

# Теория

## СТЕПЕНЬ СООТВЕТСТВИЯ МЕЖДУ ТЕЗАУРУСОМ УЧАЩЕГОСЯ И ТЕЗАУРУСОМ УЧЕБНОГО КОНТЕНТА КАК МЕТРИКА ПРОЦЕССА УСВОЕНИЯ ДИСТАНЦИОННОГО УЧЕБНОГО КУРСА

**Александр Рыбанов,**

Волжский политехнический институт (филиал) ФГБОУ ВПО  
«Волгоградский государственный технический университет»  
rybalex@rambler.ru

**В основе мониторинга и контроля за процессом изучения дистанционного учебного курса находится решение задачи выставления обучаемому адекватной интегральной оценки усвоения всего учебного курса. Это делается по результатам тестирования. В качестве интегральной оценки степени усвоения дистанционного учебного курса в статье предлагается использовать степень соответствия между тезаурусом обучаемого и тезаурусом учебного курса.**

**В статье рассмотрены метрики сложности тезауруса учебного курса, построенные на основе теории графов и теории информации. В качестве метрики сложности тезауруса учебного курса предлагается использовать количество информации, содержащейся в его графе.**

*Ключевые слова:* дистанционное образование, thesaurus, система дистанционного обучения, метрики тезауруса, измерение знаний, контент учебного курса, тестирование знаний.

## 1

*Bechhofer, S.,  
Goble C.*  
Thesaurus construction  
through knowledge repre-  
sentation. *Data &  
Knowledge Engineering*,  
2001. Vol. 37, № 1,  
p. 25–45.

## 2

*Soergel D.,  
Lauser B.,  
Liang A.,  
Fisseha F.,  
Keizer J.,  
Katz S.*  
Reengineering thesauri  
for new applications: the  
AGROVOC example.  
*Journal of Digital  
Information*, 2004. Vol. 4,  
available at  
[ftp://ftp.fao.org/docrep/  
fao/008/af234e/  
af234e00.pdf](http://ftp.fao.org/docrep/fao/008/af234e/af234e00.pdf).

## 3

*Abdulmanova I.R.*  
Formation of professional  
personal thesaurus as a  
goal of the professional  
education. *Tomsk State  
Pedagogical University  
Bulletin*, 2010, № 2,  
p. 36–39.

## Введение

Процесс обучения предполагает целенаправленное воздействие на тезаурус обучаемого. В настоящее время в системах дистанционного обучения отсутствует возможность мониторинга знаний, основанного на степени соответствия между тезаурусом обучаемого и тезаурусом учебного курса. В системе дистанционного обучения процесс изучения учебного курса состоит из последовательности тактов предъявления обучаемому и усвоения им содержания предмета (контента). Такт усвоения обучаемым определённого учебного контента приводит к расширению его тезауруса.

Процесс формирования тезауруса на основе применения методов представления знаний подробно изложен в работе (Bechhofer S., Goble C., 2001)<sup>1</sup>. Его особенностью является понимание не только признаков объекта тезауруса, но и его связей с другими объектами (Soerge I.D., Lauser B., Liang A., Fisseha F., Keizer J., Katz S., 2004)<sup>2</sup>.

## Понятия

Определение понятия «профессиональный тезаурус личности» приведено в работе (Abdulmanova I.R., 2010)<sup>3</sup>. Те-

заурус учебного курса — это совокупность объектов учебного курса (понятий, законов, теорем, утверждений и т.п.) с указанием связей между ними.

Тезаурус обучаемого рассматривается как объект измерения учебного контента, усвоенного на семантическом уровне. Оценивается количеством информации, содержащейся в его графе, с учётом коэффициентов усвоения объектов тезауруса.

Пусть система дистанционного обучения по учебному курсу, описываемому тезаурусом  $I_d$ , предъявляет учебный контент обучаемому, обладающему тезаурусом  $I_s$ . Возможность обучаемого воспринять учебный контент, описываемый с тезаурусом  $I_f \subseteq I_d$ , определяется следующими вариантами:

1) Если  $I_f \subseteq I_d$ , то в процессе обучения никаких изменений в тезаурусе обучаемого не произойдет, т.к. предъявляется уже известная ему информация.

2) Если  $I_f \cap I_s \neq \emptyset$  и  $I_s \subset I_f$ , то учебный контент при желании может быть воспринят обучаемым, и как следствие, его тезаурус будет расширен.

Максимальное количество семантической информации обучаемый приобретает при согласовании его тезауруса с тезаурусом контента учебного курса, т.е. если учебный кон-

тент понятен обучаемому и не сёт сведения, отсутствующие в его тезаурусе.

Тезаурусное представление учебного контента, а также текущего состояния знаний обучаемого обеспечивает адаптивный подбор и упорядочение учебной информации.

## Проблема

Для количественной оценки сложности тезауруса, представленного в виде графа, могут быть использованы метрики, описанные в работах Bonchev D., Buck G.A.<sup>4</sup>, а также Gangemi A., Catenacci C., Ciaramita M., Lehmann J.<sup>5</sup>. Для сравнительного анализа тезаурусов можно использовать метрики, применяемые для онтологий, т.к. тезаурусы можно рассматривать как типы онтологий. Но для сравнительного анализа тезауруса учебного курса и тезауруса обучаемого, метрики сравнения онтологий, описанные в работах (Lozano-Tello A., Gomez-Perez A., 2004)<sup>6</sup> и (Maedche A., Staab S., 2002)<sup>7</sup> требуют доработки, т.к. результатом сравнения тезауруса обучаемого и тезауруса учебного курса должна быть оценка, характеризующая не только соответствие их структуры, но и степень полноты усвоения учебного курса.

В системах дистанционного обучения степень усвоения

учебного курса оценивается по результатам тестирования обучаемых (Murgick J., 2010)<sup>8</sup>. В настоящее время большое внимание уделяется повышению точности оценки результатов обучения в системах дистанционного обучения. В работе (Рубанов А.А., 2013)<sup>9</sup> для этого предлагается учитывать процесс формирования пользователем конечного ответа на тестовые задания, а в работе (Scalise K., Gifford B., 2006)<sup>10</sup> предлагаются инновационные формы тестовых заданий для компьютерного тестирования знаний.

Интегральная оценка качества усвоения дистанционного учебного курса рассчитывается на основе оценок, полученных обучаемым в результате прохождения всех тестов учебного курса. Например, в системе Moodle предусмотрены следующие подходы к вычислению интегральной оценки качества усвоения учебного курса (Nash S.S., Rice W., 2010)<sup>11</sup>: среднее оценок, взвешенное среднее оценок, простое взвешенное среднее оценок, среднее оценок (с учётом дополнительных баллов), медиана оценок, низшая оценка, высшая оценка, мода оценок, сумма оценок.

Из всех подходов, только «взвешенное значение оценок» учитывает сложность усвоения учебного модуля путем определения весового коэффициента

## Теория

4

*Bonchev D., Buck, G.A.*  
Quantitative Measures of Network Complexity. In Complexity in Chemistry, Biology, and Ecology, D. Bonchev and D.H. Rouvray, Eds., Springer, New York, 2005, p. 191–235.

5

*Gangemi A., Catenacci C., Ciaramita M., Lehmann, J.* (2005). Ontology evaluation and validation. An integrated formal model for the quality diagnostic task. Technical Report, available at [http://www.loa-cnr.it/Files/OntoEval4OntoDev\\_Final.pdf](http://www.loa-cnr.it/Files/OntoEval4OntoDev_Final.pdf).

6

*Lozano-Tello A., Gomez-Perez A.*  
Ontometric: A method to choose the appropriate ontology. Database Management, 2004, 15(2), p. 1–18.

7

*Maedche A., Staab S.*  
Measuring similarity between ontologies. Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, 13th International Conference, EKAW 2002. Sigüenza, Spain, Vol. 2473, p. 251–263.

**ПЕД**  
**измерения**

8

*Myrick J.*  
Moodle 1.9 Testing and  
Assessment. Birmingham:  
Packt Publishing Ltd.,  
2010.

9

*Rybanov A.A.*  
Set of criteria for efficiency of the process forming the answers to multiple-choice test items. Turkish Online Journal of Distance Education. 2013. Vol. 14, No. 1, 75–84.

10

*Scalise K.,  
Gifford B.*  
Computer-Based Assessment in E-Learning: A Framework for Constructing «Intermediate Constraint» Questions and Tasks for Technology Platforms. Journal of Technology, Learning, and Assessment, 2006, Vol. 4(6). Retrieved [date] from <http://www.jtla.org>.

11

*Nash S.S.,  
Rice W.*  
Moodle 1.9 Teaching Techniques. Birmingham: Packt Publishing Ltd, 2010.

для теста, ассоциированного с этим модулем. Возникает проблема выбора весовых коэффициентов учебных модулей в рамках дистанционного учебного курса. Определение коэффициентов методом субъективного взвешивания, т.е. автором дистанционного учебного курса, приводит к погрешности в значении итоговой оценки.

Тезаурусное представление учебного курса позволит более обоснованно и объективно определять вес каждого объекта тезауруса. Сопоставлением тестовых заданий с объектами тезауруса учебного курса можно определять веса тестов.

На основе тезаурусного представления учебного контента и сравнительного анализа результатов тестирования обучаемых могут быть идентифицированы трудные в усвоении объекты тезауруса. Множество таких объектов тезауруса может быть использовано для более обоснованной стратегии корректировки учебного контента и тестов.

Используя в качестве метрики учебных достижений оценку степени соответствия между тезаурусом обучаемого и тезаурусом учебного курса можно построить более точные кривые обучаемости (рис. 1). Кривые обучаемости лежат в основе классификации обучаемых на экстравертов и интровертов: интровертированные субъекты

имеют вогнутую кривую обучаемости, обусловленную длительной фазой скрытого накопления знаний и навыков.

Основу всех выше указанных направлений мониторинга и контроля за процессом изучения дистанционного учебного курса составляет решение задачи выставления обучаемому адекватной интегральной оценки усвоения всего учебного курса по результатам тестирования. Данная задача может быть решена путем измерения степени соответствия между тезаурусом обучаемого и тезаурусом учебного курса.

## Модель тезауруса дистанционного учебного курса

*Тезаурус, описывающий систему объектов учебного курса,* представим в виде ориентированного графа  $G = (V, E)$ , где  $V$  — множество вершин (объектов тезауруса учебного курса),  $E$  — множество дуг (ориентированных ребер, описывающих логику изучения объектов учебного курса). Введем следующие обозначения:  $n = |V|$ ,  $m = |E|$ .

Рассмотрим множество логических отношений между объектами тезауруса учебного курса. Пусть  $(v_i, v_j) \in E$ , если  $v_i$  является непосредственной семантической составляющей  $v_j$ . Пусть  $A$  — матрица смежности

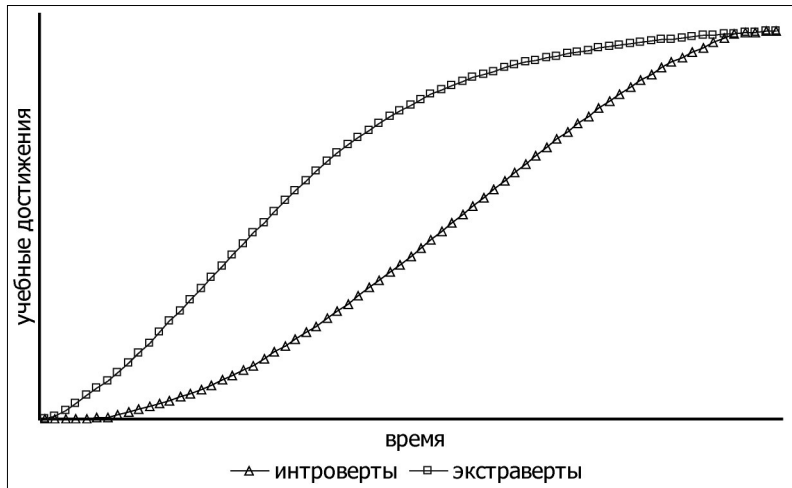


Рис. 1. Динамика изменения учебных достижений в процессе обучения

Теория

графа тезауруса учебного курса  $G$ , где элемент матрицы  $a_{ij} = 1$ , если  $(v_i, v_j) \in E$ , и  $a_{ij} = 0$ , в противном случае. Тогда  $A^L$  — матрица, показывающая, сколько путей длиной в  $L$  тактов имеется между двумя любыми объектами  $v_i$  и  $v_j$ . Количество этих путей определяется цифрой на пересечении  $i$ -й строки и  $j$ -го столбца матрицы  $A^L$ . Обозначим элемент матрицы  $A^L$  как  $a_{ij}^{(L)}$ , тогда:

$$a_{ij}^{(L=1)} = a_{ij}, \quad (1)$$

$$a_{ij}^{(L+1)} = \sum_{ik} a_{ik}^{(L)} a_{kj}. \quad (2)$$

Граф  $G$ , описывающий тезаурус учебного курса, должен удовлетворять следующим правилам:

1) В графе тезауруса учебного курса не должно быть изолированных вершин:

$$\sum_{i=j}^n a_{ik}^{(L=1)} + \sum_{j=1}^n a_{kj}^{(L=1)} \neq 0, \quad \forall k = \overline{1, n}. \quad (3)$$

2) В графе тезауруса учебного курса не должно быть контуров, т.е. для любой матрицы  $A^L$  выполняется условие:

$$\sum_{k=1}^n a_{kk}^{(L)} \neq 1. \quad (4)$$

3) В графе тезауруса учебного курса не должно быть дублирующих связей между вершинами, т.е. если существуют дуги  $(v_i, v_j)$ ,  $(v_j, v_k)$ , и  $(v_i, v_k)$ , то дуга  $(v_i, v_k)$  может быть удалена, так как она в соответствии со свойством транзитивности дублирует требования к последовательности изучения объектов тезауруса  $v_i$  и  $v_k$ .

Входными объектами тезауруса учебного курса назовём все объекты  $v_k$ , для которых выполняется условие:

$$\sum_{i=1}^n a_{ik}^{(L=1)} = 0. \quad (5)$$

Выходными объектами тезауруса учебного курса назовём все объекты  $v_k$ , для которых выполняется условие:

$$\sum_{j=1}^n a_{kj}^{(L=1)} = 0. \quad (6)$$

При анализе тезауруса предметной области важно знать, какие объекты используются для формирования других объектов, и каких именно. Для описания относительных длительностей формирования объектов тезауруса учебного курса используется матрица достижимости  $D$ :

$$D = \sum_{L=1}^N A^L, \quad (7)$$

где  $d_{ij}$  — элемент матрицы  $D$ , который показывает — через какое количество тактов после объекта  $v_i$  будет сформирован объект  $v_j$ ;  $N$  — порядок графа тезауруса учебного курса:  $A^N \neq 0$ ,  $A^{N+1} = 0$ .

### Метрики сложности тезауруса дистанционного учебного курса на основе теории графов

Для описания *характеристик тезауруса учебного курса*, представленного в виде графа  $G$ , могут приме-

няться следующие метрики графов:

1) Порядок графа тезауруса учебного курса:  $n(G) = n$ .

2) Размер графа тезауруса учебного курса:  $s(G) = m$ .

3) Диаметр тезауруса учебного курса — длина максимального пути между входным  $v_i$  и выходным  $v_j$  объектами тезауруса, выраженная числом дуг, составляющих этот путь:

$$\text{diam}(G) = \max_{d_y \in D} d_{ij}. \quad (8)$$

4) Структурная избыточность  $R(G)$  графа тезауруса учебного курса показывает превышение общего количества связей между вершинами графа  $G$  над минимальным количеством связей:

$$R(G) = \frac{m}{n-1} - 1. \quad (9)$$

5) Реберная плотность  $Q(G)$  — характеризует близость графа  $G$  к полносвязному графу:

$$Q(G) = \frac{2m}{n(n-1)}. \quad (10)$$

6) Абсолютная глубина графа  $H(G)$ :

$$H'(G) = \sum_j^{|P|} N_{j \in P}, \quad (11)$$

где — длина  $j$ -го пути, принадлежащего множеству всех путей  $P$  в графе  $G$ .

7) Средняя глубина графа  $h(G)$ :

$$h(G) = \frac{1}{|P|} \sum_j^{|P|} N_{j \in P}. \quad (12)$$

Количественные характеристики объектов тезауруса учебного курса могут быть описаны следующими метриками:

1) Вес объекта тезауруса учебного курса, ассоциированного с вершиной  $v_k$ , определим как количество всех путей, проходящих через вершину  $v_k$ :

$$w_k = \sum_{i=1}^n d_{ik} + \sum_{j=1}^n d_{jk}, \quad (13)$$

где  $d_{ij}$  — элемент матрицы достижимости  $D$ , которая показывает сколько путей, независимо от их длины, существует между вершинами  $v_i$  и  $v_j$ .

2) Ранг объекта  $v_j$  тезауруса учебного курса — это число, равное количеству дуг, входящих в путь максимальной длины на графе  $G$ , от входного объекта тезауруса учебного курса до объекта  $v_j$ :

$$P_j = L \text{ при } \sum_{i=1}^n a_{ij}^{(L)} > 0 \text{ и } \sum_{i=1}^n a_{ij}^{(L+1)} = 0. \quad (14)$$

Когда ранги всех объектов учебного курса определены, можно построить граф тезауруса учебного курса, упорядоченный по тактам.

3) Степень объекта тезауруса учебного курса определяется как сумма полустепеней исхода и захода вершины  $v_k$ , ассоциированной с объектом тезауруса:

$$a_k = \sum_{j=1}^n a_{kj} + \sum_{i=1}^n a_{ik}. \quad (15)$$

Представленные выше метрики позволяют оценить топологическую сложность графа тезауруса учебного курса и дают наглядное представление о степени сложности освоения дистанционного учебного курса.

### Метрика сложности тезауруса дистанционного учебного курса на основе теории информации

Метрику сложности тезауруса учебного курса опишем на основе теории информации Шеннона (Shannon C.E, Weaver W., 1949)<sup>12</sup>. В соответствии с теорией информации, информационная энтропия  $H(a)$  сообщения из символов, разделенных согласно некоторому критерию на  $k$  групп из  $N_1, N_2, \dots, N_k$ , символов, рассчитывается по следующей формуле:

$$\begin{aligned} H(\alpha) &= - \sum_{i=1}^k p_i \log_2 p_i = \\ &= - \sum_{i=1}^k \frac{N_i}{N} \log_2 \frac{N_i}{N}, \end{aligned} \quad (16)$$

где  $p_i = \frac{N_i}{N}$  — вероятность появления символов  $i$ -й группы в сообщении.

Граф тезауруса учебного курса задан конечным набором

<sup>12</sup> Shannon C.E, Weaver W. (1949). The Mathematical Theory of Communication. The University of Illinois Press, Urbana, Illinois.

элементов (вершин, рёбер, дуг, клик и т.д.). Пусть  $N$  — количество элементов графа тезауруса учебного курса. Каждому элементу графа тезауруса учебного курса соответствует его вес  $w_i, i = \overline{1, N}$ .

Общий вес графа тезауруса учебного курса определим следующим выражением:

$$W = \sum_{i=1}^N w_i. \quad (17)$$

Вероятность появления в графе тезауруса учебного курса  $i$ -го элемента с весом  $w_i$  рассчитывается следующим образом:

$$p_i = \frac{w_i}{W}. \quad (18)$$

Таким образом, вероятностная схема графа тезауруса учебного курса может быть описана табл. 1.

Таблица 1

**Вероятностная схема графа тезауруса учебного курса**

Элемент	1	2	...	$N$
Вес	$w_1$	$w_2$	...	$w_N$
Вероятность	$p_1$	$p_2$	...	$p_N$

Энтропия графа тезауруса учебного курса с общим весом  $W$  и весами элементов  $w_i, i = \overline{1, N}$ , для заданной вероятностной схемы (табл. 1) определяется следующим выражением:

$$H = - \sum_{i=1}^N \frac{w_i}{W} \log_2 w_i + \sum_{i=1}^N \frac{w_i}{W} \log_2 W = \log_2 W - \quad (19)$$

$$- \frac{1}{W} \sum_{i=1}^N w_i \log_2 w_i.$$

Согласно теории информации Шеннона, количество информации определяется как уменьшение энтропии системы относительно максимальной энтропии, которая может существовать в системе с таким же числом элементов:

$$I = H_{max} - H. \quad (20)$$

Информационная энтропия графа тезауруса учебного курса принимает максимальное значение при  $w_i = 1$  (формула 19), и определяется как:

$$H_{max} = \log_2 W. \quad (21)$$

Таким образом, выражение для определения количества информации, содержащейся в графе тезауруса учебного курса принимает вид:

$$I = \frac{1}{W} \sum_{i=1}^N w_i \log_2 w_i. \quad (22)$$

Полученное выражение является метрикой сложности тезауруса учебного курса и может быть использовано для оценки степени соответствия между графом тезауруса обучаемого и графом тезауруса учебного курса.

## Модель тезауруса обучаемого

Графом тезауруса обучаемого  $G' = (U, E')$  назовём подграф,



порождённый множеством вершин графа тезауруса учебного курса  $G = (V, E)$ , где  $U \subseteq V$ , и  $E'$  состоит из всех тех дуг графа  $G$ , у которых оба конца принадлежат  $U$ . Каждая вершина графа  $G'$  ассоциирована с изученным объектом дистанционного учебного курса. В качестве количественной характеристики, описывается степень усвоения  $\lambda_k \in [0;1]$  учебного материала, связанного с понятием  $u_k$ .

*Динамику процесса изучения учебного контента*, описываемого тезаурусом  $G = (V, E)$ , представим конечной упорядоченной последовательностью графов тезауруса учебного:

$$\rho = \{G'_1, \dots, G'_i, \dots, G'_r\}$$

$$G'_i \cap G'_{i+1} = G'_i, \quad i = \overline{1, r-1},$$

$$G'_r \cap G = G'_r,$$

$$\Delta_{i+1} = n(G'_{i+1}) - n(G'_i),$$

где  $G'_i$  — подграф порождённый множеством вершин графа  $G'_{i+1}$ ,  $U_i \subset U_{i+1}$ ;  $G'_r$  — подграф, порождённый множеством вершин графа  $G$ ,  $U_r \subseteq V$ ;  $\Delta_{i+1}$  — количество новых объектов, которыми был расширен граф тезауруса учебного  $G'_i$ .

Множество  $\rho$  описывает процесс изменения тезауруса учебного, связанный с усвоением новых объектов тезауруса учебного курса. В процессе

изучения учебного курса происходит расширение понятийной базы тезауруса учебного, которое приводит к увеличению отношений между понятиями.

Вес объекта в тезаурусе учебного, ассоциированного с вершиной  $u_k$ , определим как произведение степени усвоения  $\lambda_k \in [0;1]$  учебным объектом тезауруса учебного курса на вес этого объекта в графе тезауруса учебного курса:

$$w'_k = \lambda_k w_k. \quad (23)$$

В качестве метрик сложности тезауруса учебного можно использовать метрики тезауруса учебного курса.

Степень соответствия  $\delta(G'_i)$  между графом тезауруса учебного  $G'$  и графом тезауруса учебного курса может быть определена как:

$$\delta(G'_i) = \frac{I'(G'_i)}{I(G)} \cdot 100\%, \quad (24)$$

где  $I'(G'_i)$  — количество информации в графе тезауруса учебного  $G'$ , которое рассчитывается по формуле 22.

## Результаты моделирования

Для анализа предлагаемых в работе метрик (формулы: 8–12, 22, 24) был проведён эксперимент, в рамках которого моде-

ПЕД  
измерения

лировался процесс изучения учебного контента. Граф тезауруса учебного контента, представленный на рис. 2, состоит из 50 объектов.

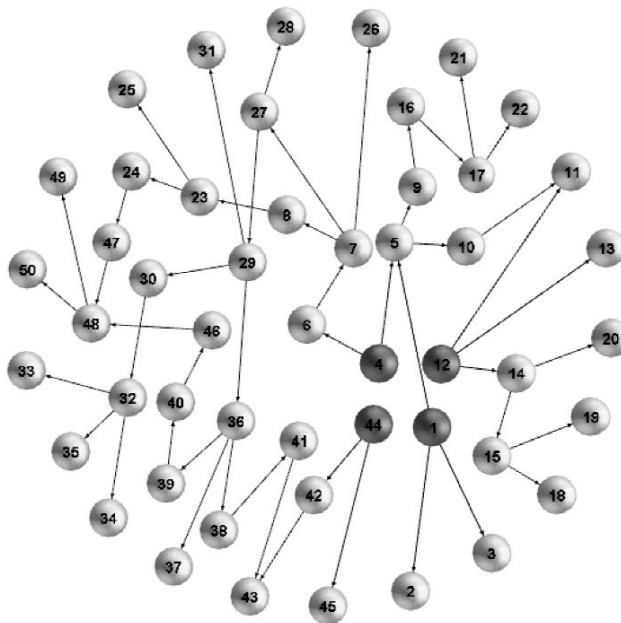


Рис. 2. Граф тезауруса учебного контента  $G$

Входными объектами тезауруса учебного контента являются объекты №1, №4, №12, №44. Граф тезауруса учебного контента имеет следующие значения метрик:  $diam(G) = 8$ ,  $n(G) = 50$ ,  $R(G) = 0,020$ ,  $Q(G) = 0,041$ ,  $H(G) = 809$ ,  $h(G) = 3,487$ ,  $I(G) = 1796,002$ .

Цель эксперимента заключалась в том, чтобы проследить изменение метрических характеристик графа тезауруса обучаемого, формируемого в процессе изучения учебного контента.

Исходные данные для проведения эксперимента:

1) Динамика процесса изучения учебного контента, представленного графом тезауруса  $G$ , описывается следующей последовательностью изменения графа тезауруса обучаемого (см. приложение):

$$\rho = \{G'_1, G'_2, G'_3, G'_4, G'_5, G'_6, G'_7, G'_8, G'_9, G'_{10}\}.$$

Метрики графов тезауруса обучаемого приведены в табл. 2.

Таблица 2

Теория

Метрики графов тезауруса обучаемого

Граф тезауруса обучаемого	$diam(G_i)$	$n(G_i)$	$R(G_i)$	$Q(G_i)$	$H'(G_i)$	$h(G_i)$
$G_{i1}$	1	5	0	.400	4	1.000
$G_{i2}$	3	10	0	.200	26	1.625
$G_{i3}$	3	15	0	.133	39	1.560
$G_{i4}$	4	20	0	.100	77	1.833
$G_{i5}$	5	25	0	.080	157	2.309
$G_{i6}$	5	30	0	.067	204	2.345
$G_{i7}$	7	35	0	.057	324	2.723
$G_{i8}$	7	40	0	.050	430	2.886
$G_{i9}$	8	45	0	.044	499	2.970
$G_{i10}$	8	50	.020	.041	809	3.487

Метрики (формулы 8–12), основанные на теории графов, не учитывают степень усвоения обучаемым объектов тезауруса учебного контента и, следовательно, будут иметь одинаковые значения для идентичных по топологии графов тезауруса различных обучаемых. Кроме того, процесс расширения понятийной базы тезауруса обучаемого не всегда означает изменение значений метрик  $diam(G_i)$ ,  $R(G_i)$  и увеличение метрики  $h(G_i)$  (например, графы тезауруса обучаемого  $G_2$  и  $G_3$ ). Метрики  $n(G_i)$ ,  $Q(G_i)$  не учитывают веса объектов тезауруса учебного контента. Как

следствие, степень соответствия данных метрик для тезауруса обучаемого и тезауруса учебного контента линейно зависит от количества объектов, которыми расширена понятийная база тезауруса обучаемого. Метрика  $H'(G_i)$  учитывает веса объектов тезауруса учебного контента и может быть использована только для оценки сложности тезауруса учебного курса.

2) Степень усвоения  $\lambda_k$  учебного материала, связанного с понятием  $u_k$  графа тезауруса  $G$  учебного курса:  $\lambda_k \in [0; 1]$ ,  $\forall k = 1, n$ . Величина  $\lambda_k$  описывается следующими категориями усвоения понятия: неудовле-

ПЕД  
измерения

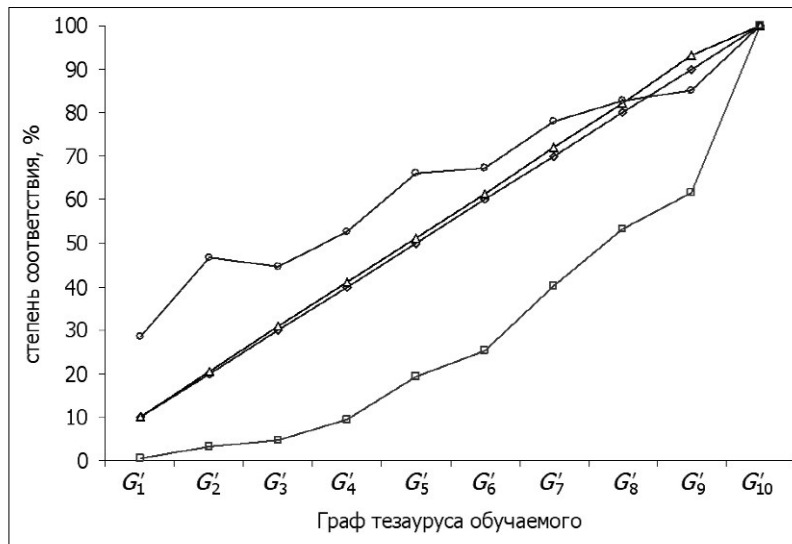


Рис. 3. Динамика изменения степени соответствия метрик  $n(G'_i)$ ,  $Q(G'_i)$ ,  $H'(G'_i)$ ,  $h(G'_i)$  аналогичным метрикам графа тезауруса учебного контента

творительно —  $[0; .61]$ ; удовлетворительно —  $[.61; .76]$ ; хорошо —  $[.76; .90]$ ; отлично —  $[.90; 1]$ . Поскольку в случае неудовлетворительного освоения понятия учебного курса, все понятия, выходящие из него, не могут быть усвоены, величина  $\lambda_k$  изменялась в интервале  $\lambda_k \in [0.61; 1]$ .

3) Расширение понятийной базы тезауруса обучаемого осуществляется с шагом  $\Delta = 5$  понятий.

Результаты изменения метрики  $I(G'_i)$ , приведены в табл. 3, а степени  $\delta(G'_i)$  соответствия в табл. 4.

Сравнительный анализ экспериментальных данных (рис. 4) показал, что:

1) с возрастанием степени усвоения объектов тезауруса учебного контента увеличивается степень соответствия  $\delta(G'_i)$  между графом тезауруса обучаемого  $G'_i$  и графом тезауруса учебного контента  $G$ ;

2) увеличение сложности топологии графа  $G'_i$ , обусловленное расширением понятийной базы тезауруса обучаемого, приводит к увеличению степени соответствия  $\delta(G'_i)$  между графом тезауруса обучаемого  $G'_i$  и графом тезауруса учебного контента  $G$ ;

3) чем больше вес объектов тезауруса учебного курса, усвоенных обучаемым, тем больше приращение  $\Delta\delta$ .

Таблица 3

Теория

Значения метрики  $I(G'_i)$ , бит

Граф тезауруса обучаемого	$\lambda_k$					
	[.61;.76)	[.61;.90)	[.76;.90)	[.76;1]	[.90;1]	1
$G'_{i1}$	1.675	2.208	3.103	5.009	5.917	6.755
$G'_{i2}$	28.339	33.437	38.247	47.923	52.808	58.894
$G'_{i3}$	46.415	57.029	69.461	79.970	91.289	97.400
$G'_{i4}$	127.073	166.283	167.941	191.029	220.562	240.865
$G'_{i5}$	265.555	276.504	349.899	379.486	425.265	465.161
$G'_{i6}$	337.307	389.449	443.160	481.549	540.448	589.141
$G'_{i7}$	459.584	551.091	637.198	725.756	767.073	830.528
$G'_{i8}$	610.144	690.862	798.965	922.250	996.544	1068.03
$G'_{i9}$	677.172	809.849	922.813	1001.127	1121.270	1208.266
$G'_{i10}$	1005.884	1221.713	1353.899	1457.382	1667.390	1796.002

Таблица 4

Значения степени соответствия  $\delta(G'_i)$ , %

Граф тезауруса обучаемого	$\lambda_k$					
	[.61;.76)	[.61;.90)	[.76;.90)	[.76;1]	[.90;1]	1
$G'_{i1}$	.093	.123	.173	.279	.329	.376
$G'_{i2}$	1.578	1.862	2.130	2.668	2.940	3.279
$G'_{i3}$	2.584	3.175	3.868	4.453	5.083	5.423
$G'_{i4}$	7.075	9.258	9.351	10.636	12.281	13.411
$G'_{i5}$	14.786	15.396	19.482	21.129	23.678	25.900
$G'_{i6}$	18.781	21.684	24.675	26.812	30.092	32.803
$G'_{i7}$	25.589	30.684	35.479	40.410	42.710	46.243
$G'_{i8}$	33.972	38.467	44.486	51.530	55.487	59.467
$G'_{i9}$	37.704	45.092	51.382	55.742	62.431	67.275
$G'_{i10}$	56.007	68.024	75.384	81.146	92.839	100

ПЕД  
измерения

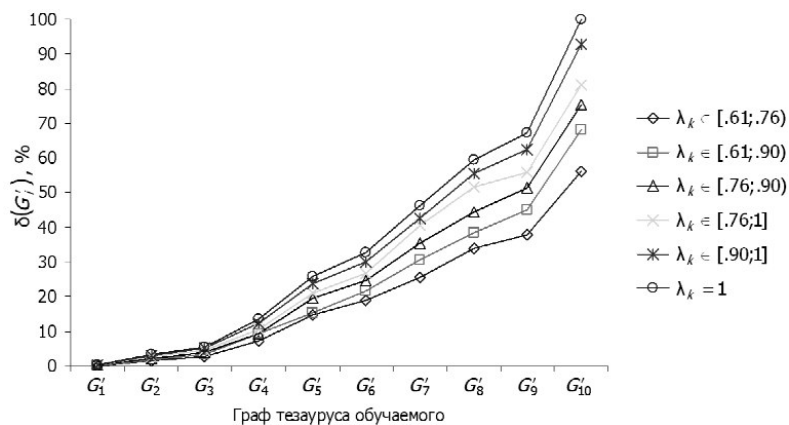


Рис. 4. Динамика изменения  $d(G_i)$  в процессе расширения тезауруса обучаемого

Данные факты позволяют сделать вывод, что метрика  $\delta(G_i)$ , построенная в соответствии с теорией информации, является объективной оценкой степени усвоения дистанционного учебного курса.

### Заключение

Сравнительный анализ метрик графа тезауруса учебного контента показал, что для оценки сложности дистанционных учебных курсов может быть использована метрика

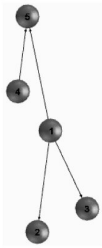
$H'(G)$  (формула 11) или  $I(G)$  (формула 22). Для оценки степени соответствия между графом тезауруса учебного контента и графом тезауруса обучаемого рекомендуется использовать метрику  $\delta(G_i)$  (формула 24), построенную на основе меры количества семантической информации, т.к. она учитывает степень усвоения обучаемым объектов тезауруса учебного контента. Предлагаемая метрика может быть использована для мониторинга за процессом дистанционного обучения.



Приложение

Теория

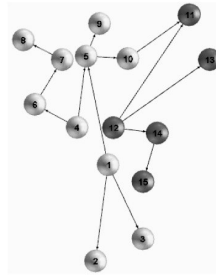
Динамика процесса изучения учебного контента



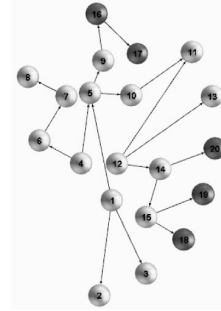
$G_1'$



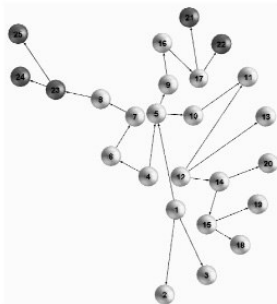
$G_2'$



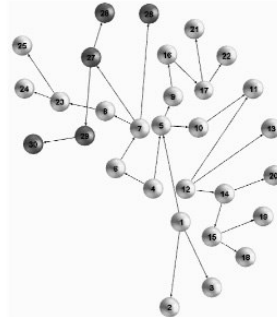
$G_3'$



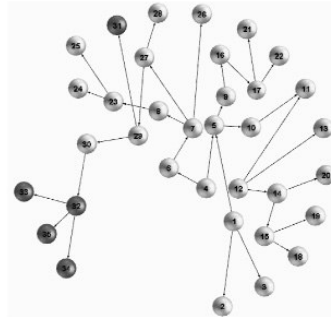
$G_4'$



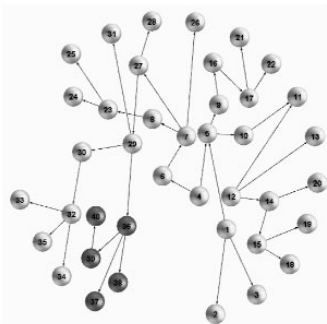
$G_5'$



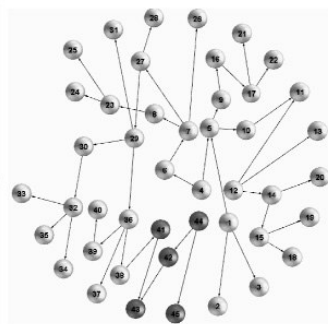
$G_6'$



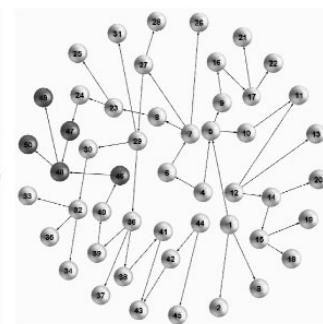
$G_7'$



$G_8'$



$G_9'$



$G_{10}'$

