М.И.Рубинштейн, А.М.Черкашин

ФОРМИРОВАНИЕ ОБОБЩЕННОЙ ОЦЕНКИ ОБЪЕКТОВ, КОМПЛЕКСНО ОЦЕНИВАЕМЫХ НАБОРАМИ ПОКАЗАТЕЛЕЙ

I. Постановка задачи

При управлении сложными объектами, характеризующимися определенными наборами показателей, возникает задача формирования обобщенной их оценки. В работе предлагается новый подход к формализации и решению этой задачи, использующий метод автоматической классификации (см., например, [1]), допустимый в большинстве практических ситуаций.

Для формальной постановки задачи введем ряд обозначений:
i - индекс отдельных объектов, число которых равно N ;
j - индекс отдельных показателей, число которых равно M;
$\bar{K}_i$ - вектор размерности M , с компонентами K_{ij} , представляющими собой значение j -го показателя для i -го объекта; K_i - обобщенная оценка i-го объекта.

Кроме того сделаем два предположения. Во-первых, будем считать, что обобщенные оценки K_i принадлежат некоторой бальной шкале $\mathcal{J}_L = \{1, 2, \ldots, L\}$, включающей L градаций. Во-вторых, будем считать, что на множестве исходных показателей задана функция свертки $f(x_1, x_2, \ldots, x_M) = f(\bar{x})$ (в дальнейшем эта функция f предполагается неубывающей по каждому из M её аргументов).

Теперь рассматриваемая задача может быть сформулирована следующим образом: по заданным векторам $\bar{K}_i (i=\overline{1,N})$ и по заданной функции свертки f сформировать обобщенные оценки $K_i (i=\overline{1,N})$, принадлежащие бальной шкале $\mathcal{J}_L$.

Основная идея предлагаемого метода решения поставленной задачи состоит в том, чтобы, используя значения исходных показателей K_{ij} , разбить множество всех объектов $\mathcal{J}_N=\{1,2,\ldots,N\}$ на L групп (по числу градаций в шкале $\mathcal{J}_L$) и присвоить всем объектам каждой группы соответствующий балл (этот балл и определяет значение обобщенной оценки). Данный метод определения обобщенной оценки будет достаточно адекватным, если в формируемые группы попадают объекты достаточно близкие по значениям исходных параметров K_{ij}. Таким образом, реализация метода требует задания меры близости на множестве объектов $\mathcal{J}_N$, а также выработки критерия, характеризующего качество получаемого разбиения.

2. Задание меры близости на множестве оцениваемых объектов

Рассмотрим два способа задания меры близости.

Первый способ состоит в том, что предварительно формируются обобщенные оценки $\widetilde{K}_i$ для каждого объекта с использованием заданной функции свертки, т.е.

$$\widetilde{K}_i = f(K_{i_1}, K_{i_2}, \ldots, K_{i_N}). \qquad (I)$$

После этого, мера близости $a_1(i_1, i_2)$ для любых двух объектов i_1 и i_2 определяется соотношением:

$$a_1(i_1, i_2) = |K_{i_1} - K_{i_2}|, \quad i_1, i_2 \in \mathcal{J}_N. \qquad (2)$$

Второй способ состоит в том, что при определении меры близости учитываются все локальные приращения функции свертки, получаемые при переходе от значений показателей для одного из сравниваемых объектов к значениям тех же показателей для другого объекта. Точнее, в данном случае мера близости $a_2(i_1, i_2)$ для любых двух объектов i_1 и i_2 определяется соотношением:

$$a_2(i_1, i_2) = \sum_{j \in \mathcal{J}_N} \Big\{ \big| f[\bar{K}_{i_1}(K_{i_2 j})] - f[\bar{K}_{i_1}] \big| + \big| f[\bar{K}_{i_2}(K_{i_1 j})] - f[K_{i_2}] \big| \Big\}, \qquad (3)$$

где через $\bar{K}_i(K_{ij})$ обозначен вектор, отличающийся от вектора $\bar{K}_i$ только j-й компонентой, которая в первом из векторов совпадает с величиной K_{ij}.

Оба способа задания меры близости на множестве оцениваемых объектов представляются вполне адекватными. При этом, использование в (I),(2) и в (3) функции f содержательно

(по смыслу функции свертки) является вполне оправданным.

Будем в дальнейшем число $a_p(i_1, i_2)$, где $p = 1,2$ и $i_1, i_2 \in \mathcal{I}_N$, определяемое с помощью (1), (2) или (3), называть величиной связи между объектами i_1 и i_2. Все величины связей будем сводить в матрицу A_p $(p = 1,2)$ размера $N \times N$. Будем называть объекты одинаковыми, если соответствующие им наборы исходных показателей совпадают. Отметим, что при обоих способах задания меры близости (см.(1), (2) и (3)) для одинаковых объектов величина связи равна нулю. Однако, при первом способе задания меры, это может иметь место и для неодинаковых объектов. Вообще, в силу (1),(2) и (3) следует считать объекты тем ближе друг другу, чем меньше для них величина связи.

3. Критерии оптимальности разбиения объектов на группы (классификации объектов)

Критерий оптимальности разбиения на группы множества объектов с заданной на нем мерой близости должен быть связан с некоторой агрегированной характеристикой близости объектов во всем разбиении. Представляется, что существуют лишь несколько разумных способов задания такой агрегированной характеристики. Прежде чем описать их введем ряд обозначений: $I = (I_1, I_2, \ldots, I_L)$ - разбиение множества $\mathcal{I}_N$ на L подмножеств $I_1, I_2, \ldots, I_L$ (точнее, I-совокупность подмножеств I_ℓ множества $\mathcal{I}_N$, для которых $I_p \cap I_q = \Phi$ при всех $p, q \in \mathcal{I}_L$ и объединение которых дает $\mathcal{I}_N$); $a_{1p}(\mathcal{I})$ ($\mathcal{I}$ является произвольным подмножеством $\mathcal{I}_N$ и $p = 1,2$) - среднее значение величин связей для объектов, индексы которых составляют $\mathcal{I}$, т.е.

$$a_{1p}(\mathcal{I}) = \frac{1}{|\mathcal{I}|(|\mathcal{I}|-1)} \sum_{i_1, i_2 \in \mathcal{I}} a_p(i_1, i_2); \qquad (4)$$

$a_{2p}(\mathcal{I})$ (снова $\mathcal{I}$ является произвольным подмножеством $\mathcal{I}_N$ и $p = 1,2$) - максимальная величина связей для объектов, индексы которых составляют $\mathcal{I}$, т.е.

$$a_{2p}(\mathcal{I}) = \max_{i_1, i_2 \in \mathcal{I}_N} \{a_p(i_1, i_2)\}. \qquad (5)$$

Рассмотрим теперь пять функций $F_j(A_p, I)$, где $j = \overline{1,5}$ и $p = \overline{1,2}$, определяющих для заданных матрицей A_p величин связей и для выбранного разбиения I агрегированную величину связи (см. (4), (5)):

$$F_1(A_p, I) = \Pi\{|I_\ell|(|I_\ell|-1)\}^{-1} \sum_{\ell\in\mathcal{J}_L} \sum_{i_1, i_2\in\mathcal{J}_N} a_p(i_1, i_2), \quad (6)$$

$$F_2(A_p, I) = \sum_{\ell\in\mathcal{J}_L} a_{1p}(I_\ell), \quad (7)$$

$$F_3(A_p, I) = \max_{\ell\in\mathcal{J}_L} \{a_{1p}(I_\ell)\}, \quad (8)$$

$$F_4(A_p, I) = \sum_{\ell\in\mathcal{J}_L} a_{2p}(I_\ell), \quad (9)$$

$$F_5(A_p, I) = \max_{\ell\in\mathcal{J}_L} \{a_{2p}(I_\ell)\}. \quad (10)$$

Из (6) – (10) легко вывести два важных свойства данных функций. Во-первых, они монотонны по A_p (точнее, для любых двух матриц A'_p, A''_p из $A'_p \leq A''_p$ вытекает, что $F_j(A'_p, I) \leq F_j(A''_p, I)$), где $j = \overline{1,5}$. Во-вторых, все эти функции обращаются в 0 на разбиениях I, отдельные группы которых состоят только из одинаковых элементов. Эти свойства позволяют считать критериями оптимальности разбиений I минимальность для них одной из функций, определенных в (6)-(10).

Полученные таким образом критерии оптимальности разбиений являются различными, поскольку при одних и тех же матрицах A_p могут давать разные оптимальные разбиения I. Первые три критерия явным образом зависят от структуры разбиения, а именно, от величин $|I_\ell|(\ell\in\mathcal{J}_L)$. Такая зависимость введена в функции (6), (7) и (8) для того, чтобы множество оптимальных решений не менялось при преобразованиях матрицы A_p вида: $A_p + \alpha E$, где E – матрица соответствующего размера, состоящая из одних единиц, а α – произвольное число. Действительно, содержательно было бы сложно объяснить почему при указанном преобразовании матрицы меняется множество оптимальных разбиений. В критериях, определяемых функциями из (9) и (10), прямой зависимости от структуры разбиения нет. Тем не менее легко проверить, что указанные выше преобразования A_ρ не меняют для них множества оптимальных разбиений. Заметим кстати, что требование инвариантности этого множества по отношению к указанному преобразованию матрицы A_ρ резко сужает множество возможных критериев. Основные из них представлены в (6)–(10).

Независимо от того, какие из критериев используются при решении задачи группировки, практика во многих случаях накладывает на структуру выделяемого разбиения определенные ограничения. Наиболее распространенный вид этих ограничений таков (через m_ℓ здесь и далее обозначается количество элементов в подмножестве I_ℓ):

$$1 \leq n_{\ell\, min} \leq m_\ell \leq n_{\ell\, max} \leq N-L, \quad \ell \in \mathcal{J}_L . \qquad (11)$$

К ограничениям (11) часто добавляется ограничение такого содержания: выделено множество пар элементов (i_1, i_2), которые не могут включаться в одну и ту же группу. В дальнейшем рассмотрении это последнее ограничение отсутствует, однако его включение фактически не усложняет соответствующие задачи.

Таким образом, в связи с проблемой формирования обобщенной оценки можно сформулировать следующую задачу (задача I): минимизировать функцию $F_j(A_\rho, I)$, где $\rho = \overline{1,2}$ и $j = \overline{1,5}$, при ограничениях (11).

Решив данную задачу, мы еще не решим исходной задачи. Для полного решения последней необходимо поставить в соответствие группам оптимального разбиения I_ℓ баллы шкалы $\mathcal{J}_L$. Для этого необходимо определенным образом упорядочить группы оптимального для задачи I разбиения и в соответствии с полученным упорядочением присвоить группам баллы. При решении задачи упорядочения (будем называть ее задачей II) необходимо сформировать некоторый агрегированный обобщенный показатель $\widetilde{K}(I_\ell)$, характеризующий уровень значений обобщенных показателей $\widetilde{K}_i$ (см. (1)), для объектов группы I_ℓ . Простейшие способы формирования величин $\widetilde{K}(I_\ell)$ определяются следующими соотношениями:

$$\widetilde{K}_1(I_\ell) = \Big(\sum_{i \in \mathcal{J}_\ell} \widetilde{K}_i\Big) / |I_\ell|, \qquad (12)$$

$$\widetilde{K}_2(I_e)=\max_{i\in \mathcal{J}_e}\{\widetilde{K}_i\},\quad \widetilde{K}_3(I_e)=\min_{i\in I_e}\{\widetilde{K}_i\};$$
$$\widetilde{K}_4(I_e)=\max_{i\in I_e}[\widetilde{K}_2(I_e)+\widetilde{K}_3(I_e)].$$

С использованием введенных величин решение задачи II состоит в упорядочении групп разбиения $\widetilde{I}$ в соответствии с монотонным уменьшением величин $K_2(I_e)$ (уменьшение указано для определенности, важна лишь монотонность).

Таким образом исходная задача фактически свелась к задаче I (далее называется основной). В следующих разделах будут описаны алгоритмы решения этой задачи. Предварительно рассматривается модификация основной задачи, для которой описываются достаточно эффективные точные алгоритмы. Незначительная модификация этих алгоритмов позволяет использовать их для решения самой задачи I. Получаемое при этом решение в одних случаях совпадает с оптимальным, а в других является достаточно близким к нему. Последнее обуславливается тем, что задача I и её модификация являются формально и содержательно весьма близкими.

4. Алгоритмы решения модифицированной основной задачи

Рассматриваемая в данном разделе модификация задачи I связана с модификацией целевых функций, определенных в (6)-(10). Для задания модифицированных целевых функций введем величины $a_1(\mathcal{J})$ и $a_2(\mathcal{J})$, где $\mathcal{J}$ - произвольное подмножество множества $\mathcal{J}_0$:

$$a_1(\mathcal{J})=\sum_{i\in\mathcal{J}}\Big[\Big(\frac{1}{|\mathcal{J}|}\sum_{i_1\in\mathcal{J}}\widetilde{K}_{i_1}\Big)-\widetilde{K}_i\Big],\qquad (13)$$
$$a_2(\mathcal{J})=\max_{i\in\mathcal{J}}\{\widetilde{K}_i\}-\min_{i_1\in\mathcal{J}}\{\widetilde{K}_{i_1}\}.$$

Смысл величин, определенных в (13), достаточно ясен: $a_1(\mathcal{J})$ - сумма на множестве $\mathcal{J}$ отклонений величин показателей K_i от среднего их значения в том же множестве; $a_2(\mathcal{J})$ - диапазон изменений величин $\widetilde{K}_i$ во множестве $\mathcal{J}$.

Далее определим функции $F_j(\widetilde{K}, I)$:

$$\widetilde{F}_1(\widetilde{K},I)=\sum_{\ell\in\mathcal{J}_L}a_1(I_e);\qquad (14)$$

$$\widetilde{F}_2(\widetilde{K}, I) = \sum_{e \in \mathcal{J}_L} a_1(I_e); \qquad (15)$$

$$\widetilde{F}_3(\widetilde{K}, I) = \max_{e \in \mathcal{J}_L} \left\{ \frac{a_1(I_e)}{|I_e|} \right\}; \qquad (16)$$

$$\widetilde{F}_4(\widetilde{K}, I) = \sum_{e \in \mathcal{J}_L} a_2(I_e); \qquad (17)$$

$$\widetilde{F}_5(\widetilde{K}, I) = \max_{e \in \mathcal{J}_L} \left\{ a_2(I_e) \right\}. \qquad (18)$$

Легко видеть, что модифицированные целевые функции $\widetilde{F}_j(\widetilde{K}, I)$ совпадают с целевыми функциями основной задачи только для j =4,5 при P =I (сравните (I7),(I8) с (9), (I0)). Тем не менее и в остальных случаях по своему физическому смыслу исходные и модифицированные целевые функции являются весьма близкими.

Будем задачу I, в которой целевые функции из (6)-(I0) заменены соответственно целевыми функциями из (I4)-(I8), называть модифицированной задачей I. Для последней справедлива следующая теорема.

<u>Теорема I</u>. Пусть I_0^j – оптимальное разбиение модифицированной задачи I с целевой функцией $\widetilde{F}_j(\widetilde{K}, I)$. Пусть $I^j_{e_1 0}$ и $I^j_{e_2 0}$ – любые два подмножества этого разбиения и пусть для этих подмножеств определены величины:

$$\widetilde{K}^*_{min}(e_1) = \min_{i \in I^*_{e_1 0}} \{\widetilde{K}_i\}, \quad \widetilde{K}^*_{max}(e_1) = \max_{i \in I^*_{e_1 0}} \{\widetilde{K}_i\};$$

$$\widetilde{K}^*_{min}(e_2) = \min_{i \in I^*_{e_2 0}} \{\widetilde{K}_i\}, \quad \widetilde{K}^*_{max}(e_2) = \max_{i \in I^*_{e_2 0}} \{\widetilde{K}_i\}.$$

Тогда, возможны лишь два взаимоисключающих случая: либо $\widetilde{K}_{max}(e_1) \leq \widetilde{K}^*_{min}(e_2)$, либо $\widetilde{K}^*_{min}(e_1) > K^*_{max}(e_2)$.

Результат, аналогичный содержащемуся в теореме I, доказывается в [2]. Поэтому на доказательстве теоремы I мы не останавливаемся. Отметим, что результат теоремы I позволяет использовать при решении модифицированной задачи I простые процедуры динамического программирования. Действительно, в силу теоремы I все объекты достаточно упорядочить в порядке монотонного убывания величин $\widetilde{K}_i$, а затем в полученном ряду чисел провести границы групп с учетом выбранного критерия. Задача же о границах легко приводится к виду, допускающему использование алгоритмов динамического программирования (см. [2,3]).

Алгоритм решения модифицированной задачи I включает три этапа.

Этап I состоит в упорядочении исходных объектов в соответствии с монотонным убыванием величин $\tilde{R}_i$ и в последующей перенумерации объектов в соответствии с найденной упорядоченностью.

Этап 2 состоит в определении величин $B_{i\ell j}$ и индексов $m_{i\ell j}$ (для $i_{\ell\,min} \le i \le i_{\ell\,max}$ и $\ell \in \mathcal{J}_L$), которое осуществляется с помощью рекуррентной процедуры, выполняемой последовательно для $\ell = 2, 3, \ldots, L$ и описываемой следующими соотношениями:

$$B_{i1j} = b_j(1, i), \quad i_{1\,min} = n_{1\,min}, \quad i_{1\,max} = n_{1\,max} \tag{19}$$

$$\left.\begin{aligned} B_{i\ell j} &= \min_{i'_\ell \le i_1 \le i''_\ell} \left\{ B_{i_1\,\ell-1\,j} \oplus b_j(i_1+1, i) \right\}, \\ i'_\ell &= \max\left\{ i_{(\ell-1)\,min},\ i - n_{\ell\,max} \right\}, \\ i''_\ell &= \min\left\{ i_{(\ell-1)\,max},\ i - n_{\ell\,min} \right\}. \end{aligned}\right\} \tag{20}$$

$$\begin{aligned} i_{\ell\,min} &= i_{\ell-1\,min} + n_{\ell\,min}, \\ i_{\ell\,max} &= \min\left\{ i_{\ell-1\,max} + n_{\ell\,max},\ N - (n_{L\,min} + n_{L-1\,min} + \ldots + n_{\ell+1\,min}) \right\}. \end{aligned} \tag{21}$$

Индекс $m_{i\ell j}$ совпадает с индексом i_1, для которого в (20) достигается минимум. При этом, в (19),(20) используются величины $b_j(i_1, i_2)$, где $i_1, i_2 \in \mathcal{J}_N$, $i_1 < i_2$, которые соответствуют целевым функциям $F_j(\tilde{R}, I)$ и определяются следующим образом:

$$\begin{aligned} &\mathcal{J}[i_1, i_2] = \left\{ i \mid i_1 \le i \le i_2 \right\}; \\ &b_1(i_1, i_2) = a_1\left\{ \mathcal{J}[i_1, i_2] \right\}; \\ &b_2(i_1, i_2) = b_3(i_1, i_2) = \frac{a_1\left\{ \mathcal{J}[i_1, i_2] \right\}}{i_2 - i_1 + 1}; \\ &b_4(i_1, i_2) = b_5(i_1, i_2) = a_2\left\{ \mathcal{J}[i_1, i_2] \right\}. \end{aligned} \tag{22}$$

Кроме того, в (20) используется операция $\oplus$, которая для $j = 1,2,4$ совпадает с обычной операцией сложения, а для $j = 3,5$ дает максимальную из величин операндов. Отметим, что при $\ell = L$ процедуру достаточно выполнить лишь для $i = N$.

Этап 3 состоит в построении разбиения I. Указанное разбиение строится по результатам выполнения этапа 2, а точнее с использованием системы индексов $m_{i\ell j}$. Предварительно, в рекуррентной процедуре для ℓ , последовательно принимающих значения $L, L-1, \ldots, 2$, определяются индексы $m_{\ell j}$:

$$m_{\ell j} = m_{i'\ell j}, \quad i' = m_{\ell+1 j}, \tag{23}$$

где $m_{L+1 j} = N$. Затем в предположении $m_{1j} = 0$, определяются подмножества $\tilde{I}_{\ell o}^{j}$ $(\ell \in \mathcal{J}_L)$ подмножества разбиения $\tilde{I}_o^j$:

$$\tilde{I}_{\ell o} = \mathcal{J}\{[m_{\ell j}+1,\ m_{\ell+1 j}]\}, \quad \ell \in \mathcal{J}_L. \tag{24}$$

Имеет место следующая теорема.

Теорема 2. Разбиение $\tilde{I}_o^j$, полученное с помощью процедуры (19)-(24), является оптимальным в модифицированной основной задаче с целевой функцией $\tilde{F}_j(\tilde{K}, I)$.

Доказательство теоремы 2 опирается на теорему I и соотношения (19)-(24), по которым строится разбиение $\tilde{I}^j$. Отметим, что аналогичные результаты содержатся в [2,3].

5. Алгоритмы решения основной задачи

Рассмотрим два подхода к решению задачи I.

В первом подходе существенно используются результаты предыдущего раздела. Суть подхода в следующем. Упорядочим объекты по убыванию величин K_i . В рамках новой нумерации, соответствующей этому упорядочению, выполним процедуру (19)-(24), в которой вместо величин $b_j(i_1,i_2)$ будем использовать величины $b_{jp}(i_1, i_2)$, определяемые следующим образом $(p = 1,2)$:

$$b_{1p}(i_1, i_2) = \sum_{i_3=i_1}^{i_2} \sum_{i_4=i_1}^{i_2} a_p(i_3, i_4), \tag{25}$$

$$b_{2p}(i_1, i_2) = b_{3p}(i_1, i_2) = \frac{b_{1p}(i_1, i_2)}{(i_2 - i_1 + 1)}; \tag{26}$$

$$b_{4p}(i_1, i_2) = b_{5p}(i_1, i_2) = \max_{i_1 \leq i_3, i_4 \leq i_2} \{ a_p(i_3, i_4) \}. \tag{27}$$

В соответствующей процедуре параметры $B_{i\ell j}, m_{i\ell j}, m_{\ell j}, i_{\ell min}, i_{\ell max}$ и получаемое разбиение $\tilde{I}_0^j$ приобретают еще один индекс Р. Разбиение $\tilde{I}_0^j$ принимаем в качестве решения задачи I. Это решение является оптимальным для j =4,5 и p =I, и является приближенным в остальных случаях.

Решение задачи I в рамках первого подхода автоматически решает и задачу П. Точнее, получаемая при определении разбиения $\tilde{I}_0^{jP}$ нумерация его подмножеств $\tilde{I}_{\ell_0}^{jP}$ (см.(24)) совпадает с нумерацией, которая получилась бы при упорядочении этих подмножеств в соответствии с монотонным убыванием величин $\widetilde{K}_z(\tilde{I}_{\ell_0}^{jP})$ для любого $z = \overline{1,4}$.

Что касается второго подхода к решению задачи I, то он не связан с предварительным упорядочением объектов и состоит в непосредственном решении данной задачи. Для этого может быть использован любой из известных методов решения дискретной задачи классификации (см., например, обзор [4]). После решения задачи I для полученного разбиения решается задача П.

Отметим, что качество решения задачи П можно характеризовать некоторой функцией $\Phi(\pi, \widetilde{K}_z, I)$, где I – разбиение, $\widetilde{K}_z$ – вектор размерности L с компонентами $\widetilde{K}_z(I_\ell)$ ($z = \overline{1,4}$), π – перестановка из L элементов, упорядочивающая множества I_ℓ в соответствии с монотонным уменьшением величин $\widetilde{K}_z(I_\ell)$.

Два простейших и разумных способа задания функции Φ определяются следующими соотношениями :

$$\Phi_1(\pi, \widetilde{K}_z, I) = \min_{\ell \in \mathcal{J}_L \setminus L} [\widetilde{K}_z(I_{\pi_\ell}) - \widetilde{K}_z(I_{\pi_{\ell+1}})], \tag{28}$$

$$\Phi_2(\pi, \widetilde{K}_z, I) = \sum_{\ell \in \mathcal{J}_L \setminus L} [\widetilde{K}_z(I_{\pi_\ell}) - \widetilde{K}_z(I_{\pi_{\ell+1}})]. \tag{29}$$

Таким образом, исходную задачу можно трактовать как задачу векторной оптимизации с двумя критериями F (см. (6) - (10)) и Φ (см. (28)-(29)), что может оказаться полезным при разработке новых подходов к формированию обобщенной оценки объектов, комплексно оцениваемых наборами показателей.

Литература

1. Дорофеюк А.А. Алгоритм автоматической классификации (обзор литературы).-В кн.: Проблемы расширения возможностей автоматов. Вып. I, М., ИАТ, 1971.
2. Рубинштейн М.И. Эффективные алгоритмы решения некоторых задач классификации. Труды Всесоюзной школы-семинара по управлению большими системами. Тб., "Мецниереба", 1976.
3. Лившиц Э.М., Рублинецкий В.И. Об оптимальном разбиении упорядоченного множества на интервалы. - В кн.: Вычислительная математика и вычислительная техника. Вып. III. Харьков, Физико-техн. ин-т низких температур, 1972.
4. Миркин Б.Г. Дискретные задачи классификации взаимосвязанных объектов (обзор). - В кн.: Вопросы анализа сложных систем, Новосибирск, СО "Наука", 1974.