 проектирования
 Тел.: (908) 513-55-12
 Эл. почта: oblebedev@sfedu.ru
 Южный федеральный университет
 www.sfedu.ru

Oleg B. Lebedev,
 Post-graduate student, Department of
 Computer Aided Design
 Tel.: (908) 513-55-12
 E-mail: oblebedev@sfedu.ru
 Southern Federal University
 www.sfedu.ru

достатком предложенного подхода является фактически ручной режим полного или частичного перебора маршрутов для выбора наилучшего из них в семантической сети концептов, опирающийся на минимизацию булевых функций [3]. В работе предлагается роевой алгоритм построения оптимального маршрута обучения, позволяющий получать решения, близкие к оптимальным.

Модель образовательного пространства в онтологических базах образовательных ресурсов

При описании модели образовательного пространства и метода синтеза учебных пособий будем использовать термины и определения, приведенные в работе **Норенкова И.П., Соколова Н.К.** [4].

Целевой концепт – концепт (понятие), который должен быть изучен с помощью синтезируемого пособия.

Исходный концепт – концепт, уже изученный пользователем.

Семантическая сеть концептов – сеть, вершинам которой соответствуют концепты, а дугам – отношения между ними. Так, если X – определяющее понятие, а Y – определяемое, то дуга направлена от вершины X к вершине Y.

Семантическая сеть модулей – сеть, вершинам которой соответствуют модули. Дуга X ->Y имеет место, если некоторое понятие определено в модуле X и используется в модуле Y.

Понятие, определяемое в модуле, называется выходным понятием этого модуля, а понятия, используемые в модуле для пояснения выходного понятия, называются входными понятиями этого модуля.

Модель образовательного пространства представляет собой И-ИЛИ-граф G, образованный объединением семантических сетей концептов и модулей.

Отметим, что вершинам концептов соответствует связка ди-

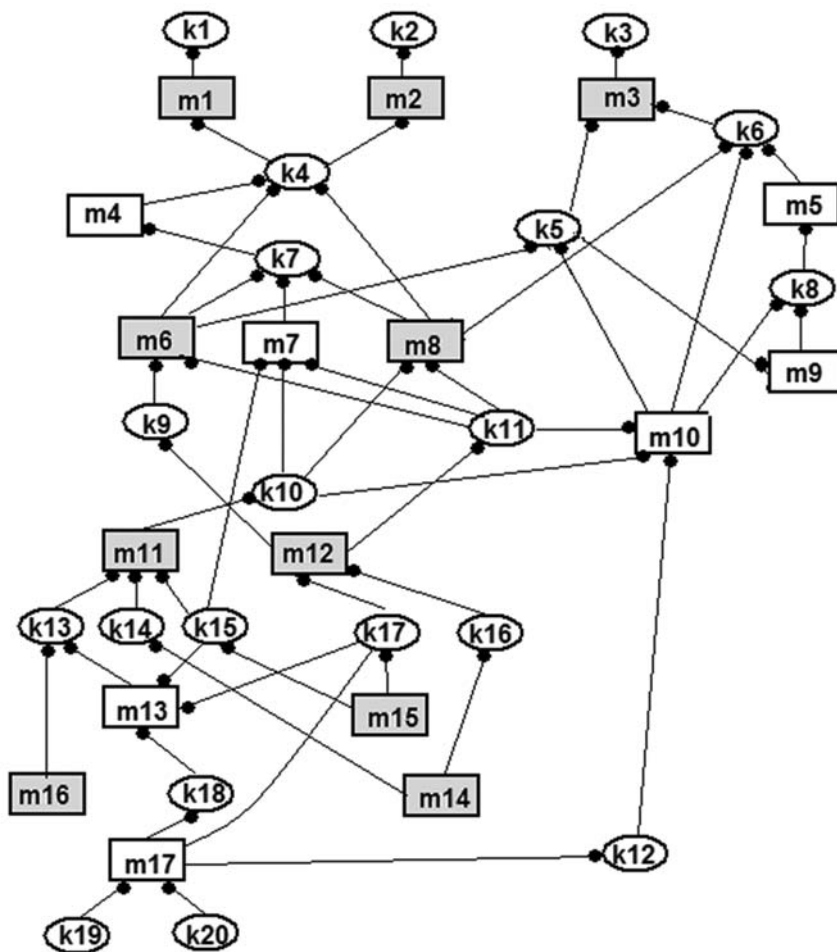


Рис. 1. Модель образовательного пространства (И-ИЛИ-граф G)

зъюнкции, и потому эти вершины названы вершинами типа ИЛИ, а вершинам модулей соответствует связка конъюнкции, и потому эти вершины – вершины типа И.

Пример модели приведен на рисунке 1, взятом из работы [4], где вершины концептов показаны в виде овалов, а вершины модулей – в виде прямоугольников.

Алгоритм синтеза маршрутов обучения

Синтез маршрутов обучения начинается с выбора множества целевых и исходных концептов. Этот выбор выполняет преподаватель, ориентируясь на индивидуальные особенности обучаемого и/или на учебную программу курса. При самообразовании задание цели исходит от самого обучаемого, которому нужно пособие для ее достижения.

В основе предлагаемого алгоритма лежит процедура «Индивидуальная траектория» выделения в модели образовательного пространства (И-ИЛИ-граф G) подграфа OM , включающего индивидуальную траекторию обучения.

Вычислительный процесс формирования траектории обучения осуществляется путем распространения волны на модели образовательного пространства. Распространение волны осуществляется путем последовательного перемещения фронта вершин Φ_i на графе G . Фронт формируется и перемещается пошагово. В начальном состоянии фронт Φ^k_0 включает множество целевых концептов. На первом шаге формируется фронт Φ^m_1 , в который включают модули, выходы которых покрывают множество концептов Φ^k_0 . Множество модулей Φ^m_1 связывается с множеством целевых концептов Φ_0 в соответствии со структурой связей графа G .

На втором шаге формируется фронт Φ^k_2 , в который включают множество концептов, покрывающих все входы модулей множества Φ^m_1 . Множество модулей Φ^m_1 связывается с множеством целевых концептов Φ^k_2 в соответствии со структурой связей графа G . В даль-

нейшем, на каждом нечетном шаге ($i = 3, 5, 7, \dots$) формируется фронт Φ^m_i , в который включаются модули, выходы которых покрывают множество концептов, входящих в фронт Φ^k_{i-1} . А на каждом четном шаге ($i = 4, 6, 8, \dots$) формируется фронт Φ^k_i , в который включаются концепты, выходы которых покрывают входы множества модулей, входящих в фронт Φ^m_{i-1} . Особенность реализации связей между элементами двух соседних фронтов Φ^k_{i-1} и Φ^m_i в том, что в соответствии с исходным графом G один и тот же концепт k_j может быть выходом у нескольких модулей. Возникает проблема выбора для включения в маршрут только одной связи между концептом k_j , входящем в состав фронта Φ^k_{i-1} и одним из модулей, входящих в состав фронта Φ^m_i . Именно этим и объясняется существование множества маршрутов обучения. Структура связей между элементами двух фронтов Φ^m_{i-1} и Φ^k_i остается неизменной в соответствии со структурой связей графа G . Этот процесс продолжается до тех пор, пока на входах модулей не будет других концептов, кроме исходных, или же их не будет.

Таким образом, построение конкретного маршрута обучения сводится к выбору альтернативных вариантов связей между фронтами при распространении волны.

Рассмотрим процесс построения одного из маршрутов на примере модели образовательного пространства, взятом из работы [4].

Вначале формируется фронт Φ^k_0 , включающий исходное множество целевых концептов $\Phi^k_0 = \{k_1, k_2, k_3\}$.

На первом шаге формируется фронт Φ^m_1 , в который включают модули, выходы которых покрывают множество концептов Φ^k_0 . $\Phi^m_1 = \{m_1, m_2, m_3\}$. m_1 связывается с $k_1, m_2 - c k_2, m_3 - c k_3$.

На втором шаге формируется фронт Φ^k_2 , в который включают концепты, выходы которых покрывают множество входов модулей Φ^m_1 . $\Phi^k_2 = \{k_4, k_5, k_6\}$. k_4 связывается с m_1 и m_2, k_5 и $k_6 - c m_3$.

На третьем шаге формируется фронт Φ^m_3 , в который включают модули, выходы которых покрыв-

вают множество концептов Φ^k_2 . $\Phi^m_3 = \{m_4, m_6, m_5\}$. m_4 связывается с $k_4, m_6 - c k_5, m_5 - c k_6$.

На четвертом шаге формируется фронт Φ^k_4 , в который включают концепты, выходы которых покрывают множество модулей Φ^m_3 . $\Phi^k_4 = \{k_7, k_9, k_{11}, k_8\}$. k_7 связывается с m_4, k_9 и $k_{11} - c m_6, k_8 - m_5$.

На пятом шаге формируется фронт Φ^m_5 , в который включают модули, выходы которых покрывают множество концептов Φ^k_4 . $\Phi^m_5 = \{m_6, m_{12}, m_9\}$. m_6 связывается с $k_7, m_{12} - c k_9$ и $k_{11}, m_9 - c k_8$.

На шестом шаге формируется фронт Φ^k_6 , в который включают концепты, выходы которых покрывают множество модулей Φ^m_5 . Отметим, что модуль m_6 уже вошел в Φ^m_3 и его входы уже покрыты, поэтому m_6 исключаем из Φ^m_5 . $\Phi^k_6 = \{k_{16}, k_{17}\}$. k_{16} и k_{17} связываются с m_{12} .

На седьмом шаге формируется фронт Φ^m_7 , в который включают модули, выходы которых покрывают множество концептов Φ^k_6 . $\Phi^m_7 = \{m_{14}, m_{15}\}$. m_{14} связывается с $k_{16}, m_{15} - c k_{17}$.

Поскольку на входы модулей Φ^m_7 ничего не поступает процесс формирования маршрута завершается.

На рис. 2 представлен синтезированный подграф И-ИЛИ-графа, содержащий пути от исходных концептов (или от концептов, не имеющих входов) к каждому целевому концепту (маршрут обучения). Концепты и модули, вошедшие в искомые пути, должны поясняться в пособии. Пояснения концептов содержатся в модулях. Для того, чтобы был изучен концепт k_i , нужно в маршруте обучения иметь хотя бы один модуль, поясняющий этот концепт. С другой стороны, чтобы модуль m_j с пояснением концепта k_i (пусть это модуль m_{ij}) был понятен обучаемому, обучаемый должен знать или предварительно изучить входные для модуля m_{ij} концепты. Очевидно, что в общем случае задача имеет множество решений. Среди них нужно выбрать одно оптимальное по одному из критериев, которыми могут быть те или иные функции метаданных модулей, вошедших в маршрут, например,

сложность освоения, современность материала, степень взаимосвязи с модулями других предметных областей и т.п. Например, если пользователя интересует наиболее короткое учебное пособие и длина пособия измеряется числом входящих в него модулей.

В работе для синтеза оптимального маршрута обучения предлагается роевой алгоритм, основанный на идеях адаптивного поведения муравьиной колонии [6, 7]. В отличие от канонической парадигмы муравьиного алгоритма, решением которого является маршрут минимальной стоимости в графе поиска решений, в новом роевом алгоритме решением является подграф с динамической структурой, выделенный в графе поиска решений [8]. Ключевым механизмом в новой парадигме роевого интеллекта, как и в алгоритме муравьиной колонии является **в первую очередь механизм непрямого обмена – стигмержи (stigmergy)**.

Представление оптимизационной задачи в виде **парадигмы роевого интеллекта** опирается на два ключевых момента: формирование графа поиска решений (ГПР) в виде *модели образовательного пространства* (И-ИЛИ-граф G) и формирование роем агентов на графе поиска решений допустимых альтернативных маршрутов обучения с помощью описанной выше волновой процедуры.

Сформируем множество W_ϕ всех ребер графа G , входящих во все вершины модули и множество W_a всех ребер графа G , входящих во все вершины концепты. Назовем множество W_ϕ множеством фиксированных связей, а множество W_a множеством альтернативных связей.

В общем случае поиск решения задачи осуществляется коллективно агентами $A = \{a_k | k = 1, 2, \dots, n_k\}$. Процесс поиска решений итерационный. Каждая итерация l включает три этапа. На первом этапе каждой итерации роевого алгоритма каждый агент a_k на *модели образовательного пространства* формирует маршрут обучения S_k . На втором этапе каждый агент a_k откладывает феромон на ребрах

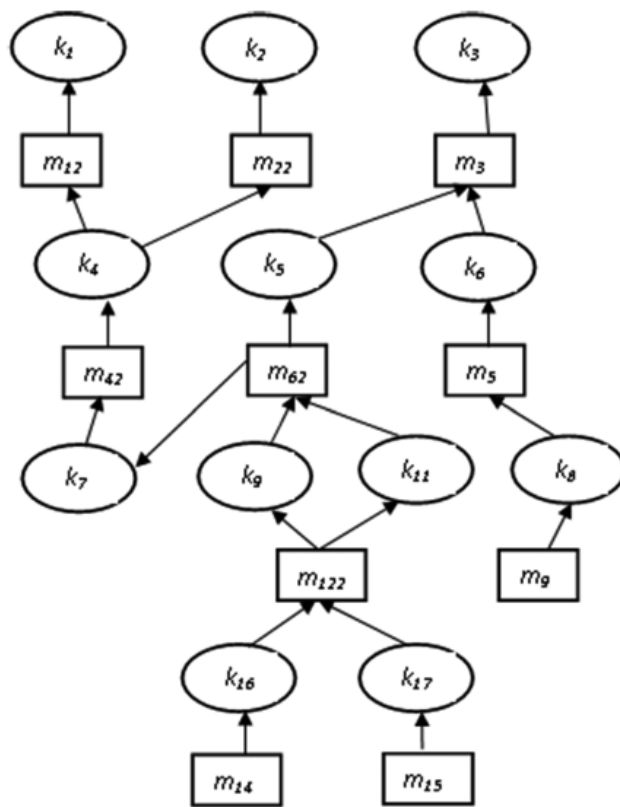


Рис. 2. Построенный маршрут обучения

графа G , входящих в состав построенного агентом маршрута обучения S_k . На третьем этапе осуществляется испарение феромона на ребрах графа G . В работе используется циклический (ant-cycle) метод муравьиных систем. В этом случае феромон откладывается агентом на ребрах графа G после полного формирования решения.

После построения ГПР G на всех его ребрах множества W_a откладывается начальное количество феромона Q/v , где $v = |U|$. В качестве исходных данных формируются множества целевых и исходных концептов. Целевые концепты включаются во фронт Φ^k_0 . Пошаговый процесс формирования траектории обучения осуществляется путем распространения волны на *модели образовательного пространства* начиная со стартового фронта Φ^k_0 . На шаге i распространения волны агент формирует фронт Φ_i и устанавливает связи между фронтом Φ_i и фронтом Φ_{i-1} . Возможны два случая.

В первом случае на шаге $i-1$ сформирован фронт модулей Φ^{m}_{i-1} , а на шаге i формируется фронт концептов Φ^k_i . В этом случае структу-

ра связей между Φ^{m}_{i-1} и Φ^k_i остается неизменной, то есть точно такой же как и в графе G , так как входы модулей, входящих во фронт Φ^{m}_{i-1} , представляют связку «И».

Во втором случае на шаге $i-1$ сформирован фронт концептов Φ^k_{i-1} , а на шаге i формируется фронт модулей Φ^m_i . Так как каждая вершина k_n (концепт) графа G представляет связку «ИЛИ», то есть все входы соответствуют одному и тому же концепту, то для избегания дублирования на входе вершины k_n остается одна связь. Выбор связи для каждого концепта выполняет агент. Пусть R_n – множество ребер u_{nm} входящих в вершину $k_n \in \Phi^k_{i-1}$. Для каждого ребра $u_{nm} \in R_n$ определяется параметр f_{nm} – суммарный уровень феромона на этом ребре. Вероятность P_{nm} выбора ребра u_{nm} определяется следующим соотношением

$$P_{nm} = f_{nm} / \sum_n(f_{nm}).$$

Агент с вероятностью P_{nm} выбирает одно из ребер, которое остается в структуре связей строящегося маршрута обучения S_k между Φ^k_{i-1} и Φ^m_i . Процесс выбора связей продолжается до тех пор, пока на

входах всех концептов фронта $\Phi^{k_{i-1}}$ (вершин графа G) останется только по одной связи.

В построенном маршруте S_k выделяется множество $W_{ak} \subset W_a$ всех ребер, входящих в вершины концепты.

На втором этапе итерации, каждый агент a_k откладывает феромон на ребрах графа G , принадлежащих множеству $W_{ak} \subset W_a$.

Количество феромона $\tau_k(l)$, откладываемое агентом a_k на каждом ребре $u_{nm} \in W_{ak}$ графа G , пропорционально базовому (опорному) количеству феромона Δ и определяется следующим образом

$$\tau_k(l) = \Delta / F_k(l),$$

где l -номер итерации, $F_k(l)$ – целевая функция для решения, соответствующего маршруту обучения S_k . Чем меньше $F_k(l)$, тем больше феромона откладывается в графе G на ребрах множества W_{ak} и, следовательно, тем больше вероятность выбора этих ребер при построении маршрута обучения на следующей итерации.

После того, как каждый агент сформировал решение (маршрут обучения) S_k и отложил феромон, на третьем этапе происходит общее испарение феромона на ребрах множества W_a графа G в соответствии с формулой

$$f_{nm} = f_{nm}(1 - \rho),$$

где ρ – коэффициент обновления. После выполнения всех действий на итерации находится агент с лучшим решением, которое запоминается. Далее осуществляется переход на следующую итерацию полного или частичного перебора маршрутов для выбора наилучшего из них.

Алгоритм поведения роя агентов

1. Формируется модель образовательного пространства (И-ИЛИ-граф G) $G = (M \cup K, U)$, где M – множество модулей, K – множество концептов.

2. Выбирается множество целевых $K_u \subset K$ и исходных $K_u \subset K$ концептов.

3. Задается: начальное количество феромона – Q ;

ρ – коэффициент обновления феромона на ребрах графа G ;

V – количество агентов в рое;

L – число итераций.

4. На множестве ребер W_a графа G , входящих в вершины (концепты) $k_j \in K$ откладывается начальное количество феромона Q/v , где $v = |U|$.

5. Формируется фронт Φ^k_0 , включающий множество целевых концептов.

6. $l = 1$. (l – номер итерации).

7. $k = 1$. (k – номер агента).

8. Выполняется «Алгоритм агента».

9. Если $k < V$, то $k = k + 1$ и переход к пункту 8, иначе переход к пункту 10.

12. Каждый агент a_k в графе G на ребрах множества $W_{ak} \subset W_a$ в построенного им маршруте обучения S_k откладывает феромон в количестве

$$\tau_k(l) = \Delta / F_k(l), \quad (4)$$

Общее количество феромона $h_i(l)$, отложенного на ребре $u_i \in D_k(l)$ после выполнения l итераций, определится как

$$h_i(l) = h_i(l-1) + \sum_{k|u_i \in D_k(l)} \tau_k(l)$$

13. После того, как каждый агент отложил феромон, происходит общее испарение феромона на ребрах графа $G = (M \cup C, U)$ в соответствии с нижеприведенной формулой.

$$h_i = h_i(1 - \rho), \quad (5)$$

где ρ – коэффициент обновления, h_i – суммарное количество феромона, отложенного муравьями на ребре $u_i \in U$ графа $G = (M \cup C, U)$,

14. Выбор лучшего решения, полученного на протяжении всех выполненных итераций.

15. Если все итерации выполнены, то конец работы алгоритма, в противном случае $l = l + 1$ и переход к пункту 8.

Алгоритм агента

1. $i = 1$ (i – номер шага распространения волны).

2. Фронт $\Phi^{k_{i-1}}$, включается в состав маршрута обучения S_k .

3. Формируется фронт модулей Φ^m_i , имеющих связи в графе G с концептами фронта $\Phi^{k_{i-1}}$. Фронт

модулей Φ^m_i включается в маршрут обучения S_k .

4. Выбирается очередная, начиная с первой, вершина концепт $k_n \in \Phi^{k_{i-1}}$.

5. На графе G определяется множество R_n ребер u_{nm} входящих в вершину $k_n \in \Phi^{k_{i-1}}$.

6. Для каждого ребра $u_{nm} \in R_n$ определяется параметр f_{nm} – суммарный уровень феромона на этом ребре в графе G .

7. Для каждого ребра $u_{nm} \in R_n$ рассчитывается вероятность P_{nm} его выбора

$$P_{nm} = f_{nm} / \sum_n(f_{nm}).$$

8. Агент a_k с вероятностью P_{nm} выбирает одно из ребер $u_{nm} \in R_n$, которое остается в структуре связей строящегося маршрута обучения S_k между $\Phi^{k_{i-1}}$ и Φ^m_i .

9. Если просмотрены все вершины $k_n \in \Phi^{k_{i-1}}$, то переход к пункту 10, иначе переход к пункту 4.

10. Если на входы модулей Φ^m_i ничего не поступает, то переход к пункту 15, иначе переход к пункту 11.

11. $i = i + 1$.

12. Формируется фронт вершин концептов Φ^k_i , имеющих связи в графе G вершинами фронта Φ^m_{i-1} . Фронт концептов Φ^k_i включается в маршрут обучения S_k .

13. Агент a_k фиксирует в маршруте обучения S_k все связи между фронтами Φ^m_{i-1} и Φ^k_i , содержащиеся в графе G .

14. $i = i + 1$. Переход к пункту 3.

15. Рассчитывается оценка $F_k(l)$ маршрута обучения S_k , построенного агентом a_k на итерации l .

16. Конец работы алгоритма.

Время работы этого алгоритма зависит от времени жизни колонии l (число итераций), количества вершин n графа G , числа агентов m , и определяется как (lnm) .

Заключение

Предложенный алгоритм позволяет генерировать варианты маршрутов обучения, состоящие из подмножества небольших модулей (разделяемых единиц контента) учебного материала. Рассмотрен метод оптимизации траектории обучения на основе онтологического подхода. Предлагается роевой

алгоритм синтеза оптимального маршрута обучения, основанный на моделях адаптивного поведения муравьиной колонии. Отличительной особенностью является наличие графического представления маршрута обучения, удобного для

анализа полученных результатов. Экспериментальные исследования проводились на ЭВМ типа IBM PC. Эксперименты показали, что для графов, содержащих до 1000 вершин, лучшие результаты решения были получены в среднем на 75

итерации роем из 100 агентов. Вероятность получения оптимального решения составила 0,9. Временная сложность алгоритма (BCA), для рассмотренных тестовых задач составляет ($BCA \approx O(n^2)$), где n – число вершин графа G .

Литература

1. Бова В.В. Концептуальная модель представления знаний при построении интеллектуальных информационных систем // Известия ЮФУ. Технические науки. – 2014. – № 7 (120). – С. 109–117.
2. Норенков И.П. Онтологические методы синтеза электронных учебных пособий // Научно-практический журнал «Открытое образование». – 2010. – № 6. – С. 39–44.
3. Кравченко Ю.А., Марков В.В. Онтологический подход формирования информационных ресурсов на основе разнородных источников знаний // Известия ЮФУ. Технические науки. – 2013. – № 7 (144). – С. 116–120.
4. Норенков И.П., Соколов Н.К. Синтез индивидуальных маршрутов обучения в онтологических обучающих системах // Информационные технологии, 2009, № 3, с. 74–77.
5. Тельнов Ю.Ф. Интеллектуальные информационные системы. – М.: Московский государственный университет экономики, статистики и информатики, 2003. – 26 с.
6. Лебедев О.Б. Модели адаптивного поведения муравьиной колонии в в задачах проектирования. Таганрог. Изд-во ЮФУ, 2013.
7. Лебедев Б.К., Лебедев В.Б. Оптимизация методом кристаллизации россыпи альтернатив // Известия ЮФУ. Изд-во ТТИ ЮФУ, 2013, №7. С. 11–17.
8. Лебедев Б.К., Лебедев О.Б. Моделирование адаптивного поведения муравьиной колонии при поиске решений, интерпретируемых деревьями // Известия ЮФУ. Изд-во ТТИ ЮФУ, 2012, №7. С. 27–35.