УДК 004.82
ВАК 05.25.05
РИНЦ 14.35.09

Ю.А. Кравченко

Управление знаниями как одно из направлений развития технологий открытого образования[1]

В статье рассмотрена проблема создания сценариев управления знаниями, как принципиальной организации информационных процессов в технологиях открытого образования, а также принципов их проектирования и эволюции. Решение проблемы информационного переполнения в интеллектуальных системах обучения описано с позиции перехода от хранения и обработки данных к накоплению и обработке знаний. Увеличение вероятности продуцирования нового знания в процессах обучения предполагается за счет использования междисциплинарного и мультидисциплинарного подходов к накоплению информации. Предложена нечеткая прогнозная модель оценки уровня знаний обучаемого с учетом естественной потери некоторой части знаний в течение длительного промежутка времени.

Ключевые слова: *интеллектуальные системы обучения, информационные процессы, сценарии управления знаниями, накопление и обработка знаний, нечеткая модель, коэффициент забывания.*

KNOWLEDGE MANAGEMENT AS ONE OF OPEN EDUCATION TECHNOLOGIES' DEVELOPMENT DIRECTIONS

The article deals with the problem of creating knowledge management scenarios as the principal organization of information processes in the open education technologies, as well as the principles of their design and evolution. Solution the problem of information overflow in intelligent learning systems was described from the point of transition from the storage and processing of data to the accumulation and processing of knowledge. Increase the probability of producing new knowledge in learning is expected due to the use of interdisciplinary and multidisciplinary approaches to gathering information. Fuzzy forecast model of the student's knowledge level estimation was proposed, taking into account the natural loss of some knowledge for a long time period.

Keywords: *intelligent learning systems, information processes, knowledge management scenarios, knowledge accumulation and processing, fuzzy model, forgetting factor.*

Введение

Развитие современных технологий открытого образования предполагает улучшение результатов обучения за счет значительного усиления синергетического эффекта, обеспечивающего дополнительное качество полученных в процессе обучения знаний, не присутствующее в представленных образовательной программой слагаемых. Данный принцип подразумевает активное развитие интеллектуальных методов и систем обучения и самообучения. Основным препятствием на пути развития подобных интеллектуальных инструментов становится проблема информационного переполнения, усложняющего естественный процесс семантической концентрации и освоения знаний.

Сложность проблемы состоит в определении сценариев управления знаниями, как принципиальной организации информационных процессов и их взаимоотношений, а также принципов их проектирования и эволюции. Под управлением знаниями будем понимать – непрерывное управление процессами проверки знаний для выявления новых закономерностей с целью создания и удовлетворения спроса на новые знания. Управление знаниями отличается от управления информацией наличием метаданных и метаутверждений, как оценок достоверности, дающих обучаемому аргументы и доказательства, стоящие за этими утверждениями. Гипотеза разработки основана на следующем предположении: для решения проблемы информационного переполнения необходим переход от хранения и обработки данных к накоплению и обработке

[1] Исследование выполнено за счет гранта Российского научного фонда (проект №14-11-00242) в Южном федеральном университете.

Юрий Алексеевич Кравченко,
к.т.н., доцент
Тел.: (928) 908-01-51
Эл. почта: krav-jura@yandex.ru
Южный федеральный университет
www.sfedu.ru

Yury A. Kravchenko,
PhD in Technical Sciences,
Associate Professor
Tel.: (928) 908-01-51
E-mail: krav-jura@yandex.ru
Southern federal university,
www.sfedu.ru

знаний. Ожидаемые конкретные научные результаты могут проявляться в увеличении вероятности продуцирования нового знания за счет использования междисциплинарного и мультидисциплинарного подходов накопления информации. Также необходимо учитывать то, что результат процессов накопления и обработки знаний находится в объективной зависимости от естественных процессов, связанных с потерей некоторой части знаний в течение длительного промежутка времени. Поэтому процессы накопления и обработки знаний должны проектироваться с учетом корректных прогнозных моделей формирования остаточных знаний в долговременной памяти обучаемого.

1. Выявление новых знаний на основе циклического сценария

Накопление и обработка знаний – сложные многоэтапные процессы, требующие соответствующего управления для обеспечения необходимой последовательности, согласованности и эффективности. Отталкиваясь от того мнения, что знания не существуют в готовом виде [1], сформулируем **аксиому 1**: *Новые знания непрерывно производятся в процессе обработки знаний.* Неверным будем считать подход к управлению знаниями, как к фиксированию, кодированию и распространению имеющейся информации. Управление знаниями отличается от управления информацией наличием метаданных и метаутверждений, как оценок достоверности, дающих исследователю аргументы и доказательства, стоящие за этими утверждениями.

Говоря об архитектуре циклического сценария управления знаниями, как принципиальной организации информационных процессов и их взаимоотношений, а также принципов их проектирования и эволюции, выделим следующие основные этапы (рис. 1):

1) накопление (выявление) неформализованных знаний из разнородных источников информации;

2) классификация знаний (идентификация предметной области, кластеризация, определение законов распределения, анализ таблиц частот возникновения событий и т.д.);

3) формализация (представление) знаний на основе создания онтологии, семантических, нечетких и других моделей;

4) обработка (преобразование) знаний (создание унифицированных хранилищ данных на основе разработанной онтологии, семантических концентраторов и нечетких моделей);

5) распространение (реализация) знаний (траектории принятия решений, системы нечеткого вывода);

6) использование (применение) знаний (оптимизационные процедуры, интеллектуальный анализ данных, технологии data mining);

7) анализ и оценка эффективности знаний (создание имитационных моделей и графических диаграмм информационных потоков).

Сформулируем ***концепцию*** интеллектуальной поддержки принятия решений, основанную на процессах цикла управления знаниями. *Выявление* новых знаний связано с изучением и поиском отношений между разнородными источниками знаний, так как именно междисциплинарный и мультидисциплинарный подходы накопления информации с большей вероятностью после обработки приведут к продуцированию нового знания. Результатом следующего этапа проведения *классификации* полученных знаний на основе методов и технологий разведочного анализа данных является разработка визуальной объектно-ориентированной модели управления знаниями.

Представление знаний основано на построении онтологической модели поддержки принятия решений, а также семантических и нечетких моделей формирования знаний. Данная задача связана с целым рядом подзадач разработки: баз знаний, содержащих блоки правил принятия решений и прецедентов; множеств объектных, онтологических, нечетких, семантических и аналитических моделей, реализующих процессы принятия решений; модулей выбора моделей и формирования решений на осно-

ве базы знаний, математического и имитационного моделирования [2]. Создание подобных моделей позволяет адаптировать исследуемые информационные потоки к изменяющимся внешним условиям и интегрировать разнородные предметные области в рамках совместного решения задач управления знаниями в реальном времени.

В неформальной трактовке онтология представляет собой описание понятий, принадлежащих исследуемой предметной области, и отношений между ними, которое состоит из терминов и правил их использования, ограничивающих значения этих терминов в рассматриваемой предметной области.

Формально онтология состоит из организованных в таксономию терминов словаря, их определений и атрибутов, а также связанных с ними аксиом и правил вывода. Зачастую набор утверждений, входящих в онтологию, представляют формулами логики первого порядка. При этом термины служат именами унарных и бинарных предикатов, называемых концептами и отношениями соответственно.

Таким образом, онтология – это система, состоящая из набора понятий и набора утверждений об этих понятиях, на основе которых можно строить классы, объекты, отношения, функции и теории. Такая система может быть реализована как база знаний, описывающая факты, которые являются всегда истинными в рамках определенной предметной области на основе общепринятого смысла, фиксируемого используемым словарем [3].

На основе разработанной онтологии происходит *преобразование* знаний в хранилище данных, обеспечивающее возможность поддержки принятия решений с применением методов и алгоритмов поиска решений на основе семантической концентрации знаний, а также моделей и методов оптимизации, представленных в онтологии. Развитие данного этапа в дальнейшем позволяет *реализовать* полученные новые знания за счет формирования траекторий принятия решений в виде интеграции цепочек логических рассуждений и стратегий формирования решений на основе технологий нечеткого вывода.

В процессе *применения* знаний происходит адаптация базы знаний с целью последующего обучения системы поддержки принятия решений новым знаниям на основе оптимизационных процедур, методов интеллектуального анализа данных и методов data mining. Анализ и оценка эффективности накопленных знаний проводится на основе построенных имитационных моделей управления знаниями и графических диаграмм управления информационными потоками.

На основе рассмотренных процессов циклического сценария управления знаниями сформулируем ***аксиому 2***: *Цель управления знаниями состоит в повышении их прозрачности*. Абстрактность данной аксиомы подразумевает определение широкого спектра задач, основным критерием эффективности решения которых будет прозрачность процессов сценария управления знаниями. Среди таких задач можно выделить: ***идентификацию, приобретение, совершенствование, распределение, использование и сохранение*** знаний.

Классическая задача прямой идентификации заключается в определении функциональной зависимости выходного сигнала y от входного – x ($y = F(x)$). Когда у исследователя нет непосредственного доступа к выходному сигналу, используют метод компараторной идентификации, позволяющей формулировать основные интеллектуальные выводы с помощью дедуктивного способа, исходя из физически имеющихся наблюдений. Для обработки больших массивов поступающей информации исследователи чаще всего используют технологии Data Mining и Text Mining. Data Mining предусматривает анализ структурированной информации, Text Mining анализирует сверхбольшие массивы неструктурированной информации на основе методов семантического анализа информационного поиска и управления. Компараторная идентификация (метод сравнения) является одним из способов технологии Text

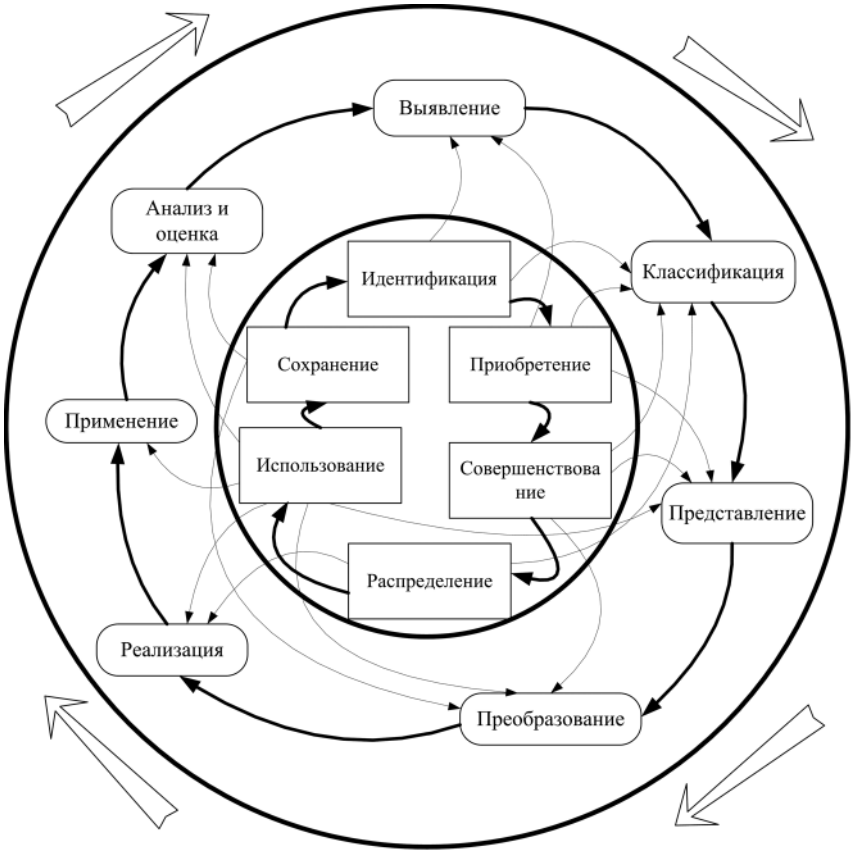

Рис. 1. Архитектура циклического сценария этапов и задач управления знаниями

Mining. В рассматриваемой задаче идентификации знаний предметной областью является множество процессов, которые составляют основу обработки информации. Компараторная идентификация применяется для формального описания периферических механизмов интеллекта (восприятие, узнавание, понимание), обрабатываемые с ее помощью информационные объекты должны быть дискретными, конечными и детерминированными. Для создания интегрированных моделей априорных знаний методом компараторной идентификации используют алгебру предикатов и предикатных операций. Системы предикатных уравнений решаются с помощью универсальных программ, написанных на языке высокого уровня. Любое уравнение алгебры предикатов может быть представлено в виде переключательной цепи как основы интеллектуального процессора обработки знаний.

Под приобретением знаний будем понимать процесс переноса знаний из разнородных источников в хранилище данных путем использования различных методов, моделей, алгоритмов и инструментальных средств. Данный процесс требует максимальной автоматизации, т.к. при иных стратегиях приобретения знаний всегда существует проблема передачи информации при совместной работе эксперта в предметной области и инженера знаний. Существует целый ряд стратегий и моделей приобретения знаний на основе использования экспертных систем, интеллектуальных редакторов и индуктивных программ, автоматически формирующих значимые отношения и правила, описывающие предметную область.

Любая интеллектуальная информационная система должна обладать способностью верифицирования и совершенствования имеющейся у нее базы знаний (БЗ), т.е. должна иметь инструменты анализа имеющихся формализованных знаний и их использования. Целью анализа является эволюция (улучшение) знаний. Для этого система должна анализировать сценарии рассуждений под углом их успеха или неудачи и проводить сравнение с имеющимися прототипами в базе ретроспективных прецедентов. На основе проведенного сравнительного анализа происходит доработка БЗ в части улучшения имеющихся сценариев использования знаний и рассуждений. Реализация данной функции может быть эффективно организована на основе многоагентного моделирования. Выбор многоагентных систем в качестве средства моделирования обусловлен эффективностью их применения для построения моделей сценариев идентификации, приобретения и усовершенствования знаний, возникающих в результате локальных взаимодействий простых реактивных агентов.

При разработке подобных многоагентных систем необходимо решить следующий ряд проблем:

– создание оригинальной архитектуры системы;

– проектирование онтологии предметной области;

– выбор инструментария и среды взаимодействия агентов;

– определение программных средств для реализации интеллектуального наполнения агента;

– интеграция всех компонентов в единую интеллектуальную информационную систему.

Следовательно, актуальной является разработка подхода, совмещающего в себе методы преодоления отмеченных проблем для определенного ряда типовых задач управления знаниями. Можно выделить множество таких задач, связанных с распределением, использованием и сохранением знаний в интеллектуальных информационных системах, для решения которых многоагентная архитектура представляется излишне общей. При построении многоагентных систем внимание акцентируется на унифицированном внешнем представлении знаний с целью обмена между агентами различной природы, а онтология представляет собой некоторое внешнее по отношению к агентам определение. При этом задание структуры предметной области будет проходить на основе иерархической комбинации фреймовых субиерархий, являющихся естественной реализацией таксономии онтологии. Для задания динамики предметной области распределенная фреймовая иерархия может быть расширена продукционными правилами.

2. Нечеткая прогнозная модель оценки уровня знаний

Для проведения процедуры выявления и подтверждения знаний используется цепочка вопросно-ответных структур нетривиального вида (множественный выбор ответов, многоуровневая структура вопросов и др.). В качестве основных данных имеем множество ответов *ANSW*, множество вопросов *QUES*, взаимосвязь которых определяет нечеткое отношение $\tilde{R}_{answ} \subset R_{answ}$, характеризующее правильные ответы на вопросы из подмножества $ANSW' \subseteq ANSW$, указанных ответов. Правильность ответа субъекта отражает условное нечеткое подмножество $\widetilde{QUES}_{ANSW} \subset QUES$, индуцируемое подмножеством ANSW′ и нечетким отношением $\tilde{R}_{answ}$ [4,5]. Для проведения оценки всех указанных субъектом обучения ответов необходимо определить функцию принадлежности нечеткого множества $\widetilde{QUES}_{ANSW}$.

При определении функции принадлежности $\mu_{\widetilde{QUES}_{ANSW}}(ques)$ необходимо учесть коэффициенты:

– забывания (потери возможности воспроизводить приобретенную ранее информацию) $K_{frgtA}(ques)$;

– усредненности оценок ответов $K_{middleA}(ques)$, данных субъектом на вопрос *ques*.

Тогда выражение определения функции принадлежности нечеткого множества $\widetilde{QUES}_{ANSW}$ примет вид:

$$\mu_{\widetilde{QUES}_{ANSW}}(ques) = \frac{K_{frgtA}(ques)}{K_{middleA}(ques)} \cdot \cdot \sum_{answ} \left(\mu_{ANSW'}(answ) \cdot \mu_{\tilde{R}_{answ}}(answ, ques) \right).$$

Согласно когнитивной теории с точки зрения структурной организации память делится на долговременную, кратковременную и оперативную. Полученные знания оставляют след в памяти, что и яв-

ляется признаком их сохранения. Забывание – активный и непроизвольный процесс потери возможности воспроизводить приобретенные ранее знания.

Из литературных источников известны три основные теории забывания:

1. *Ретроактивное и проактивное торможение* – забывание, связанное с получение новых (ретроактивное) или доминированием уже имеющихся знаний (проактивное);

2. *Мотивируемое забывание* – намеренное исключение из памяти знаний, связанных с отрицательным опытом их приобретения;

3. *Систематическая деформация следов памяти* – ослабление структурных связей в тканях головного мозга, что является причиной постепенного разрушения следов памяти и утере имеющихся знаний [6–12].

В случае с профессиональными знаниями, ни ретроактивное, ни проактивное торможение к процессу забывания информации непричастны. Мотивированное забывание также исключено, т.к. имеется потребность в знаниях. Таким образом, забывание знаний объясняется систематической деформацией, т.е. разрушением следа памяти и интерференцией (замещением) информации.

При определении коэффициента забывания необходимо учитывать то, что, согласно исследованиям немецкого ученого Г. Эббингауза и американского психолога М. Джонса, процесс забывания приобретенных во время обучения знаний не бесконечен (рис. 2), т.е. по истечении некоторого времени t_{frgt} (в среднем от месяца до полугода), активный процесс забывания прекращается, а в памяти субъекта остаются знания V_{frgt}, которые в дальнейшем (на протяжении многих лет – t_{const}) уменьшаются незначительно (3–5%), и в итоге могут быть представлены постоянной величиной V_{const} [7]. С точки зрения анализа и оценки уровня профессиональных знаний субъекта исследователей в первую очередь интересует коэффициент забывания для уменьшения уровня

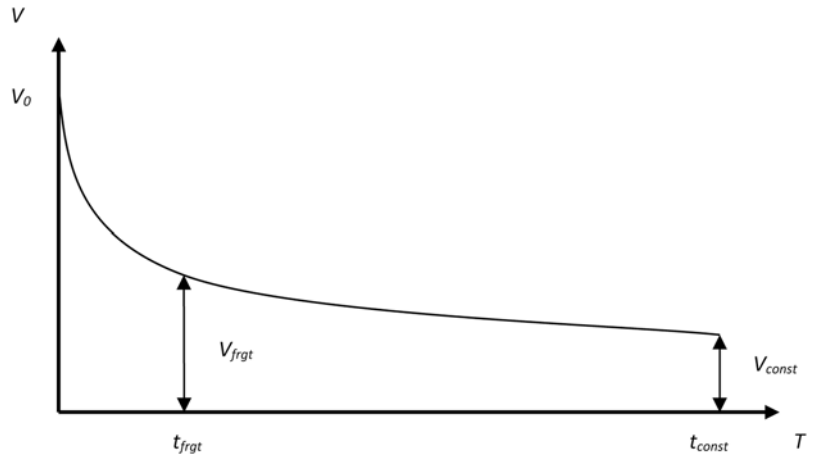

Рис. 2. Кривая забывания Джонса-Эббингауза

знаний в долговременной памяти $\Delta V = V_{frgt} - V_{const}$. Тогда с учетом известных экспериментальных данных $\Delta V_{min} = 0{,}03$, $\Delta V_{max} = 0{,}05$.

Для получения аналитического выражения вычисления значения коэффициента забывания K_{frgt} воспользуемся формулой

$$\Delta V = e^{-K_{frgt}\Delta t}$$

где $\Delta t = t_{const} - t_{frgt}$, логарифмируя, получим

$$K_{frgt} = \ln(\Delta V) / -\Delta t.$$

При вычислении числового значения K_{frgt} задается произвольное Δt, соответствующее известным статистическим данным экспериментальных исследований. В таком случае,

$$K_{frgtmax} = \ln(0{,}03) / -\Delta t,$$

$$K_{frgtmin} = \ln(0{,}05) / -\Delta t.$$

Таким образом, чем меньше, принимаемое за эталон, значение ΔV, тем больше значение коэффициента K_{frgt}, и наоборот. При этом, в любом случае, коэффициент забывания является понижающим. Определение коэффициента забывания как величины обратно пропорциональной скорости потери знаний $\beta = \Delta V / \Delta t$ позволяет при уменьшении скорости потери знаний увеличивать значение функции принадлежности нечеткого множества $\widetilde{QUES}_{ANSW}$ (отражающего уровень приобретенных в ходе обучения субъектом знаний) по отношению к функции принадлежности данного множества в условиях более высокой скорости забывания.

Коэффициент усредненности $K_{middleA}(ques)$ необходим для адекватности оценки степени правильности комплексного ответа на вопрос. При наличии Z абсолютно правильных ответов на вопрос X, необходимо также положительно оценивать случаи $Z - Y$ и $Z + Y$ правильных ответов, где Y – некоторая целочисленная переменная, причем $1 \le Y < Z$. Определим данный коэффициент следующим образом:

$$K_{middleA}(ques) = \\ = max\left\{\begin{array}{l}\sum\limits_{answ\in ANSW'} \mu_{R_{answ}}(answ, ques), \\ \sum\limits_{answ\in ANSW'} \mu_{core\tilde{R}_{answ}}(answ, ques)\end{array}\right\}$$

где $core\tilde{R}_{answ}$ – ядро нечеткого множества $\tilde{R}_{answ}$, а $\mu_{core\tilde{R}_{answ}}$ – его функция принадлежности.

Таким образом, значение функции принадлежности нечеткого множества $\widetilde{QUES}_{ANSW}$ позволит в полной мере оценить уровень и качество знаний субъекта профессиональной деятельности в данный момент, учесть естественные процессы, связанные с потерей некоторой части знаний в течение длительного промежутка времени, и создать корректную прогнозную модель остаточных знаний в долговременной памяти.

Заключение

Данная работа посвящена решению актуальной проблемы создания сценариев управления знаниями, как принципиальной организации информационных про-

цессов в технологиях открытого образования, а также принципов их проектирования и эволюции. Основным препятствием на пути разработки эффективных интеллектуальных инструментов организации информационных процессов становится проблема информационного переполнения, усложняющего естественный процесс семантической концентрации и освоения знаний. Для решения проблемы информационного переполнения необходим переход от хранения и обработки данных к накоплению и обработке знаний.

В статье представлена и описана архитектура циклического сценария этапов и задач управления знаниями, основанная на утверждениях необходимости повышения прозрачности знаний и осуществления обработки имеющихся знаний с целью получения новых. Сформулирована концепция интеллектуальной поддержки принятия решений, на основе описания процессов цикла управления знаниями.

Также рассмотрена проблема объективной зависимости результатов процессов накопления и обработки знаний от естественных процессов, связанных с потерей некоторой части знаний в течение длительного промежутка времени. Поэтому процессы накопления и обработки знаний должны проектироваться с учетом корректных прогнозных моделей формирования остаточных знаний.

Предложена нечеткая прогнозная модель оценки уровня знаний обучаемого. Функция принадлежности нечеткого множества, отражающего уровень приобретенных в ходе обучения субъектом знаний, позволяет оценить уровень и качество знаний обучаемого в данный момент, учесть естественные процессы, связанные с потерей некоторой части знаний в течение длительного промежутка времени.

Литература

1. *Тузовский А.Ф., Чириков С.В., Ямпольский В.З.* Системы управления знаниями (методы и технологии) / Под общ. Ред. В.З. Ямпольского. – Томск: Изд-во НТЛ, 2005. – 260 с.
2. *Черняховская Л.Р., Малахова А.И.* Разработка моделей и методов интеллектуальной поддержки принятия решений на основе онтологии организационного управления программными проектами // Онтология проектирования. – Самара: Изд-во «Новая техника», №4 (10), 2013. – С. 42–52.
3. *Гладун А.Я., Рогушина Ю.В.* Онтологии в корпоративных системах // Корпоративные системы. – М.: Комиздат, 2006. – №1.
4. *Денисова И.Ю.* Математические модели онтологии базы знаний информационной обучающей системы [Текст] / И.Ю. Денисова, П.П. Макарычев // Научный журнал «Онтология проектирования». – Самара: Изд-во «Новая техника», 2012. – №3(5). – С. 62–78.
5. *Кравченко Ю.А.* Интегрированная модель представления знаний на основе метода нечеткого моделирования // Известия ЮФУ. Технические науки. – Таганрог: Изд-во ЮФУ, 2013. – № 7 (144). – С. 97–102.
6. *Глазунов Ю.Т.* Математическая модель развития потребностей эстетического характера // Вестник МГТУ. – М.: Изд-во МГТУ, №4 (14), 2011. – С. 811–815.
7. *Метешкин А.А.* Апробация формулы для вычисления времени забывания студентами учебного материала / А.А. Метешкин, В.А. Шевченко, М.Н. Шарыгин // Автомобильный транспорт. – М.: №26, 2010. – С. 126–128.
8. *Курейчик В.В.* О правилах представления решений в эволюционных алгоритмах / В.В. Курейчик, С.И. Родзин // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – Таганрог: Изд-во ТТИ ЮФУ, 2010. – № 7 (108). – С. 13–21.
9. *Курейчик В.М.* Особенности построения систем поддержки принятия решений // Известия ЮФУ. Технические науки. – Таганрог: Изд-во ЮФУ, 2012. – № 7 (132). – С. 92–98.
10. *Курейчик В.В., Родзин С.И., Родзина Л.С.* Мобильное обучение: контекстная адаптация и сценарный подход // Открытое образование. 2013. № 4. С. 75–82.
11. *Гладков Л.А.* Новые подходы к построению систем анализа и извлечения знаний на основе гибридных методов / Л.А. Гладков, Н.В. Гладкова // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – Таганрог: Изд-во ТТИ ЮФУ, 2010. – № 7 (108). – С. 146–154.
12. *Кравченко Ю.А.* Способы интеллектуального анализа данных в сложных системах / Ю.А. Кравченко, Д.Ю. Запорожец, А.А. Лежебоков // Российская академия наук. Научный журнал. Известия КБНЦ РАН. – Нальчик: Изд-во КАБАРДИНО-БАЛКАРСКОГО НАУЧНОГО ЦЕНТРА РАН, 2012. – №3 (47). – С. 52–57.