

С.А. Петров

Категориально-информационная модель адаптивной системы непрерывного обучения

Рассмотрена категориально-информационная модель адаптивной системы поддержки непрерывного обучения, построенная в рамках теории распознавания образов и информационного критерия функциональной эффективности. Показан процесс оптимизации технологических и дидактических параметров системы поддержки обучения. Предложен вероятностный алгоритм дообучения системы в процессе ее функционирования и определения момента ее полного переобучения.

A categorial-information model of the adaptive system to support the lifelong learning built in the framework of the theory of pattern recognition and an information criterion of performance is considered. The process of optimization of technological and didactic parameters of the system of learning support is described. A probabilistic algorithm to adapt the system in its functioning and to determine the moment of its full retraining is suggested.

Розглянуто категоріально-інформаційну модель адаптивної системи підтримки неперервного навчання, побудовану в межах теорії розпізнавання образів та інформаційного критерію функціональної ефективності. Показано процес оптимізації технологічних та дидактичних параметрів системи підтримки неперервного навчання. Запропоновано ймовірнісний алгоритм донавчання системи в процесі її функціонування та визначення моменту її повного перенавчання.

Введение. Современные системы поддержки дистанционного образования основаны на экспертно-эмпирических моделях, в основу которых положены концепции, разрабатываемые отечественными и зарубежными учеными [1–3]. Развитие информационных технологий значительно опережает уровень формального описания и математического обоснования существующих систем поддержки непрерывного образования, трансформируя их в системы транспортировки контента к потребителю данных систем.

Серьезная проблема, сдерживающая развитие дистанционных систем и систем управления непрерывным обучением (СУНО), это недостаточное развитие формальных общепризнанных математических концепций их построения и функционирования. Жизненный цикл функционирования систем поддержки принятия решения, систем сопровождения учебного процесса состоит из следующих этапов:

- сбор данных и экспертных оценок в предметной области функционирования СУНО;
- формирование математического описания конструирование параметров функционирования системы и их оптимизация;
- работы в реальном времени (тестирование) с возможностью адаптации к обучаемому и дообучения системы [4, 5].

В силу различных подходов к реализации этапов функционирования системы не существует единых стандартов в синтезе СУНО. Обобщая рассмотренные системы можно разделить их на следующие группы:

- системы, базирующиеся на аддитивных схемах;
- системы, базирующиеся на байесовских алгоритмах принятия решений;
- нейросетевые и генетические алгоритмы;
- нечеткие системы принятия решений;
- алгоритмы, базирующиеся на анализе информационной компоненты системы.

Существует несколько различных классификаций систем приобретения знаний [6], топология которых показана на рис. 1.

Синтезом системы поддержки непрерывного обучения являются первые два этапа, непосредственно влияющие на эффективность работы системы на третьем этапе ее жизненного цикла. Особенно это важно, когда система предлагает пользователю индивидуальную траекторию обучения, оценку его уровня знаний, адаптивное формирование вспомогательного образовательного материала.

Постановка задачи

Рассмотрим обобщенную постановку задачи синтеза систем поддержки непрерывного обучения. Пусть множество входных данных сис-

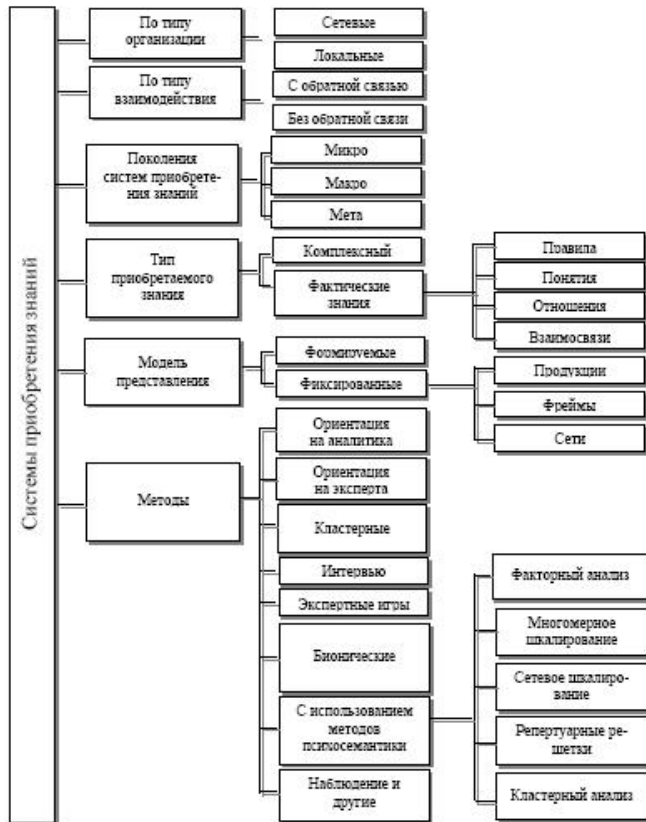


Рис. 1. Обобщенная классификация систем приобретения знаний

темы будет представлено: G – множество входных сигналов системы – множество данных, которые могут быть получены в результате начального функционирования системы, а именно статистических данных о действиях пользователя во время использования системы, ответы на контрольные вопросы, замеры времени, затраченного на различные виды работ, данных о структуре образовательного материала и др.; T – множество моментов сбора информации, которые функционально разделимы в контексте взаимодействия пользователя с системой; Ω – открытое множество характеристик объекта (слушателя), формируемых путем применения определенного оператора над множествами G и T ; Z – множество возможных функциональных состояний объекта (слушателя). S – множество выходных реакций системы с возможностью формирования корректирующего воздействия на систему. Матрица вида «объект-свойство» Y формируется оператором Φ , применяемым для четверки $\langle G, T, \Omega, Z, S \rangle$.

Оператор Φ представляет собой алгоритм гибридного кластер-анализа, который используя обобщенные метрики мехаланобиуса формирует классы характеристик, параметрически описывающих возможные функциональные состояния объекта [7, 8].

Построение модели СУНО

Для определения функциональной эффективности системы сопровождения непрерывного обучения используется обобщенная информационная мера Кульбака, имеющая вид [6]

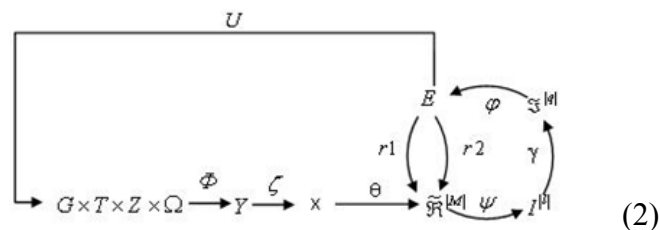
$$E_m = \frac{1}{2} \log_2 \left(\frac{D_1 + D_2}{\bar{b} + \bar{v}} \right) [(D_1 + D_2) - (\bar{b} + \bar{v})] = \log_2 \left(\frac{2 - (\bar{b} + \bar{v})}{\bar{b} + \bar{v}} \right) [2 - (\bar{b} + \bar{v})]. \quad (1)$$

Для ее вычисления нам необходимо перейти в пространство Хемминга. Оператор \mathcal{J} , действующий на обучающую матрицу Y , проводит ее бинаризацию формируя эквивалентную ей матрицу X , использует систему контрольных допусков, начальная база значений которой формируется во время работы гибридного алгоритма кластерного анализа. Оператор \mathcal{J} действуя на матрицу X формирует геометрические параметры гиперсфер, описывающих множество значений векторов $X_i, i = 1 \dots N$, которые относятся к определенному функциональному состоянию.

Исходя из определения (1) значение критерия функциональной эффективности (КФЭ) может лежать в диапазоне от нуля до единицы. Учитывая то, что КФЭ показывает адекватность разработанной модели системы поддержки непрерывного обучения, задачей оператора Ψ по оптимизации параметров функционирования системы является поиск таких значений параметров функционирования, при которых значение КФЭ будет достигать асимптотического максимума в области определения его функции на основе обучающей выборки Y . В качестве параметров оптимизации могут быть выбраны: система контрольных допусков, геометрические параметры гиперсфер, дидактические и технологические параметры тестиро-

вания. Множество I хранит значение КФЕ в процессе оптимизации параметров функционирования системы на i -м шаге. Оператор γ определяет множество $\mathfrak{Z}^{|q|}$ точностных характеристик, необходимых для вычисления КФЕ. Оператор φ вычисляет максимальное значение КФЕ, являющееся элементом терм-множества E . Оператор U выполняет дообучение СУНО в процессе ее функционирования: во время работы в режиме тестирования для определения принадлежности вектора-реализации к некоторому классу, характеризующему определенное функциональное состояние объекта [1], ставит в соответствие вектору-реализации числовое значение, равное расстоянию от геометрического центра контейнера до заданной точки в пространстве характеристик объекта, описываемого множеством Ω . В случае массового скопления векторов-реализаций в значительном удалении от геометрического центра контейнера необходимо проводить процедуру переобучения системы [7].

Ниже показана категориальная модель функционирования СУНО.



Практический эксперимент

Для экспериментального подтверждения корректности построенной модели выбран курс «Программирование». В качестве входных данных взяты ответы студентов на тестовые вопросы по первому модульному циклу. Базовым классом функционального состояния СУНО выбран класс «отлично», первый класс – «хорошо», второй класс – «удовлетворительно», третий – «неудовлетворительно». Числовые значения КФЕ и точностных характеристик показывают значительное пересечение классов функциональных состояний объекта:

- для базового класса $E^* = 0,70751$; $D_1 = 0,75$; $\beta = 0,11$;

- для первого – $E^1 = 4,39232$; $D_1 = 1,00$; $\beta = 0,00$;
- для второго – $E^2 = 4,39232$; $D_2 = 1,00$; $\beta = 0,00$;
- для третьего – $E^3 = 1,24119$; $D_1 = 0,79$; $\beta = 0,14$.

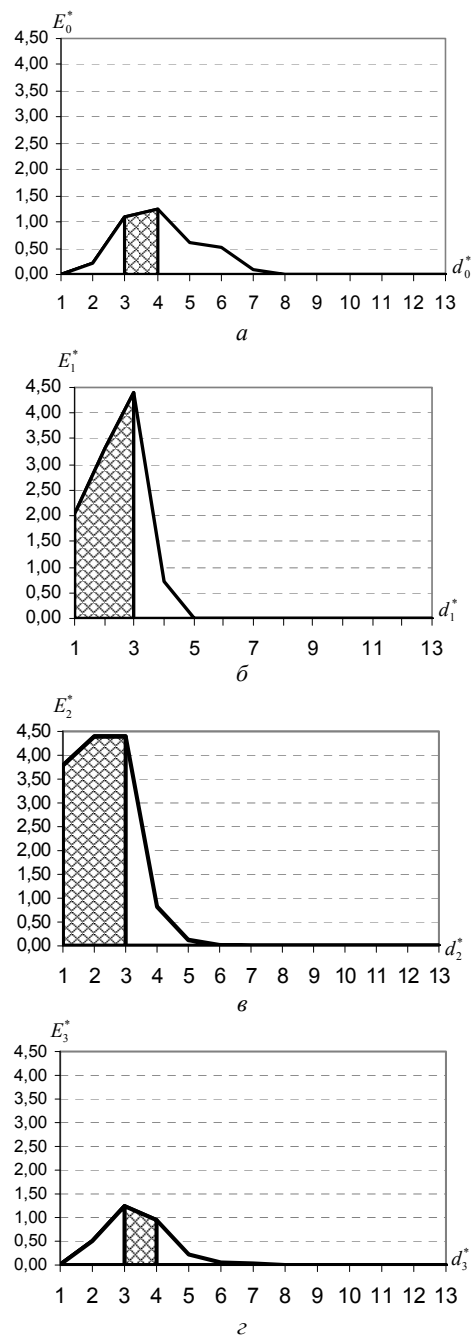


Рис. 2. График изменения КФЕ при обучении СУНО для четырех классов: a – базовый класс – «отлично»; b – первый класс – «хорошо»; v – второй класс – «удовлетворительно»; z – третий класс – «неудовлетворительно»

Учитывая малое значение КФЕ необходимо провести оптимизацию геометрических параметров контейнеров классов функциональных состояний объекта и дидактических параметров тестовых вопросов [8–9].

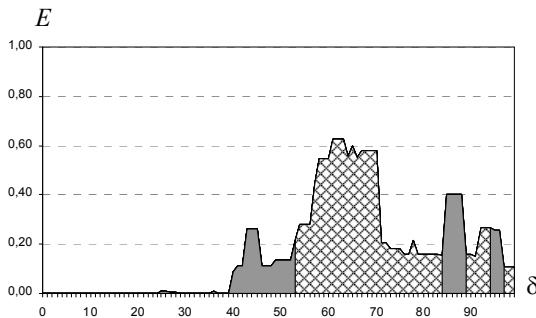


Рис. 3. График оптимизации геометрических и дидактических параметров функционирования СУНО

Значение нормированного значения КФЕ (рис. 3) повысилось и составило $E = 0,62$. Процедура оптимизации не позволила достигнуть асимптотического максимума значения КФЕ в области определения его функции, в связи с чем необходимо проводить дополнительную оптимизации параметров функционирования СУНО [10].

Заключение. В результате проведенной работы проанализированы существующие подходы к формированию входных данных систем поддержки принятия решений, проведена классификация соответствующих систем.

Предложен универсальный подход к проектированию СУНО, основанный на информационно-экстремальной интеллектуальной технологии синтеза слабо формализованных систем управления [7]. Рассмотрен алгоритм дообучения системы в процессе ее функционирования с возможностью определения момента переобучения системы. Недостаток разработанной концепции – метод не гарантирует достижения асимптотического максимума КФЕ в рабочей области определения ее функции, что указывает на необходимость поиска дополнительных параметров оптимизации СУНО [11]. Однако отметим, что при достаточно длительном функ-

ционировании СУНО удастся достигнуть близких к асимптотическому значению КФЕ.

1. Гриценко В.И., Довбыш А.С., Любчик В.А. Информационный синтез адаптивной мультиагентной системы управления дистанционным обучением // УСиМ. – 2006. – № 6 – С. 4–6, 25.
2. Федорук П.И. Использование адаптивных и интеллектуальных технологий в системах дистанционного обучения // УСиМ. – 2006. – № 5. – С. 68–73.
3. Peter Kazik, Igor Sivý. Intelligent e-learning systems // 5th International Conference on Emerging e-learning Technologies and Applications, Stara Lesna, the High Tatras, Slovakia. 2007. – P. 241–244.
4. SCORM 2004 Sequencing & Navigation // Learning Technology Publication of IEEE Comp. Society, 7. Issue 1 ISSN 1438-0625. – http://lttf.ieee.org/learn_ech/issues/january2005.
5. Rumelhart A.E., Hinton G.E., Williams R.J. Learning internal representations by error propagation / Ed. by D.E. Rumelhart, J.L. McClelland editors // Parallel Distributed Processing, Explorations in the Microstructure of Cognition. – 1, chap. 8. – P. 318–362. – MIT Press, 1986. – 523 p.
6. Korn F., Pagel B., Faloutsos C. On the 'Dimensionality Curse' and the 'Self-Similarity Blessing' // IEEE Trans. on Knowledge and Data Engineering, Jan., 2001. – N 1, 13. – P. 96–111.
7. Сороколетов В.П. Анализ, проблемы и состояние моделей представления знаний в системах принятия решений // Перспективные информационные технологии и интеллектуальные системы. – 2003. – № 3(12). – С. 16–24.
8. Сироджа И.Б. Квантовые модели и методы инженерии знания в задачах искусственного интеллекта // Искусственный интеллект. – Донецк, 2002. – № 3. – С. 161–171.
9. Краснопоясковский А.С. Информационный синтез интеллектуальных систем керування: Підхід, що ґрунтується на методі функціонально-статистичних випробувань. – Суми: Вид-во СумДУ, 2004. – 261 с.
10. Довбыш А.С., Любчик В.О., Петров С.О. Машинна оцінка знань студентів у системі керування дистанційним навчанням // Вісн. СумДУ. Сер. тех. наук. – 2007. – № 1. – С. 167–178.
11. Алексахин С.В., Николаев А.Б., Строганов В.Ю. Модели адаптивного тестового контроля в системе дистанционного образования. – <http://ito.edu.ru/2001/ito/VI/VI-0-17.html>