УДК 519.6

Р. А. Иванов

Оптимизация поиска минимума информационной энтропии на конфигурациях разбиения при реализации ИКТ группового сотрудничества в процессе обучения

В работе рассмотрены вопросы оптимизации алгоритма поиска минимума информационной энтропии на конфигурациях разбиения обучаемого контингента на коалиции для реализации ИКТ эффективного группового сотрудничества в процессе обучения в малых группах.

Ключевые слова: оптимизация поиска, минимум информационной энтропии, ИКТ, процесс обучения, групповое сотрудничество, пространство конфигураций, разбиение, распределение вероятностей, тестирование.

R. A. Ivanov

Optimal Search of Minimum of Informational Entropy on Configurations of Clustering in Realization of IT Group Cooperation in the Teaching Process

In this work are regarded optimization problems of searching algorithm on configurations of teaching contingent clustering into coalitions, for realization of IT group cooperation in the teaching process of small undergroups.

Keywords: optimal search, minimum of informational entropy, IT, a teaching process, group cooperation, space configurations, clustering, probability distribution, testing.

1. Теоретико-информационные аспекты оптимизации группового сотрудничества в учебном процессе

При организации технологии группового сотрудничества в учебном процессе наиболее важным моментом является формирование такого разбиения обучаемого контингента на коалиции, при котором обеспечивается оптимальный учебный эффект. Процедура оптимизации в этом случае исходит из следующей информационной модели [2–7].

Пусть $A=\{a_1;a_2;\ldots;a_m\}$ – конечное множество, представляющее некоторый обучаемый контингент, которому предлагается выполнить некоторое задание (тест), при этом измеряется время его выполнения отдельными учащимися. В результате устанавливается некоторая цепочка неравенств $0<t_1\le t_2\le \ldots \le t_m<T$, где t_i – общее время выполнения задания i-м учащимся, в котором учтено качество проделанной работы; $i=\overline{1;m}$; T – временной регламент, определяемый параметрами теста. Пусть установленная цепочка неравенств – это некоторое устойчивое статистическое среднее, с достаточной точностью реализующееся при многократном испытании.

После этого определяются вероятности $\alpha_i=1-t_i/T$, характеризующие уровень обученности i-го учащегося, следуя положению классической тестологии [7]: более высоким скоростям интеллектуальных процессов (меньшие t_i) отвечает больший интеллектуальный потенциал тестируемого (большие α_i). По распределению вероятностей α_i строится распределение нормированных вероятностей:

$$p(a_i)=\frac{\alpha_i}{\alpha}=\frac{\alpha_i}{\alpha_1+\alpha_2+\ldots+\alpha_m},\ i=\overline{1,m}, \tag{1}$$

которые образуют полную систему, характеризующую рейтинги отдельных учащихся при выполнении данного задания.

Пусть теперь для улучшения показателей при обучении контингента A задействована технология группового сотрудничества. Формально, это выражается посредством разбиения множества

$$A=A_1\cup A_2\cup\ldots\cup A_n,\ A_j\cap A_k=\varnothing,\ j\neq k,\ j;k=\overline{1,n}, \tag{2}$$

причем параметры этого разбиения, связанные с формированием классов $A_j \subset A$, в данном случае выступают как параметры оптимизации рассматриваемой технологии обучения. Проведение самой оптимизации в рамках излагаемой модели осуществляется следующим образом.

Заметим, что для мощностей $|A_j|$ классов разбиения (2) выполняется соотношение:

$$|A_1| + |A_2| + \ldots + |A_n| = |A| = m \qquad (3)$$

Для проведения процедуры оптимизации в рамках излагаемой модели определим групповые вероятности: $p_j = \sum p(a_i), \forall a_i \in A_j$, (4)

где p_j есть вероятность того, что некоторый элемент из A входит в класс A_j, с вероятностями p_j связывается групповая информационная энтропия

$$H(p) = -\sum_{j=1}^{n} p_j \log_2 p_j \ . \qquad (5)$$

Оптимум в рассматриваемой информационной модели достигается, если минимальна энтропия $H(p)$. Следовательно, при оптимизации группового сотрудничества в учебном процессе разбиение (2) должно формироваться с учетом распределения (1) таким образом, чтобы при определении групповых вероятностей p_j обеспечивался минимум энтропии $H(p)$, то есть критерий оптимизации имеет вид: $H(p) \to min$. (6)

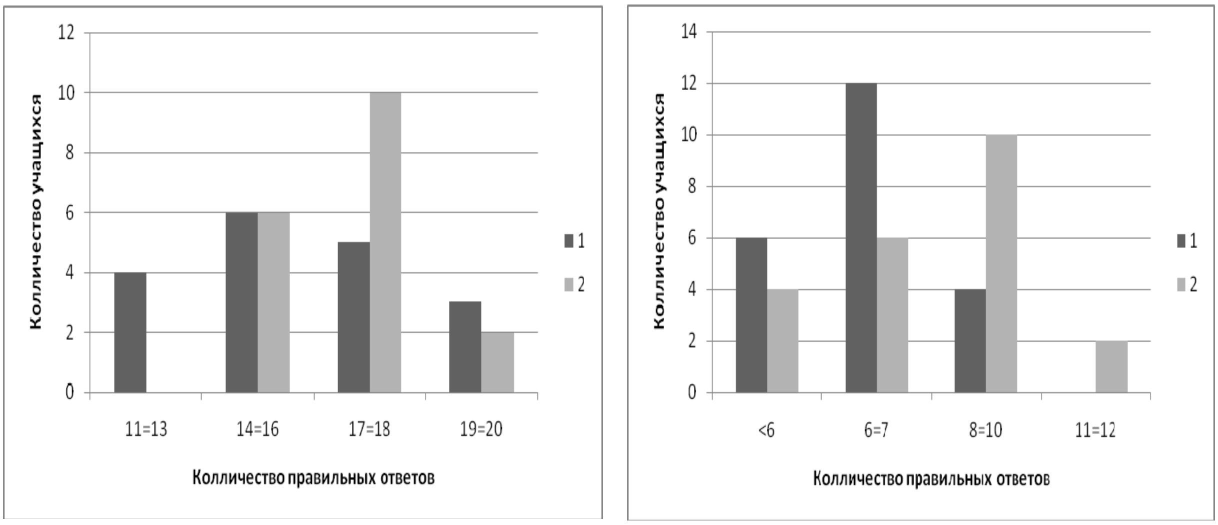
Легко показать [3, 4], что разность $\Delta H = H(A) - H(p) > 0$ между энтропией $H(A)$ при обучении данного контингента A как целого и энтропией $H(p)$ при обучении аудитории, разбитой на группы, всегда положительна, иными словами, реализация технологии сотрудничества при обучении приводит к снижению информационной энтропии в учебном процессе. Это связано с тем, что при разбиении на группы учебная информация прорабатывается не отдельным учащимся, а в процессе ее обсуждения в группе, что равносильно появлению дополнительных каналов целевого общения, реализующих режим усиления при восприятии целевой информации, то есть ее лучшее понимание и усвоение.

2. Экспериментальные данные по оптимизации группового сотрудничества в учебном процессе

Экспериментальная апробация теоретической модели оптимизации группового сотрудничества в учебном процессе (1–6) осуществлялась в 2008–2010 гг. как на уровне средней школы, так и на уровне ВПО [3]. Эксперименты проводились на базе Саратовского госуниверситета им. Н. Г. Чернышевского и в ряде школ Саратовской области, после чего полученный опыт был распространен на отдельные регионы России и аналогичные эксперименты были поставлены в ВГПУ (г. Воронеж), в КГУ г. Кострома), в БГПИ (г. Борисоглебск, Воронежской обл.), а также в Центре непрерывного математического образования при Институте прикладной математики и информатики РАН (Научный центр, г. Владикавказ, Республика Северная Осетия – Алания).

Отдельные результаты экспериментов представлены на рис. 1: цифры 1 и 2 – результаты тестирования до и после реализации ИКТ группового сотрудничества; по оси абсцисс – количество правильных ответов по тестовым заданиям; по оси ординат – соответствующее количество тестируемых респондентов. Данные на рис. 1а получены при организации группового сотрудничества на уроках математики в 4-м классе средней школы, каждый тест содержал 20 заданий; результаты на рис. 1б получены на занятиях со студентами I курса специальности 032100.00(050201) в ходе курса алгебры при изучении темы «Теория множеств», каждый тест содержал 12 заданий.

Данные на рис. 1 демонстрируют явную тенденцию улучшения показателей академической успеваемости посредством реализации ИКТ группового сотрудничества, поскольку заметно увеличивается контингент респондентов с большим количеством правильных ответов. В целом, как позволяет заметить анализ экспериментов [3], за счет реализации ИКТ группового сотрудничества наблюдается повышение показателей успеваемости по математике на 27,5 % в 4-м классе, на 25 % в 9-м классе и на 20–25 % в 10–11-х классах на фоне слабого уменьшения показателей в онтогенезе. На уровне ВПО повышение составляет 18–20 %.



а) **б)**

Рис. 1

3. Пространство конфигураций разбиений множества и числа Белла

Мощность пространства конфигураций разбиений m-элементного множества A, фигурирующего в разбиении (2), определяется количеством разбиений данного множества на классы и совпадает с числом Белла [2]:

$$B(m+1) = \sum_{i=0}^{m} C_m^i B(i), \qquad (7)$$

где $B(0) = B(1) = 1$. Числа Белла с увеличением мощности m множества A растут очень быстро, что иллюстрируется данными табл. 1.

Таблица 1

Числа Белла B(m) при значениях m от 0 до 20

m	0	1	2	3	4	5	6	7	8	9	10	11	12
B(*m*)	1	1	2	5	15	52	203	877	4140	21147	115975	678570	4213597

m	13	14	15	16	17
B(*m*)	27644437	190899322	1382958545	10480142147	82864869804

m	18	19	20
B(*m*)	682076806159	5832742205057	51724158235372

В связи с этим, как показывает опыт отладки программы перечисления конфигураций разбиения, при $m>15$ появляются проблемы с ресурсом оперативной памяти, которые требуют оптимизации сложности алгоритма перечисления конфигураций разбиения.

4. Компьютерное моделирование при определении минимума групповой информационной энтропии для реализации ИКТ группового сотрудничества

Модель, в рамках которой реализуется компьютерный поиск минимума групповой информационной энтропии, строится в рамках теоретико-информационных соображений п. 1. Пусть A – пространство событий с мощностью $|A| = m$, $B(A)$ – пространство конфигураций разбиений на A с мощностью $|B(A)| = B(m)$, и задана стохастическая мера $P(A)=\{p_1;p_2;\ldots;p_m\}$, где $p_1+p_2+\ldots+p_m=1$, $p_i \in [0;1]$, $i=\overline{1;m}$. В пространстве $B(A)$ в рамках распределения $P(A)$ требуется найти минимум групповой информационной энтропии $H(p)$, определенной соотношением (5).

В табл. 2 приведены результаты компьютерного поиска минимума энтропии $H(p)$ при $|A| = 5$ для распределения $P(A)=\{0,25; 0,35; 0,2; 0,15; 0,05\}$.

Таблица 2

<table>
<tr><th>№</th><th>Разбиение</th><th>Групповая энтропия</th><th>№</th><th>Разбиение</th><th>Групповая энтропия</th></tr>
<tr><td>1</td><td>|1||2||3||4|#|5|#####</td><td>0,286396957</td><td>27</td><td>|1||4|#|2||3|#|5|####</td><td>1,219240705</td></tr>
<tr><td>2</td><td>|1||2||3||5|#|4|#####</td><td>0,609840305</td><td>28</td><td>|1||5|#|2||3||4|#####</td><td>0,881290899</td></tr>
<tr><td>3</td><td>|1||2||3|#|4||5|#####</td><td>0,721928095</td><td>29</td><td>|1|#|2||3||4||5|#####</td><td>0,811278124</td></tr>
<tr><td>4</td><td>|1||2||3|#|4|#|5|####</td><td>0,88418372</td><td>30</td><td>|1|#|2||3||4|#|5|####</td><td>1,076297626</td></tr>
<tr><td>5</td><td>|1||2||4||5|#|3|#####</td><td>0,721928095</td><td>31</td><td>|1||5|#|2||3|#|4|####</td><td>1,406007579</td></tr>
<tr><td>6</td><td>|1||2||4|#|3||5|#####</td><td>0,811278124</td><td>32</td><td>|1|#|2||3||5|#|4|####</td><td>1,352724196</td></tr>
<tr><td>7</td><td>|1||2||4|#|3|#|5|####</td><td>0,991760148</td><td>33</td><td>|1|#|2||3|#|4||5|####</td><td>1,438758681</td></tr>
<tr><td>8</td><td>|1||2||5|#|3||4|#####</td><td>0,934068055</td><td>34</td><td>|1|#|2||3|#|4|#|5|###</td><td>1,601014306</td></tr>
<tr><td>9</td><td>|1||2|#|3||4||5|#####</td><td>0,970950594</td><td>35</td><td>|1||4||5|#|2|#|3|####</td><td>1,512887622</td></tr>
<tr><td>10</td><td>|1||2|#|3||4|#|5|####</td><td>1,188376372</td><td>36</td><td>|1||4|#|2||5|#|3|####</td><td>1,521928095</td></tr>
<tr><td>11</td><td>|1||2||5|#|3|#|4|####</td><td>1,278897903</td><td>37</td><td>|1||4|#|2|#|3||5|####</td><td>1,558871848</td></tr>
<tr><td>12</td><td>|1||2|#|3||5|#|4|####</td><td>1,352724196</td><td>38</td><td>|1||4|#|2|#|3|#|5|###</td><td>1,739353872</td></tr>
<tr><td>13</td><td>|1||2|#|3|#|4||5|####</td><td>1,370950594</td><td>39</td><td>|1||5|#|2||4|#|3|####</td><td>1,485475297</td></tr>
<tr><td>14</td><td>|1||2|#|3|#|4|#|5|###</td><td>1,533206219</td><td>40</td><td>|1|#|2||4||5|#|3|####</td><td>1,438758681</td></tr>
<tr><td>15</td><td>|1||3||4||5|#|2|#####</td><td>0,934068055</td><td>41</td><td>|1|#|2||4|#|3||5|####</td><td>1,5</td></tr>
<tr><td>16</td><td>|1||3||4|#|2||5|#####</td><td>0,970950594</td><td>42</td><td>|1|#|2||4|#|3|#|5|###</td><td>1,680482024</td></tr>
<tr><td>17</td><td>|1||3||4|#|2|#|5|####</td><td>1,188376372</td><td>43</td><td>|1||5|#|2|#|3||4|####</td><td>1,581290899</td></tr>
<tr><td>18</td><td>|1||3||5|#|2||4|#####</td><td>1</td><td>44</td><td>|1|#|2||5|#|3||4|####</td><td>1,558871848</td></tr>
<tr><td>19</td><td>|1||3|#|2||4||5|#####</td><td>0,992774454</td><td>45</td><td>|1|#|2|#|3||4||5|####</td><td>1,558871848</td></tr>
<tr><td>20</td><td>|1||3|#|2||4|#|5|####</td><td>1,234497797</td><td>46</td><td>|1|#|2|#|3||4|#|5|###</td><td>1,776297626</td></tr>
<tr><td>21</td><td>|1||3||5|#|2|#|4|####</td><td>1,44064545</td><td>47</td><td>|1||5|#|2|#|3|#|4|###</td><td>1,926120747</td></tr>
<tr><td>22</td><td>|1||3|#|2||5|#|4|####</td><td>1,457717469</td><td>48</td><td>|1|#|2||5|#|3|#|4|###</td><td>1,903701696</td></tr>
<tr><td>23</td><td>|1||3|#|2|#|4||5|####</td><td>1,512887622</td><td>49</td><td>|1|#|2|#|3||5|#|4|###</td><td>1,94064545</td></tr>
<tr><td>24</td><td>|1||3|#|2|#|4|#|5|###</td><td>1,675143246</td><td>50</td><td>|1|#|2|#|3|#|4||5|###</td><td>1,958871848</td></tr>
<tr><td>25</td><td>|1||4||5|#|2||3|#####</td><td>0,992774454</td><td>51</td><td>|1|#|2|#|3|#|4|#|5|##</td><td>2,121127473</td></tr>
<tr><td>26</td><td>|1||4|#|2||3||5|#####</td><td>0,970950594</td><td>min</td><td>|1||2||3||4|#|5|#####</td><td>0,286396957</td></tr>
</table>

Из табл. 2 видно, что min $H(p)=0{,}286396957$ наблюдается для разбиения $\{1;2;3;4\}\cup\{5\}$ (№ 1 в табл. 2). В дидактическом смысле полученный результат можно трактовать как указание на то, что с обучаемым 5 следует работать индивидуально, либо прикрепить слабых учащихся к сильным в рамках разбиения $\{1;\ 5\}\cup\{2;3;4\}$ с энтропией $H(p)=0{,}881290899$ (№ 28 в табл. 2).

5. Оптимизация поиска минимума групповой информационной энтропии в ИКТ группового сотрудничества

В рамках компьютерной модели оптимизации разбиения на группы (пп. 1, 4) удается моделировать специфику группировки в слабых, средних и сильных обучаемых контингентах. Это достигается заданием определенных статистических распределений $P(A)$. Например, ориентируя распределение $P(A)$ по закону Пуассона, можно моделировать сильные (максимум в области высоких показателей академической успеваемости) и слабые (с максимумом в области низких показателей успеваемости) контингенты; для контингента со средними показателями успеваемости распределение $P(A)$ обычно следует нормальному закону.

На рис. 2 распределения *P(A)* следуют закону Пуассона и в случае, когда |*A*|=*10*, *P(A)={0,034; 0,087; 0,145; 0,181; 0,181; 0,151; 0,108; 0,06; 0,037; 0,016}*; при |*A*|=*13*, *P(A)= {0,006; 0,023; 0,05; 0,09; 0,13; 0,15; 0,156; 0,13; 0,103; 0,074; 0,047; 0,027; 0,014}*. Оптимизация разбиения в рамках модели пп. 1,4 дает: при |*A*|=*10*, min *H(p) = 0,11835* при разбиении *{1;2;3;4;5; 6;7;8;9}*∪*{10}* и *H(p) = 0,2864* при разбиении *{1;10}*∪*{2;3;4;5;6;7;8;9}*; при |*A*|=*13*, min *H(p)=0,0529* при разбиении *{1}*∪*{2;3;4; 5;6;7;8;9;10;11;12;13}*.

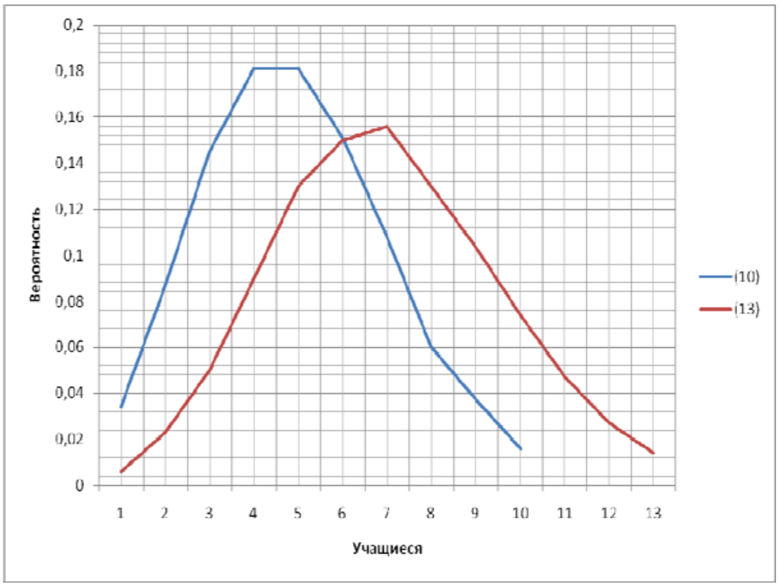


Рис. 2

Отметим, что алгоритм перечисления разбиений для условий рис. 2 сводился к прямой прогонке элементов (разбиений) пространства конфигураций *B(A)*, количество которых возрастает очень быстро в соответствии с рекуррентным соотношением (7) и данными табл. 1. Такой алгоритм минимизации групповой информационной энтропии *H(p)* является довольно трудоемким, и, например, анализ конфигураций разбиения при |*A*|=*10* занимает около 45 мин.; при |*A*|=*13* трудоемкость уже составляет около 2 часов при объеме информации примерно 1,5 Гб. Поэтому данный алгоритм перечисления конфигураций пространства *B(A)* был оптимизирован по сложности путем блочной обработки массива целевой информации, и, например, при |*A*|=*13* общий массив в 1,5 Гб был разбит на 238 блоков по 6,5 Мб и 1 блок объемом 2,4 Мб, которые обрабатывались 4-ядерным процессором, что позволило снизить трудоемкость алгоритма минимизации групповой информационной энтропии примерно вдвое.

Заключение

Как представляется, возможности для оптимизации алгоритма поиска минимума информационной энтропии в рассматриваемой ИКТ эффективной организации группового сотрудничества в процессе обучения не ограничены полученными результатами, которые могут быть улучшены. Дальнейшая работа в этом направлении проводится на основе анализа сетевой структуры алгоритма поиска и является текущей проблемой.

Отметим, что, кроме «интеллектуальных» показателей (1), при формировании групп должна также учитываться психологическая совместимость учащихся. Этот фактор контролируется по экспертным оценкам или методом социометрической матрицы [7].

Библиографический список

1. Гуревич, К. М. Тесты интеллекта в психологии [Текст] / К. М. Гуревич // Вопросы психологии. – 1980. – № 2. – С. 53–64.

2. Новиков, Ф. А. Дискретная математика для программистов [Текст] / Ф. А. Новиков. – СПб. : Питер, 2002. – 304 с.

3. Фирстов, В. Е. Кибернетическая концепция и математические модели управления дидактическими процессами при обучении математике в школе и вузе [Текст] : монография / В. Е. Фирстов. – Саратов : Издательский Центр «Наука», 2010. – 511 с.

4. Фирстов, В. Е. Количественные меры информации и оптимизация группового сотрудничества при обучении [Текст] / В. Е. Фирстов // Вестник Саратовского госуд. техн. ун-та. – 2008. – № 3 (34), Вып. 1. – С. 105–109.

5. Фирстов, В. Е. Математические модели управления дидактическими процессами при обучении математике в средней школе на основе кибернетического подхода [Текст] : автореф. дисс. … д-ра пед. наук / В. Е. Фирстов. – Ярославль, 2010. – 48 с.

6. Firstov, V.E. Semantic Model and Optimization of Creative Processes at Mathematical Knowledge Formation // Natural Science, 2010, Vol.2, No.8. – P. 915-922.

7. White, H.C. Social structure from multiple networks, I: blockmodels of roles and positions // Amer. J. Sociol., 1976, v. 81. – P. 730–780.