

ДИФФЕРЕНЦИАЛЬНАЯ КЛАСТЕРИЗАЦИЯ ФУНКЦИОНАЛЬНЫХ ХАРАКТЕРИСТИК ОПЕРАТОРОВ АСУ

Рассматривается методика кластеризации функциональных характеристик операторов АСУ на основе дифференциальной модели обучения специального вида. Исследована эффективность типовых операционных средств. Приведены результаты численных экспериментов.

Повышение эффективности деятельности операторов и лиц, принимающих решение (ЛПР) в АСУ является важной и актуальной задачей, так как возникновение большинства ситуаций неудовлетворительного функционирования промышленных и энергетических объектов критического применения связано с так называемым человеческим фактором. Анализ функциональных задач операторов АСУ проводился многими авторами в самых различных постановках этой многоплановой деятельности. Изучалось влияние условий труда операторов на их показатели работоспособности, степени обученности на интегральные характеристики качества их деятельности. На наш взгляд, эта задача в концептуальной постановке должна в первую очередь опираться на системные характеристики тех функциональных задач, которые выполняет оператор АСУ. В большинстве случаев такими задачами слежения операторов в системах являются следующие:

- за уровнем входного сигнала и обеспечение своевременной его компенсации по мере возникновения отклонений от эталонного;
- за производной входного сигнала по текущим значениям самого сигнала в режиме так называемого «бессознательного дифференцирования» и командного реагирования в том случае, когда абсолютная величина производной превысит некоторый заданный порог;
- за интегралом от входного сигнала, т.е. работа оператора в режиме «бессознательного интегрирования».

Безошибочность работы операторов в таких режимах определяется не только характеристиками входных сигналов и эталонных функций, величиной порогов, превышение которых обязывает оператора перейти к задачам ЛПР, но и во многом индивидуальными характеристиками психофизиологической личностной организации, присущими тому типу, к которому относится данный оператор. Априори следует ожидать, что для определенного типа операторов некоторые из указанных задач являются естественными и вызывают у него минимум дискомфорта и ошибок, другие же задачи приводят к его стойкому некорректному поведению и быстрой утомляемости.

Целью данной работы является анализ характеристик идентификации оператора на принадлежность его к определенной классификационной группе, т.е. задача кластеризации операторской деятельности в зависимости от типа решаемых им задач в АСУ. Экспериментальный анализ результатов операторской деятельности приводит к необходимости рассмотрения временных последовательностей. Задача кластеризации временных последовательностей исследовалась в работах [1,2], авторы которых использовали различные классификационные признаки. Так, в работе [1] в качестве таких признаков были использованы значения самих последовательностей, в работе [2] использовался вектор статистических оценок числовых характеристик. В данной работе предлагается использовать в качестве классификационных признаков результаты идентификации дифференциальных моделей обучения операторов [3].

Общий метод получения математического описания процессов обучения операторов АСУ на основе дифференциальных моделей состоит в следующем:

- в соответствии с типом задания, выполнению которого обучается оператор, ситуационной обстановкой в системе и индивидуальными психофизиологическими характеристиками оператора выбирается одна из типовых моделей, структурно заданных;
- на основании априорных данных определяются начальные условия и вид правых частей дифференциальных уравнений, входящих в модель;
- на основе экспериментальных данных производится параметрическая идентификация моделей;
- строится аналитическое решение систем уравнений, описывающих процессы обучения.

В наиболее общем виде структура дифференциальной модели обучения 2-го порядка имеет вид [3]:

$$\begin{cases} dL/dt + \gamma_1 L + \gamma_2 I + \gamma_5 dI/dt + \gamma_3 K = 0; \\ dI/dt + \gamma_4 I = L + \gamma_6 dL/dt. \end{cases} \quad (1)$$

где $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6$ — коэффициенты, которые определяются по экспериментальным данным на этапе параметрической идентификации и зависят от режима обучения оператора, а также функций, исполняемых им, эргономических показателей, индивидуальных психофизиологических характеристик и т.д.;

$L(t)$ — интенсивность потока информации, усваиваемой оператором при его обучении;
 $I(t)$ — интенсивность потока обучающей информации;
 $K(t)$ — интенсивность потока информационных помех (учитывающая потери информации при ее передаче оператору, искажения и забывание).

Система (1) эквивалентна неоднородному дифференциальному уравнению второго порядка вида

$$T^2 \frac{d^2 L}{dt^2} + 2\xi T \frac{dL}{dt} + L = \omega \sigma(t) \quad (2)$$

где T — постоянная времени обучения оператора; ξ — коэффициент забывания; ω — коэффициент передачи информации.

На основании решения дифференциального уравнения (2) можно построить семейство эталонных функций обучения оператора, определенное на отрезке $t = [0, T]$, $t_i \in [0, T]$, $i = \{0, N\}$, имеющих общий вид:

$$L(t) = a_0 - a_1 e^{-\alpha t} - a_2 e^{-\beta t} \quad (3)$$

$$\text{где } \alpha = \frac{\xi - \sqrt{\xi^2 - 1}}{T}, \beta = \frac{\xi + \sqrt{\xi^2 - 1}}{T}, \xi^2 - 1 > 0, a_1 = -\frac{\omega \alpha}{\beta - \alpha}, a_2 = -\frac{\omega \beta}{\alpha - \beta}.$$

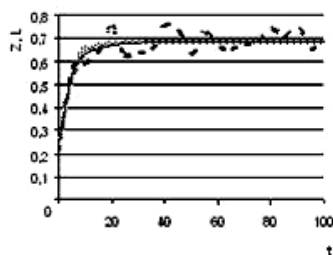
Каждая эталонная функция определяет соответствующий класс оператора АСУ. Возникает задача классификации неизвестного оператора, которая состоит в том, что необходимо по набору экспериментальных данных о функции $L(t)$ отыскать класс, к которому этот оператор принадлежит с максимальным правдоподобием. На основе экспериментальных данных, полученных за длительное время при обучении различных групп операторов указанным ранее функциональным задачам выделены 12 классов, имеющих следующие эталонные функции обучения: $L_1(t) = 0,7 - 0,25e^{-0,02t} - 0,25e^{-0,02t}$, $L_2(t) = 0,7 - 0,25e^{-0,02t} - 0,25e^{-0,2t}$, $L_3(t) = 0,7 - 0,25e^{-0,02t} - 0,25e^{-2t}$, $L_4(t) = 0,7 - 0,25e^{-0,2t} - 0,25e^{-0,02t}$, ..., $L_{11}(t) = 0,7 - 0,25e^{-2t} - 0,25e^{-0,2t}$, $L_{12}(t) = 0,7 - 0,25e^{-2t} - 0,25e^{-2t}$.

На основе экспериментальных исследований можно получить оценки $Z(t_i) = Z_i$ последовательности значений функции $L(t)$ во времени, которые являются исходными данными для решения задачи параметрического оценивания решений уравнения (2). Задача идентификации ставится как задача минимизации функции w , определяемой как сумма квадратов отклонений фактических значений от значений эталона в точках отрезка $[0, T]$, т.е.

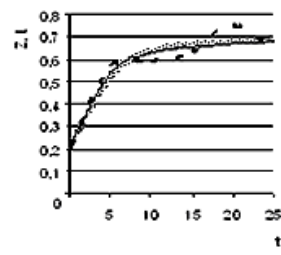
$$w(a_0, a_1, a_2, \alpha, \beta) = \frac{\sum_{i=0}^N (a_0 - a_1 e^{-\alpha t_i} - a_2 e^{-\beta t_i} - Z_i)^2}{N + 1} \rightarrow \min. \quad (4)$$

По результатам идентификации находится набор оценок параметров $R = \langle a_0, a_1, a_2, \alpha, \beta \rangle$, обеспечивающих минимум функции w . На первом этапе параметрической идентификации исследовалась эффективность различных методов решения задачи (4). Она может быть решена тремя основными методами: методом градиентного спуска с постоянным шагом, методом нелинейного регрессионного анализа и на основе процедуры Minimize системы вычислительной математики MathCad. Этим методам поставим в соответствие идентификаторы C1, C2, C3.

Конкретный набор значений величины $Z(t)$ на временном отрезке $[0, 100]$, кривые, полученные в результате параметрической идентификации этих данных, $L(t, CI)$, где CI — набор значений R, получен тремя указанными выше методами (C1, C2, C3) показаны на рисунке 1а. Из анализа полученных результатов видно, что кривая $L(t)$ имеет два характерных фрагмента: начальный, где происходит быстрый рост ее значений, и конечный, где производная $L(t)$ сравнительно мала. На рисунках 1б и 1в изображены фрагменты указанных выше кривых для временных диапазонов $[0, 25]$ и $[30, 100]$ изменения аргументов.



а)



б)

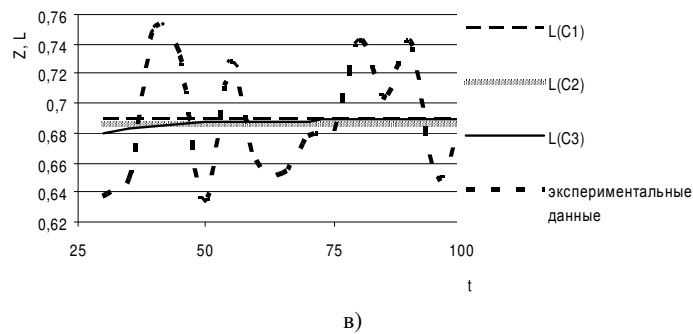


Рисунок 1 — Результаты параметрической идентификации кривых обучения

Выполним исследование качества параметрической идентификации кривых обучения при различных значениях параметров (малых и значительных дисперсиях экспериментальных значений), а также соотношениях параметров показателей обучения α и β , различных объемах выборок.

Результаты исследований сведены в таблицу 1, в которой представлены наборы значений R , оптимизирующие величину w , в различных режимах исследования. Каждому набору R соответствует аппроксимирующая кривая, структура которой задана выражением (3).

Для оценки близости функции $L(t)$, описывающей характеристики конкретного оператора и эталонной функции, определяющих выделенный класс используется мера W , вычисляемая как сумма квадратов отклонений аппроксимирующих кривых от эталона:

$$W(a_0, a_1, a_2, \alpha, \beta) = \frac{\sum_{i=0}^N (l(t_i) - L(t_i))^2}{N+1}.$$

Значения величины W приведены также в таблице 1.

Таблица 1 — Значения меры близости W

Метод определения коэффициентов аппроксимирующей линии		n=21 $\delta=10\%$ $\alpha=\beta$	n=21 $\delta=20\%$ $\alpha=\beta$	n=41 $\delta=10\%$ $\alpha=\beta$	n=21 $\delta=10\%$ $\alpha>>\beta$	n=21 $\delta=20\%$ $\alpha>>\beta$	n=41 $\delta=10\%$ $\alpha>>\beta$	n=21 $\delta=10\%$ $\alpha<<\beta$	n=21 $\delta=20\%$ $\alpha<<\beta$	n=41 $\delta=10\%$ $\alpha<<\beta$
Градиентный спуск с постоянным шагом C1	R	0,6900	0,6819	0,6968	0,7001	0,7000	0,6881	0,6994	0,6995	0,7001
		0,2515	0,2515	0,2482	0,2493	0,2496	0,2505	0,2499	0,2504	0,2496
		0,2515	0,2515	0,2482	0,2485	0,2493	0,2539	0,2500	0,2500	0,2500
		0,2004	0,2002	0,2004	2,0000	2,0000	2,0000	0,0171	0,0183	0,0223
		0,2004	0,2002	0,2004	0,1990	0,1995	0,2031	0,2000	0,2001	0,1997
	W	0,0019	0,0074	0,0014	0,0015	0,0060	0,0014	0,0012	0,0056	0,0007
Нелинейный регрессионный анализ C2	R	0,0001	0,0003	0,0000	0,0000	0,0000	0,0001	0,0002	0,0001	0,0001
		0,6861	0,6710	0,6948	0,7049	0,7099	0,6878	0,8033	0,6114	0,6783
		0,3684	0,3518	0,2601	0,1458	0,0399	0,3675	0,3086	0,0000	0,3194
		0,1092	0,1074	0,2601	0,3776	0,5069	0,1248	0,3005	0,4157	0,1429
		0,2202	0,2727	0,2120	1,9990	1,9970	0,8599	0,0074	0,1084	0,0359
	W	0,2202	0,2727	0,2120	0,1975	0,1968	0,1698	0,1820	0,1084	0,2419
Минимизация в MathCad (функция Minimize) C3	R	0,0018	0,0071	0,0014	0,0014	0,0055	0,0014	0,0012	0,0044	0,0006
		0,0002	0,0007	0,0000	0,0001	0,0005	0,0002	0,0002	0,0009	0,0002
		0,6887	0,6948	0,6948	0,7049	0,7099	0,6878	0,6916	0,5798	0,6805
		0,3730	0,3920	0,2601	0,1459	0,0400	0,3579	0,2241	-0,1643	0,3105
		0,0949	0,0999	0,2601	0,3775	0,5068	0,1353	0,2707	0,5397	0,1532
	W	0,2847	0,4711	0,2120	2,0000	1,9980	0,8900	0,0164	0,0238	0,0343
Эталон	R	0,0804	0,0321	0,2120	0,1974	0,1968	0,1789	0,2127	0,0778	0,2188
		0,0017	0,0067	0,0014	0,0014	0,0055	0,0014	0,0012	0,0043	0,0006
		0,0002	0,0008	0,0000	0,0001	0,0005	0,0001	0,0002	0,0015	0,0002
		0,7000	0,7000	0,7000	0,7000	0,7000	0,7000	0,7000	0,7000	0,7000
		0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500
	W	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500

Продолжение таблицы 1

Эталон	R	0,7000	0,7000	0,7000	0,7000	0,7000	0,7000	0,7000	0,7000	0,7000
		0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500
		0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500	0,2500
		0,2000	0,2000	0,2000	2,0000	2,0000	2,0000	0,0200	0,0200	0,0200
		0,2000	0,2000	0,2000	0,2000	0,2000	0,2000	0,2000	0,2000	0,2000
	W	0,0020	0,0078	0,0015	0,0015	0,0060	0,0015	0,0014	0,0057	0,0008

Анализ таблицы 1 показывает, что при решении задачи параметрической оптимизации все указанные методы обладают достаточно хорошими финальными характеристиками, однако требуют различных затрат машинного времени. Наиболее трудоемким является градиентный спуск с постоянным шагом, но вместе с этим для него характерны сравнительно лучшие результаты для случая небольших коэффициентов вариации данных. Меры расхождения w и W во многих случаях тесно взаимосвязаны и взаимообусловлены, что

вполне соответствует способу их конструирования. Более детальный анализ, кроме того, указывает на следующее:

- 1) увеличение коэффициента вариации экспериментальных данных вдвое приводит к ухудшению качества аппроксимации, оцениваемому по w в 7...10 раз;
- 2) при $\alpha \gg \beta$ и $\beta \gg \alpha$ метод градиентного спуска с постоянным шагом имеет подавляющее преимущество по точности аппроксимации;
- 3) при α приблизительно $=\beta$ все методы имеют один порядок ошибки идентификации;
- 4) при α приблизительно $=\beta$ даже существенное увеличение объема данных не изменяет величину w ;
- 5) при $\alpha \ll \beta$ увеличение объема данных позволяет существенно уменьшить показатель w .

Графически зависимости параметров моделей (3) для указанных случаев изображены на рисунках 2,а — 2,в.

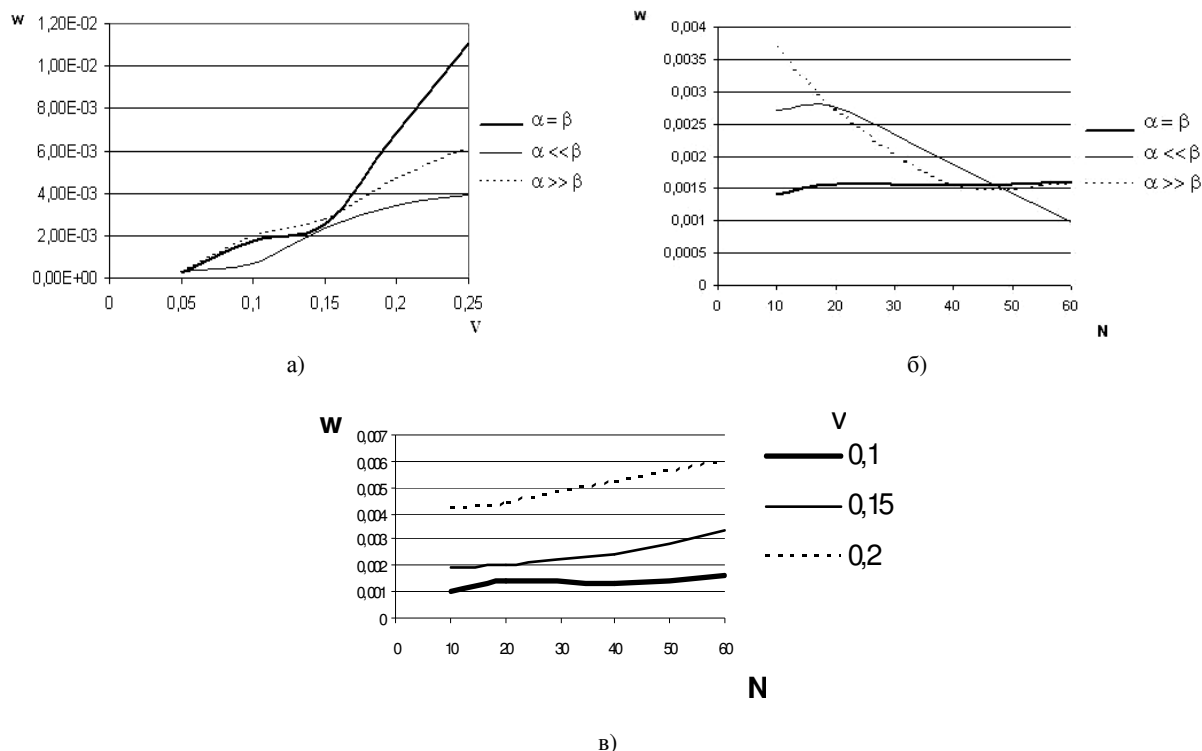


Рисунок 2 — Оценка качества параметрической идентификации кривых обучения оператора

На рисунке 2,а показаны зависимости величины w от коэффициента вариации v экспериментальных данных. Параметром является соотношение между коэффициентами α и β . На рисунке 2,б и 2,в — в зависимости от объема экспериментальных данных при различных значениях параметров.

При решении задачи кластеризации будем считать, что i -й класс оператора АСУ $K_i(\alpha, \beta)$ определяется значениями параметров α_i и β_i в (3). Весь интервал возможных значений параметров разбивается на классы со следующими центрами: $\alpha=0,2; 0,2; 1; 2$ и $\beta=0,02; 0,2; 2$. Таким образом, $K_1(0,02; 0,02)$, $K_2(0,2; 0,02)$, $K_3(1; 0,02)$, ..., $K_{12}(2; 2)$. Мера расстояний между введенными классами приведена в таблице 2.

Таблица 2 — Числовые характеристики выделенных классов операторов АСУ

	$\alpha=0,02$			$\alpha=0,20$			$\alpha=1,00$			$\alpha=2,00$		
	$\beta=0,02$	$\beta=0,20$	$\beta=2,00$	$\beta=0,02$	$\beta=0,20$	$\beta=2,00$	$\beta=0,02$	$\beta=0,20$	$\beta=2,0$	$\beta=0,02$	$\beta=0,2$	$\beta=2$
№ класса	1	2	3	4	5	6	7	8	9	10	11	12
1	0,0000	0,0107	0,0132	0,0107	0,0428	0,0473	0,0132	0,0473	0,0527	0,0132	0,0473	0,0528
2	0,0107	0,0000	0,0005	0,0000	0,0107	0,0132	0,0005	0,0132	0,0166	0,0005	0,0132	0,0166
3	0,0132	0,0005	0,0000	0,0005	0,0091	0,0107	0,0000	0,0107	0,0132	0,0000	0,0107	0,0132
4	0,0107	0,0000	0,0005	0,0000	0,0107	0,0132	0,0005	0,0132	0,0166	0,0005	0,0132	0,0166
5	0,0428	0,0107	0,0091	0,0107	0,0000	0,0005	0,0091	0,0005	0,0018	0,0091	0,0005	0,0019
6	0,0473	0,0132	0,0107	0,0132	0,0005	0,0000	0,0107	0,0000	0,0005	0,0107	0,0000	0,0005
7	0,0132	0,0005	0,0000	0,0005	0,0091	0,0107	0,0000	0,0107	0,0132	0,0000	0,0107	0,0132
8	0,0473	0,0132	0,0107	0,0132	0,0005	0,0000	0,0107	0,0000	0,0005	0,0107	0,0000	0,0005
9	0,0527	0,0166	0,0132	0,0166	0,0018	0,0005	0,0132	0,0005	0,0000	0,0132	0,0005	0,0000
10	0,0132	0,0005	0,0000	0,0005	0,0091	0,0107	0,0000	0,0107	0,0132	0,0000	0,0107	0,0132
11	0,0473	0,0132	0,0107	0,0132	0,0005	0,0000	0,0107	0,0000	0,0005	0,0107	0,0000	0,0005
12	0,0528	0,0166	0,0132	0,0166	0,0019	0,0005	0,0132	0,0005	0,0000	0,0132	0,0005	0,0000

Для примера рассмотрим результаты решения задачи кластеризации для тех же экспериментальных значений $Z(t)$. В соответствии с данными таблицы 2 рассматриваемый оператор без сомнения может быть отнесен к классу K_1 . Интересным представляется вопрос о том, насколько контрастно выглядит такое заключение по отношению к конкурирующим гипотезам, что является сутью дискриминантного анализа задачи кластеризации. С этой целью на рисунке 3 изображена зависимость, описывающая меру расхождения между результатами аппроксимации обучающей кривой донного оператора $Z(t)$ и эталоном класса K_i в зависимости от меры расстояния между эталонами (вторая строка на оси абсцисс указывает номер соответствующего класса). Результаты показывают на уверенное разделение классификационных признаков при решении задачи кластеризации для данного оператора.

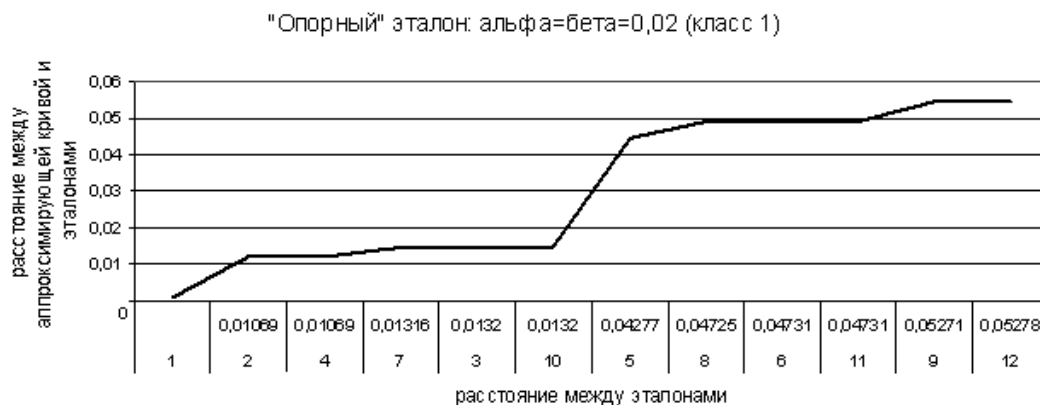


Рисунок 3 — Дискриминантный анализ при решении задачи кластеризации

Описанная методика кластеризации позволяет достаточно оперативно и в то же время вполне обосновано решать поставленную задачу для основных функциональных характеристик операторов АСУ с учетом наиболее эффективного способа решения задачи параметрической идентификации на основе типовых инструментальных средств.

Представляется перспективным продолжение исследований в этом направлении проводить путем анализа задачи в рамках теории принятия решений с рассмотрением соответствующих рисков.

Библиографический список

1. Вавилов Б.Т. О кластеризации временных рядов производственно-экономических показателей / Б.Т. Вавилов, Ю.В. Лунин // Вопросы радиоэлектроники. Серия АСУ. — 1979. — Вып. 3. — С. 71 — 80.
2. Дюран Б. Кластерный анализ / Б. Дюран, П. Оедд. — М.: Статистика, 1977. — 128 с.
3. Скатков И.А. Интеллектуальная поддержка принятия решений по организации процессов обучения операторской деятельности. / И.А. Скатков, Н.А. Скаткова // Труды Одесского политехнического университета: науч. и производ.-практ. сб. по техническим и естественным наукам. — Одесса, 2007. — Вып. 2(28). — С. 140 — 144.

Поступила в редакцию 22.04.2008 г.