

ОБРАБОТКА ДАННЫХ

УДК 519.178 : 681.3 : 51-7

Л. И. Макаров

(Новосибирск)

ОТДЕЛИМОСТЬ ВЕРШИН ВЗВЕШЕННОГО ГРАФА
И АЛГОРИТМЫ ИХ КЛАССИФИКАЦИИ

Введено понятие отделимости вершин для ребер взвешенного графа, на основе которого определена величина качества таксономии (разбиения) вершин графа на классы близких вершин. Для задачи классификации с обучением рассматривается взвешенный граф, соответствующий объединению обучающей и контрольной выборки объектов. Вершины графа соответствуют объектам, а длины ребер — расстояниям между ними. Приведены алгоритмы классификации контрольных объектов, основанные на таксономии вершин взвешенного графа, учитывающей отделимость таксонов.

Во многих областях применения компьютерных методов искусственного интеллекта (распознавание образов, классификация и др.) при анализе и прогнозе свойств объектов используются графовые модели их выборок. Моделью исследуемой выборки объектов является взвешенный граф, вершины которого соответствуют объектам, а каждое ребро — отношению (связи) между парой объектов, задаваемому количественно положительной действительной величиной — длиной ребра.

При описании выборки объектов с помощью таблицы «объект—признак» длину каждого ребра взвешенного графа определяет расстояние между соответствующими объектами в пространстве признаков (свойств) [1].

При исследовании выборок структурированных объектов — объектов, состоящих из взаимосвязанных элементов, — следует учитывать, что их свойства зависят от элементного (фрагментарного) состава, внутренней структуры объектов, взаимного положения фрагментов и т. д. Для описания таких объектов используются помеченные графы структур: с внутренними элементами объекта сопоставляются вершины с соответствующими метками, а со связями элементов — помеченные ребра графа. Части (подграфы) такого графа называют его фрагментами.

Компьютерный анализ зависимости свойств объектов от их структурных характеристик основывается на гипотезе о том, что общие свойства объектов некоторого семейства обусловлены наличием в них общего структурного фрагмента. Цель анализа — нахождение существенных фрагментов, определяющих свойства данных объектов, и прогноз свойств новых объектов в предположении, что структурно близкие объекты обладают близкими свойствами. Такая гипотеза используется в органической химии при исследовании взаимосвязи молекулярных структур и свойств химических соединений [2]. При этом граф структуры называют молекулярным.

Для определения количественной меры структурной близости (подобия) пары объектов используют величину расстояния между их графами структур, зависящую от размера их наибольшего общего фрагмента [3, 4]. Чем больше общий фрагмент, тем больше структурная близость объектов и тем меньше

расстояние между ними. Поэтому выборке структурированных объектов можно сопоставить взвешенный граф, длина ребер которого определяется расстоянием между графами структур объектов.

Для взвешенного графа выборки объектов естественно возникает задача таксономии его вершин, т. е. разбиения множества вершин взвешенного графа на таксоны (классы) — подмножества близких (в смысле длин соединяющих ребер) вершин графа. Анализ полученной таксономии и разбиения объектов (вершин графа) выборки на классы их свойств позволяет установить зависимость между принадлежностью объекта данному таксону и его свойствами. Такая зависимость дает возможность определить решающие правила прогноза свойств (классификации) объектов, не входящих в выборку.

Известные методики классификации объектов с обучением предполагают наличие обучающей (ОВ) и контрольной (КВ) выборки объектов [1, 5]. Объекты ОВ разбиты на классы в соответствии с их свойствами. Для объектов КВ производится их классификация (прогноз свойств) с помощью определенных решающих правил. При этом классификация может осуществляться как для каждого контрольного объекта в отдельности, так и для всего множества объектов КВ одновременно [1, с. 43]. В исследуемых выборках возможно присутствие «сомнительных» объектов: объектов с недостоверными значениями характеристик, ошибочно включенных объектов, объектов с уникальными свойствами и т. д. Можно предположить, что такие объекты во взвешенном графе выборки связаны с другими объектами своего класса ребрами относительно большей длины, чем, например, длина ребер в окрестностях типичных объектов или объектов — эталонов класса. Нахождение «сомнительных» объектов взвешенного графа выборки позволяет провести коррекцию состава выборки и сформировать решающие правила классификации объектов КВ.

В данной работе приведены алгоритмы таксономии и классификации вершин взвешенного графа, основанные на принципе «ближайшего соседа» и понятии отделимости вершин ребер графа. Приведенный алгоритм таксономии является алгоритмом «разбора» кратчайшего остова взвешенного графа (см. также [1, с. 32]) в отличие от алгоритма «укрупнения», описанного в [3].

Предложенные алгоритмы таксономии и классификации были реализованы и прошли успешные испытания в системе ВАСС анализа и классификации биологически активных химических соединений, разработанной в Институте математики СО РАН для IBM PC.

Основные понятия и определения. Пусть $G(V, X)$ — граф, имеющий множество вершин $V = \{v, u, \dots\}$, $|V| = p_G$, и множество ребер $X = \{x \mid x = (v, u), v, u \in V\}$, $|X| = q_G$. Граф $F(V', X')$ называют подграфом графа G , если $V' \subseteq V$, $X' \subseteq X$. Каждый подграф F однозначно задается (порождается) множеством своих ребер X' . В этом случае его называют частичным и обозначают $\langle X' \rangle$. Подграфом, порожденным в графе G множеством вершин V' , называют подграф, в котором множество ребер X' состоит из всех ребер, соединяющих в G вершины из V' , и обозначают $\langle V' \rangle$.

Граф G называют взвешенным, если каждому его ребру $x = (v, u)$ приписано положительное действительное число $l(x) > 0$ — длина ребра. Взвешенный граф G задают взвешенной матрицей смежности (матрицей расстояний) $M(G) = \|l(x)\|$. Расстояние $r_G(v, u)$ между вершинами v и u взвешенного графа G определяется как длина кратчайшего пути в G между ними.

Взаимное положение подграфов F и H в графе G можно оценивать количественно, например, величиной $R(F, H) = \left(\sum_v r_G(v, H) + \sum_u r_G(u, F) \right) / p_{F \cup H}$, $r_G(v, H) = \min_u r_G(v, u)$, $v \in V_F$, $u \in V_H$, $r_G(u, F) = \min_v r_G(u, v)$ [4].

Остовом (каркасом) связного графа G называют связный подграф (дерево), содержащий все вершины графа. Остов $O(V, X_O)$, имеющий минимальную сумму длин ребер, называют кратчайшим.

Звездой (окрестностью) вершины v графа G называют подграф $S_v(G) = \langle X_v(G) \rangle$, порожденный множеством $X_v(G) \subseteq X$ всех ребер $x = (v, u)$, содер-

жащих v . Для вершины v обозначим через $X_v^l(G)$ множество кратчайших ребер $x_v = (v, u)$, имеющих минимальную длину l^v среди ребер $X_v(G)$.

Легко установить следующие свойства кратчайших остовов взвешенных графов.

1. Каждый кратчайший остов $O(V, XO)$ для каждой вершины v содержит хотя бы одно ребро $x_v \in X_v^l(G)$ минимальной длины l^v .

2. При удалении из кратчайшего остова $O(V, XO)$ ребра $x = (v, u)$ образуются два подостова $O^v(V^v, XO^v)$ и $O^u(V^u, XO^u)$, которые являются кратчайшими остовами для порожденных в G подграфов $\langle V^v \rangle$ и $\langle V^u \rangle$ соответственно, $v \in V^v, u \in V^u$.

Определим *отделимость вершины v* по ребру $x = (v, u)$ в G как величину

$$f_v^x(G) = l(x)/l^v, \text{ где } l^v = l(x_v), x_v \in X_v^l(G).$$

Определим *отделимость вершин ребра $x = (v, u)$* в G как величину

$$f_x(G) = \max(f_v^x(G), f_u^x(G)).$$

Из свойства 1 кратчайших остовов следует

Утверждение 1. Для любого ребра $x = (v, u)$ кратчайшего остова $O(V, XO)$ взвешенного графа G выполняется равенство $f_x(G) = f_x(O)$.

Разбиение $\tilde{V} = \{V_k\}, k = 1, \dots, K$, множества V вершин связного взвешенного графа G на подмножества V_k , для каждого из которых $G_k(V_k, X_k) = \langle V_k \rangle$ связен, называют таксономией, а V_k — таксонами (классами), $|V_k| = p_k, |X_k| = q_k$.

Вершинами — эталонами таксона V_k графа G — называют вершины v , имеющие наименьшее значение критерия $\sum_u r_G(v, u), u \in V_k$, или наибольшее значение критерия $\sum_u r_G(v, u), u \in V \setminus V_k$ [3].

Типичными вершинами таксона V_k графа G назовем вершины v , для которых $n_v > n_0$, где n_v — число вершин $u \in V_k$, находящихся от v на расстоянии, не превышающем величины $l_0 = \sum_x l(x)/q_k, x \in X_k$ (или $l_0 = \sum_v \sum_u r_G(v, u)/(p_k(p_k - 1)); v, u \in V_k$), а $n_0 = \sum_v n_v/p_k$.

В качестве оценки l_{ks} расстояния между таксонами V_k и V_s примем длину кратчайшего из ребер $x(k, s) = (v, u), v \in V_k, u \in V_s$, соединяющих в G вершины этих таксонов. Оценку l_k близости вершин в таксоне V_k определим как среднюю длину ребра кратчайшего остова O_k таксона. При $|V_k| = 1$ принимаем $l_k = \min_s l_{ks}$.

Тогда *отделимость f_{ks} таксонов V_k и V_s* можно определить такими, например, величинами:

$$f'_{ks} = (l_{ks}/l_k + l_{ks}/l_s)/2 \text{ или } f''_{ks} = f_{x(k,s)} l_{ks}.$$

На основе понятия отделимости таксонов введем оценку качества таксономии вершин взвешенного графа. Предполагаем, что качество таксономии тем выше, чем больше средняя отделимость близких друг к другу таксонов.

Связный взвешенный граф $G(V, X)$ с заданной таксономией $\tilde{V} = \{V_k\}, k = 1, \dots, K, K > 1$, преобразуем в граф $G'(V, X')$ следующим образом. Порожденные таксонами подграфы $\langle V_k \rangle$ заменим их кратчайшими остовами O_k и удалим из G все ребра, соединяющие пары остовов O_k и O_s , кроме одного кратчайшего ребра $x(k, s), k, s \in \{1, \dots, K\}$, имеющего $\min f_{ks}$. Получим граф G' , имеющий множество ребер $X' = X^0 \cup X'', X^0 = \cup_k X O_k, X'' = \{x(k, s)\}$. Граф G' преобразуем в граф $G''(\tilde{V}, X'')$, вершины которого соответствуют таксонам,

а ребрами являются $x(k, s)$. Для графа G'' построим кратчайший остов $O''(\tilde{V}, XO'')$.

Тогда *качество таксономии* \tilde{V} вершин связного взвешенного графа G можно оценить величиной

$$Q(\tilde{V}) = \sum f_{ks} / (K - 1), \quad x(k, s) \in XO''.$$

Легко видеть, что для функции $f(k) = (1/k) \sum_{i=1}^k y_i$ действительной переменной y_i справедливо равенство $(k + 1)(f(k + 1) - f(k)) = y_{k+1} - f(k)$. Отсюда следует

Утверждение 2. Для упорядоченной по невозрастанию последовательности $y^0 = (y_1, y_2, \dots, y_i, \dots)$ значений действительной переменной $y_i \geq 0$ имеет место $\max_k f(k) = f(1)$.

Действительно, для функции $f(k)$ и последовательности y^0 выполняются неравенства $f(k) \geq y_k \geq y_{k+1}$. Поэтому $f(k + 1) - f(k) \leq 0$ для всех k и $\max_k f(k) = f(1)$.

Граф F называют общим подграфом помеченных графов G и H , если в них существуют подграфы G' и H' , изоморфные графу F . Общий подграф двух графов называют наибольшим, если он содержит максимальное число вершин или ребер. Большинство известных алгоритмов нахождения наибольших общих подграфов пары графов основаны на использовании алгоритма поиска с возвратом [6].

Меру структурной близости (подобия) пары помеченных графов $G_i(V_i, X_i)$ и $G_j(V_j, X_j)$ можно оценивать с помощью количественных характеристик, например, расстоянием между ними:

$$r_{ij} = cp(i, j)/p + (1 - c)q(i, j)/q,$$

где $0 \leq c \leq 1$, $p_i = |V_i|$, $p_j = |V_j|$, $q_i = |X_i|$, $q_j = |X_j|$, $p = p_i + p_j$, $q = q_i + q_j$, $p(i, j) = p_i + p_j - 2p_{ij}$, $q(i, j) = q_i + q_j - 2q_{ij}$; $G_{ij}(V_{ij}, X_{ij})$ — наибольший общий подграф графов G_i и G_j , имеющий наибольшие значения p_{ij} или q_{ij} (пересечение или перекрытие графов G_i и G_j [4]).

Утверждение 3. Расстояние r_{ij} между помеченными графами G_i и G_j не является метрикой.

Покажем это на конкретном примере. Пусть имеются графы G_i, G_j, G_k и $p_i = p_k = p > 1$, $p_{ij} = p_{jk} = p' > 0$, $p_{ik} = 0$, $p_j = 2p' + 1$, $p_i > p' + p_{ik}$, $p_k > p' + p_{ik}$, $p_{ik} = 1$, $c = 1$. Неравенство $r_{ij} + r_{jk} < r_{ik}$ в этом случае равносильно неравенству $p(p + 3)/(p - 1) < p_j$, которое выполняется, например, при $p = 8$ и $p_j = 13$. Отсюда следует, что величина расстояния r_{ij} не удовлетворяет аксиоме треугольника.

Выборкой объема p из некоторого множества объектов называют его подмножество мощности p , разбитое на $K \geq 1$ классов по их свойствам. В соответствии с принятой методикой прогноза свойств объектов их выборки разделены на обучающие и контрольные. Объединение выборок ОБ и КВ обозначим как $B = OB \cup KB$.

Для заданной выборки объектов можно найти матрицу $M(G)$ расстояний взвешенного графа $G(V, X)$, вершины v, u, \dots которого соответствуют объектам выборки, а длина $l(x)$ каждого ребра $x = (v, u)$ равна расстоянию между объектами — вершинами ребра: $l(x) = r^{vu}$ — расстоянию в пространстве признаков или $l(x) = r_{vu}$ — расстоянию между графами G_v и G_u структур объектов.

Взвешенные графы обучающей и контрольной выборок обозначим соответственно как GT и GC .

Множество вершин $GT(V_T, X_T)$, $|V_T| = p_T$, разбито на классы V_k , $k = 1, \dots, K$, $\tilde{V} = \{V_k\}$, $|V_k| = p_k$, $\sum_k p_k = p_T > 2$, в соответствии с классами свойств объектов ОБ.

Вершины графа $GC(V_C, X_C)$ составляют отдельный 0-й класс $V_0 = V_C$, $|V_C| = p_C \geq 0$. Предполагается, что 0-й класс вершин соответствует множеству объектов, содержащему объекты со свойствами классов ОБ, «сомнительные» объекты и объекты со свойствами, не представленными в ОБ.

Определим некоторые способы построения взвешенного графа G объединенной выборки В. При всех преобразованиях графов отделимость вершин каждого ребра считаем неизменной и равной отделимости в графе G^1 .

Граф G^1 является соединением графов GT и GC , т. е. состоит из объединения графов GT и GC и всех ребер $X_{TC} = \{x \mid x = (v, u), v \in V_T, u \in V_C\}$, соединяющих V_T и V_C : $G^1 = GT + GC = GT \cup GC \cup \langle X_{TC} \rangle$.

Граф GT преобразуем в граф $GT'(V_T, X_T')$, $X_T' = X_T^0 \cup X_T''$, $X_T^0 = \cup_k X_{TO_k}$, $X_T'' = \{x(k, s)\}$, где X_{TO_k} — множество ребер кратчайшего остова OT_k подграфа $\langle V_k \rangle$, порожденного k -м классом вершин; $x(k, s)$ — кратчайшее ребро, соединяющее вершины из k -го и s -го классов. Граф G^2 определим как соединение $G^2 = GT' + GC$.

Граф GT' преобразуем в граф $GT^0(V_T, X_T^0)$ удалением из X_T' множества ребер X_T'' . Граф G^3 определим как соединение $G^3 = GT^0 + GC$.

Для каждой вершины $u \in V_C$ и каждого класса V_k выберем одно кратчайшее ребро $x(u, V_k)$ из ребер, соединяющих вершину u с вершинами из V_k . Обозначим $X_{TC}^* = \cup_k \cup_u x(u, V_k)$. Граф GT' преобразуем в граф $GT''(\tilde{V}, X_T'')$, вершины которого соответствуют классам, а ребрами являются кратчайшие ребра $x(k, s)$. Граф G^4 определим как соединение $G^4 = GT'' + GC = GT'' \cup \cup GC \cup \langle X_{TC}^* \rangle$.

Граф GT'' преобразуем в граф GT_0 удалением из него всех его ребер X_T'' , т. е. граф GT_0 состоит из K изолированных вершин V_k . Граф G^5 определим как соединение $G^5 = GT_0 + GC$, ребра которого составляют множество $X_C \cup X_{TC}^*$.

Алгоритмы классификации вершин взвешенного графа. Приведенные ниже алгоритмы применяются ко взвешенному графу $G(V, X)$, являющемуся соединением графов ОБ и КВ: $G = GT + GC$.

В зависимости от конкретных особенностей проведения компьютерных экспериментов по классификации объектов КВ в качестве графа G может быть использован один из видов $G^1 - G^5$ взвешенных графов объединенной выборки В.

Если при таксономии вершин графа G допускается разбиение заданных классов V_k на подклассы вершин, то возможно использование графов G^1, G^2, G^3 . Если такое разбиение недопустимо, то необходимо использовать графы G^4, G^5 .

При заданной матрице расстояний MG взвешенного графа G объединенной выборки В вычислительная сложность приведенных алгоритмов таксономии и классификации не превышает величины $C(p_T + p_C)^2$, $C > 0$.

Алгоритм DOUBT нахождения «сомнительных» вершин в каждом классе вершин графа. В каждом классе V_k разбиения вершин взвешенного графа G находятся «сомнительные» вершины — объекты по критериям, учитывающим расстояния от исследуемой вершины до смежных вершин своего класса, наибольшие расстояния внутри этого класса, расстояния до ближайших вершин из других классов и отделимости вершин.

Множество вершин класса V_k , кроме вершин графа GT , может содержать контрольные вершины $u \in V_C$, отнесенные при классификации к классу V_k .

1. Для каждого класса вершин V_k находится кратчайший остов O_k .

2. Вершина $v \in V_k$ считается «сомнительной», если для нее выполняется условие одного из критериев:

1) $\max_u l(v, u) > \min_w l(v, w)$, где u — вершины звезды $S_v(O_k)$, $w \in V \setminus V_k$;

2) $\max_u l(v, u) > \max_x l(x)$, где u — вершины звезды $S_v(O_k)$ и ребро x принадлежит подостову O_k'' , полученному после удаления наибольшего ребра (v, u) и содержащему u ;

3) $\max_u f_v^x > \min_w f_v^{x'}$, где $x = (v, u)$, $u \in V_k$, $x' = (v, w)$, $w \in V \setminus V_k$.

Алгоритм CLASSEQ последовательной классификации объектов КВ. При классификации отдельного объекта u КВ используется взвешенный граф объединенной выборки $G = GT + GC$, в котором $V_C = \{u\}$, $p_C = 1$.

1. Для каждой вершины — объекта u из КВ среди всех ребер $x = (u, v)$, $v \in V_T$, находится ребро $x^* = (u, v^*)$, $v^* \in V_k \in \tilde{V}_T$, имеющее наименьшее значение величины $l(x)f_x$.

Вершина u относится к классу V_k , если u не является «сомнительной» в этом классе, иначе u выделяется в отдельный 0-й класс V_0 .

2. Вычисляются w — относительная ошибка последовательной классификации всех вершин КВ и качество классификации $f = 1 - (w/w')$, где $w' = K/(K + 1)$ является относительной ошибкой классификации при равновероятном отнесении вершин КВ к $K \geq 1$ классам ОВ или к отдельному 0-му классу, $-1 \leq f \leq 1$.

Алгоритм VEGSEP таксономии вершин взвешенного графа. Во взвешенном графе $G(V, X)$ производится таксономия на классы близких вершин, состоящая в последовательном удалении из кратчайшего остова $O(V, XO)$ ребер с наибольшими значениями отделимости вершин.

1. Для графа $G(V, X)$ находится кратчайший остов $O(V, XO)$.

2. Для каждого ребра $x = (v, u) \in XO$ остова O определяется отделимость его вершин f_x и все ребра упорядочиваются по невозрастанию f_x (или $l(x)$).

3. В полученном порядке $q(1)$ ребер, имеющих $f_x > 1$, по очереди удаляются из XO .

Удаление множества ребер $X(n) = \{x_i\}$, $i = 1, \dots, n$, $n \leq q(1)$, приводит к разбиению $\tilde{V}(n) = \{V^t\}$, $t = 1, \dots, n + 1$, остова O на $(n + 1)$ подостовов $O^t (V^t, XO^t)$. Каждый подостов O^t является кратчайшим остовом порожденного подграфа $\langle V^t \rangle$ и соответствует таксону (классу) V^t .

4. Для каждого из полученных разбиений $\tilde{V}(n)$, $1 \leq n \leq q(1)$, вершин вычисляется качество таксономии $Q(\tilde{V}(n)) = (1/n) \sum f_{ks}$, $x(k, s) \in X(n)$.

5. Результатом является таксономия $\tilde{V}^*(n)$, имеющая наибольшее качество $Q(\tilde{V}^*(n)) = \max_n Q(\tilde{V}(n))$.

Алгоритм CLASTAX классификации вершин КВ с помощью таксономии. Для объединенной выборки B с известным разбиением объектов ОВ на классы свойств V_k , $k = 1, \dots, K$, $|V_k| = p_k$, и 0-м классом $V_0 = V_C$, $|V_C| = p_0$, объектов КВ задается взвешенный граф G с соответствующим разбиением вершин на $K + 1$ классов V_s , $s = 0, 1, \dots, K$.

1. Алгоритмом VEGSEP производится таксономия вершин взвешенного графа G на классы U^t , $t = 1, \dots, T$, $|U^t| = p^t$, близких вершин.

2. Для всех $s = 0, 1, \dots, K$ и $t = 1, \dots, T$ находятся пересечения $Z_s^t = V_s \cap U^t$ — множества вершин, принадлежащих классу близости U^t и входящих в класс свойств V_s , $|Z_s^t| = p_s^t$.

3. Вершина $u \in V_C$, входящая в класс близости $U^t = \cup_s Z_s^t$, относится к классу свойств V_s , имеющему наибольшую оценку вероятности p_s^t/p^t .

Результаты испытаний. Предложенные алгоритмы были испытаны в системе ВАСС при прогнозе свойств химических соединений семейства акарицидов.

Для описания молекул химических соединений использовались как их структурные характеристики — молекулярные графы, так и количественные — топологические индексы графов.

Общее количество акарицидов с известными значениями токсичности, имеющееся в базе данных системы, составило 98 соединений.

Классы токсичности соединений ОВ и КВ задавались по значениям средней смертельной дозы ЛД₅₀ (мг/кг) при их введении в желудок крысы. Проводилось 14 экспериментов. Объем ОВ выбирался в пределах 17 ÷ 62 соединений, объем КВ — 6 ÷ 21, число классов свойств ОВ — 2 ÷ 5, при этом в ряде экспериментов в ОВ и КВ включались соединения, не обладающие акарицидной активностью [7]. В этих экспериментах качество классификации при последовательном прогнозе свойств отдельных соединений КВ и прогнозе свойств с помощью таксономии для графа G^1 в среднем составило примерно 0,65. Это значение качества соответствует при двух классах ОВ относительной ошибке классификации, равной 22 %.

Разработанные алгоритмы могут найти применение при решении задач классификации объектов, выборки которых можно представить в виде взвешенных графов. В частности, приведенные алгоритмы могут быть использованы для повышения эффективности компьютерного дизайна химических соединений, основанного на исследовании взаимосвязи структур и свойств соединений. Целью этих исследований является создание и применение биологически активных и экологически безопасных веществ, получение эффективных лекарственных препаратов, создание малотоксичных химических средств защиты растений и т. д.

СПИСОК ЛИТЕРАТУРЫ

1. Загоруйко Н. Г., Елкина В. Н., Лбов Г. С. Алгоритмы обнаружения эмпирических закономерностей. Новосибирск: Наука, 1985.
2. Стьюпер Э., Брюгге У., Джурс П. Машинный анализ связи химической структуры и биологической активности. М.: Мир, 1982.
3. Макаров Л. И. Алгоритм таксономии вершин взвешенного графа // Математические методы в химической информатике. Новосибирск, 1991. Вып. 140: Вычислительные системы. С. 130.
4. Макаров Л. И. Некоторые характеристики и описание молекулярных (помеченных) графов и их подграфов // Приложения теории графов в химии. Новосибирск, 1993. Вып. 149: Вычислительные системы. С. 109.
5. Макаров Л. И., Скоробогатов В. А. Комплекс программ для исследования зависимости «структура—свойство» химических соединений // Алгоритмический анализ графов и его применения. Новосибирск, 1988. Вып. 127: Вычислительные системы. С. 92.
6. Рейнгольд Э., Нивергельт Ю., Део Н. Комбинаторные алгоритмы. Теория и практика. М.: Мир, 1980.
7. Макаров Л. И. Методика нахождения информативного набора индексов молекулярных графов при прогнозе свойств химических соединений // Журнал структурной химии. 1997. 38, № 4.

Поступила в редакцию 15 апреля 1997 г.