

УДК 547:541.6:51-7

**МЕТРИЧЕСКИЕ СВОЙСТВА ФУНКЦИЙ РАССТОЯНИЙ МЕЖДУ  
МОЛЕКУЛЯРНЫМИ ГРАФАМИ**

© 2007 Л.И. Макаров\*

*Институт математики им. С.Л. Соболева СО РАН, Новосибирск**Статья поступила 3 июля 2006 г.*

Рассмотрены метрические свойства некоторых функций расстояний между молекулярными (помеченными) графами, зависящих от размера их наибольших общих подграфов. Приведены рекомендации по использованию таких функций при исследовании выборок химических соединений.

**Ключевые слова:** молекулярный граф, расстояние.

При компьютерной обработке данных о молекулярных структурах исследуемой выборки химических соединений в качестве моделей как структур соединений, так и заданной выборки часто используют помеченные графы, элементам которых (вершинам и ребрам) приписаны некоторые метки — символы, числа, списки и т.д.

Моделью структуры химического соединения служит молекулярный граф, вершины которого соответствуют атомам или группам атомов соединения, а ребра — химическим связям между его атомами. При описании молекулярной структуры каждому ее фрагменту соответствует определенный подграф молекулярного графа. Моделью выборки химических соединений служит взвешенный граф, вершины которого соответствуют соединениям, а длина (вес) ребер определяется расстояниями между их молекулярными графами.

В качестве расстояния используют количественную оценку степени структурного различия молекулярных графов, зависящую от размера их наибольшего общего подграфа [2—7]. Чем больше общий подграф (фрагмент структуры), тем больше структурное подобие соединений и тем меньше расстояние между ними.

Графовые модели молекулярных структур находят применение в различных областях исследований химических соединений, в частности, в исследованиях по проблемам количественной зависимости свойств соединений от их структуры и взаимосвязи структуры соединений и их спектральных характеристик [1—6]. Целью компьютерного анализа зависимости "структура—свойство" некоторого семейства соединений является нахождение общих структурных фрагментов, определяющих их свойства, и классификация (прогноз свойств) исследуемых соединений. Другая проблема состоит в нахождении фрагментного состава молекулы исследуемого соединения по его спектральным характеристикам, например, по его инфракрасному (ИК) спектру, нахождении фрагментов, определяющих особенности спектра, и конструировании структуры его молекулы, исходя из полученных фрагментов.

Такие исследования проводятся в Новосибирском институте органической химии (НИОХ) СО РАН, имеющем базу данных о ИК спектрах и структурах более чем 30 000 химических соединений [3, 5]. В базе данных НИОХ СО РАН применяется бинарное векторное описание молекулярных графов. Каждая компонента вектора молекулярного графа соответствует определенному фрагменту, а ее значение равно единице, если этот фрагмент входит в граф, и равно нулю, если фрагмент в графе не содержится.

---

\* E-mail: makarov@math.nsc.ru

### ПОДОБИЕ МОЛЕКУЛЯРНЫХ ГРАФОВ

Пусть  $G(V, X)$  — конечный, помеченный, связный, неориентированный граф без петель и кратных ребер, имеющий множество вершин  $V = \{v, u, \dots\}$ ,  $|V| = p > 0$ , и множество ребер  $X = \{x, y, \dots\}$ ,  $x = (v, u)$ ,  $|X| = q > 0$ .

Граф  $G'(V', X')$  называют подграфом (фрагментом) графа  $G$ ,  $G' \subseteq G$ ,  $p' > 0$ ,  $q' \geq 0$ , если его множества  $V'$  и  $X'$  являются подмножествами множеств  $V$  и  $X$ .

Далее произвольный подграф  $G'$  графа  $G$  будем называть частичным. Подграфом, порожденным в графе  $G$  множеством вершин  $V'$ , называют подграф, в котором множество ребер  $X'$  состоит из всех ребер, соединяющих в  $G$  вершины из  $V'$ .

Два помеченных графа  $G$  и  $H$  называют изоморфными, если между их множествами вершин существует взаимно однозначное отображение, сохраняющее метки вершин и их смежность по одинаково помеченным ребрам.

Граф  $F$  называют общим подграфом графов  $G$  и  $H$ , если в них существуют подграфы  $G'$  и  $H'$ , изоморфные графу  $F$ .

При нахождении общих подграфов молекулярных графов можно задавать условия эквивалентности меток в смысле изоморфизма. Например, для количественной величины можно задать интервалы ее значений, внутри которых все значения считаются равными. Для номинальных меток можно задать списки эквивалентных имен, например, список галогенов.

Общие подграфы молекулярных графов могут быть различных типов, например, подграфы помеченные или без меток, связные или несвязные, порожденные или частичные; подграфы, общие для графов нескольких соединений и наибольшие по числу вершин или ребер, и т.д. Выбираемый тип подграфов определяется спецификой задачи.

Общий подграф двух графов называют наибольшим, если этот подграф содержит максимальное число вершин или ребер среди других общих подграфов.

Пересечением графов  $G$  и  $H$  назовем наибольший по числу вершин общий подграф графов  $G$  и  $H$ , изоморфный их порожденным подграфам  $G'$  и  $H'$ . Перекрытием назовем наибольший по числу ребер общий частичный подграф графов  $G$  и  $H$ .

Меру структурного подобия молекулярных графов  $G_i(V_i, X_i)$  и  $G_j(V_j, X_j)$  можно оценивать с помощью количественных характеристик, зависящих от числа вершин или ребер их наибольшего общего подграфа. Наибольший общий подграф графов  $G_i$  и  $G_j$  обозначим через  $G_{ij}(V_{ij}, X_{ij})$ ,  $|V_i| = p_i > 0$ ,  $|X_i| = q_i > 0$ ,  $|V_j| = p_j > 0$ ,  $|X_j| = q_j > 0$ ,  $|V_{ij}| = p_{ij} \geq 0$ ,  $|X_{ij}| = q_{ij} \geq 0$ .

Следует заметить, что два графа могут иметь несколько неизоморфных наибольших общих подграфов, но в силу их определения все пересечения имеют одинаковое количество вершин, а все перекрытия — одинаковое число ребер. Однако при этом пересечения могут иметь разное число ребер, а перекрытия — разное число вершин. Поэтому в качестве наибольшего общего подграфа необходимо выбирать либо пересечение графов с максимальным числом ребер, либо перекрытие с максимальным числом вершин. Тогда для каждой пары графов будут зафиксированы величины  $p_{ij}$ ,  $q_{ij}$ . Это позволяет в качестве оценок подобия графов использовать функции расстояний между графами, зависящие от этих величин.

Рассмотрим некоторые функции расстояний: абсолютные расстояния, зависящие от размеров графов и их подграфов,

$$r_{ij}^V = p_i + p_j - 2p_{ij}, \quad r_{ij}^X = q_i + q_j - 2q_{ij}, \quad r_{ij}^{VX} = (r_{ij}^V + r_{ij}^X) / 2,$$

и относительные (нормированные) расстояния

$$R_{ij}^{V'} = r_{ij}^V / (p_i + p_j) \quad \text{и} \quad R_{ij}^{X'} = r_{ij}^X / (q_i + q_j),$$

$$\text{или} \quad R_{ij}^V = r_{ij}^V / (p_i + p_j - p_{ij}), \quad R_{ij}^X = r_{ij}^X / (q_i + q_j - q_{ij}), \quad R_{ij}^{VX} = (R_{ij}^V + R_{ij}^X) / 2.$$

При практическом использовании функций расстояний для исследования выборки соединений обычно требуется, чтобы эти функции удовлетворяли условиям метрики.

Метрикой на некотором множестве  $\{A, B, C, \dots\}$  называют числовую функцию расстояния  $f(A, B) \geq 0$ , удовлетворяющую условиям:

1.  $f(A, B) = 0$  тогда и только тогда, когда  $A = B$ ;
2.  $f(A, B) = f(B, A)$ ;
3.  $f(A, C) + f(C, B) \geq f(A, B)$  — неравенство треугольника.

Использование функций расстояний, не удовлетворяющих этим условиям, может приводить к неоднозначности результатов исследований выборок соединений.

Например, если функция расстояния не удовлетворяет условиям метрики, т.е.

1.  $f(A, B) = 0$  при  $A \neq B$ , то соединения с разными структурами будут считаться совпадающими по структуре;
2.  $f(A, B) \neq f(B, A)$ , то количественное различие (подобие) структур будет зависеть от порядка вычислений;
3.  $f(A, C) + f(C, B) < f(A, B)$ , то оценка различия (подобия) структур двух соединений  $A$  и  $B$  будет зависеть от наличия или отсутствия в выборке некоторого третьего соединения  $C$ .

Поэтому применение функций расстояния, в которых не выполняются условия метрики, нецелесообразно.

### МЕТРИЧЕСКИЕ СВОЙСТВА ФУНКЦИЙ РАССТОЯНИЙ МЕЖДУ МОЛЕКУЛЯРНЫМИ ГРАФАМИ

Рассмотрим некоторую выборку химических соединений. Пусть этой выборке соответствует множество  $G^0 = \{G_i\}$  молекулярных графов соединений  $G_i(V_i, X_i)$  и множество  $G^1 = \{G_{ij}\}$  их наибольших общих подграфов  $G_{ij}(V_{ij}, X_{ij})$ ,  $i, j \in \{1, 2, \dots, M\}$ .

Множеству графов  $G^0$  соответствуют множества вершин  $V^0 = \{V_i\}$  и ребер  $X^0 = \{X_i\}$ , а множеству  $G^1$  их подграфов — множества  $V^1 = \{V_{ij}\}$  и  $X^1 = \{X_{ij}\}$ .

Запишем систему неравенств, справедливых для любой тройки элементов из множеств  $V^0$  или  $X^0$ . При этом совместный наибольший общий подграф трех графов  $G_i, G_j, G_k$  обозначим как  $G_{ijk}(V_{ijk}, X_{ijk})$ ,  $|V_{ijk}| = p_{ijk} \geq 0$ ,  $|X_{ijk}| = q_{ijk} \geq 0$ ,  $i, j, k \in \{1, 2, \dots, M\}$ ,  $M \geq 3$ .

Поскольку для каждого графа в функциях расстояний используются только количества вершин  $p_i, p_{ij}$  или число ребер  $q_i, q_{ij}$ , то каждому графу можно сопоставить некоторое множество  $W$ , которое представляет его множества  $V$  или  $X$  в зависимости от величин, используемых в функции расстояния:  $W \in \{V, X\}$ ,  $r_{ij} \in \{r_{ij}^V, r_{ij}^X\}$ ,  $R'_{ij} \in \{R'_{ij}^V, R'_{ij}^X\}$ ,  $R_{ij} \in \{R_{ij}^V, R_{ij}^X\}$ ,  $0 \leq R'_{ij}, R_{ij} \leq 1$ .

Обозначим  $|W_i| = N_i$ ,  $|W_{ij}| = N_{ij}$ ,  $|W_{ijk}| = N_{ijk}$ , т.е. для  $W_i = V_i$ :  $N_i = p_i$ ,  $N_{ij} = p_{ij}$ ,  $N_{ijk} = p_{ijk}$ , а для  $W_i = X_i$ :  $N_i = q_i$ ,  $N_{ij} = q_{ij}$ ,  $N_{ijk} = q_{ijk}$ .

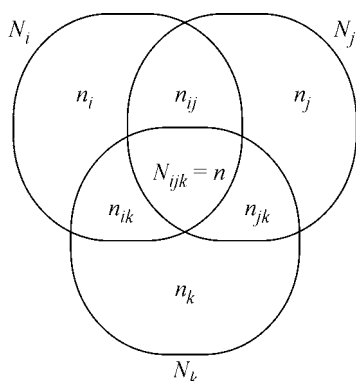
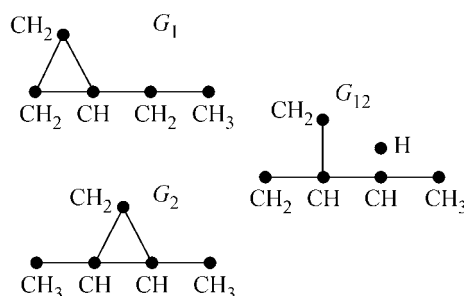
Тогда для пересечений множеств справедливы неравенства (рис. 1).

$$\begin{aligned}
 &N_i, N_j, N_k > 0, \\
 &N_i = n_i + n_{ij} + n_{ik} + N_{ijk}, \quad N_j = n_j + n_{ij} + n_{jk} + N_{ijk}, \quad N_k = n_k + n_{ik} + n_{jk} + N_{ijk}, \\
 &0 \leq n_i \leq N_i, \quad 0 \leq n_j \leq N_j, \quad 0 \leq n_k \leq N_k, \\
 &0 \leq n_{ij} \leq \min(N_i, N_j), \quad 0 \leq n_{ik} \leq \min(N_i, N_k), \quad 0 \leq n_{jk} \leq \min(N_j, N_k), \\
 &0 \leq N_{ijk} \leq \min(N_i, N_j, N_k), \\
 &0 \leq N_{ij} = n_{ij} + N_{ijk} \leq \min(N_i, N_j), \quad 0 \leq N_{ik} = n_{ik} + N_{ijk} \leq \min(N_i, N_k), \\
 &0 \leq N_{jk} = n_{jk} + N_{ijk} \leq \min(N_j, N_k).
 \end{aligned}$$

Для упрощения вида формул введем обозначения:  $n = N_{ijk}$ ,  $h = n_{ij} + n_{ik} + n_{jk} + n \geq 0$ ,  $s_{ij} = N_i + N_j - N_{ij} = n_i + n_j + h$ ,  $s_{ik} = N_i + N_k - N_{ik} = n_i + n_k + h$ ,  $s_{jk} = N_j + N_k - N_{jk} = n_j + n_k + h$ ,  $0 < s_{ij} \leq N_i + N_j$ ,  $0 < s_{ik} \leq N_i + N_k$ ,  $0 < s_{jk} \leq N_j + N_k$ .

Рассмотрим выполнение условий метрики для функции расстояния  $r_{ij} = N_i + N_j - 2N_{ij}$ .

1.  $r_{ii} = 0$ .
2.  $r_{ij} = r_{ji}$ .
3.  $r_{ik} + r_{kj} = N_i + N_k - 2N_{ik} + N_k + N_j - 2N_{jk} + 2N_{ij} - 2N_{ij} = N_i + N_j - 2N_{ij} + 2(N_k + N_{ij} - N_{jk} - N_{ik}) = r_{ij} + 2(n_k + n_{ij}) \geq r_{ij}$ .

Рис. 1. Пересечения множеств  $W$ Рис. 2. Пример циклоалканов, для которых  $r_{12}^V = 0$ 

Для расстояния  $r_{ij}^V$  выполняются условия 2 и 3 метрики, а условие 1 выполняется не всегда. Возможны случаи, когда  $p_i = p_{ij} = p_j$  и  $r_{ij}^V = 0$ , но графы  $G_i$  и  $G_j$  неизоморфны.

На рис. 2 представлены два неизоморфных молекулярных графа  $G_1$  и  $G_2$  циклоалканов, имеющих наибольший общий подграф  $G_{12}$ . Каждая метка  $\text{CH}_l$ ,  $l = 1, 2, 3$ , соответствует подграфу, соединяющему атом углерода  $C$  с  $l$  атомами водорода  $H$ . Число вершин графов  $p_1 = p_2 = p_{12} = 15$ , из них 5 соответствуют атомам  $C$ , а 10 — атомам  $H$ . Отсюда следует, что  $r_{12}^V = 0$ .

Если  $r_{ij}^V = 0$  для неизоморфных графов  $G_i$  и  $G_j$ , то для любого графа  $G_k$  наибольшие общие подграфы  $G_{ik}$  и  $G_{jk}$  имеют  $p_{ik} = p_{jk}$  и  $r_{ik}^V = r_{jk}^V$ , поэтому условие 3 метрики выполняется и в этом случае.

Поскольку каждое ребро графа определяется парой помеченных вершин и имеет собственную метку, то список ребер графа является его однозначным представлением и  $r_{ij}^X = 0$ ,  $i \neq j$ , тогда и только тогда, когда  $q_i = q_{ij} = q_j$ , т.е. когда графы  $G_i$  и  $G_j$  изоморфны.

Таким образом, для расстояния  $r_{ij}^X$  выполняются все три условия метрики.

Следует отметить, что в примере на рис. 2 число ребер графов  $q_1 = q_2 = 15$ , а  $q_{12} = 13$ .

Отсюда следует

**утверждение 1:** расстояние  $r_{ij}^X$  между молекулярными графами  $G_i$  и  $G_j$  является метрикой, а для расстояния  $r_{ij}^V$  может не выполняться условие 1 метрики.

**Утверждение 2:** для относительного расстояния  $R'_{ij}$  между молекулярными графами  $G_i$  и  $G_j$  может не выполняться условие 3 метрики.

Покажем это на конкретном примере. Пусть имеются графы  $G_i$ ,  $G_j$ ,  $G_k$  и  $N_i = N_j = N > 1$ ,  $N_{ik} = N_{jk} = N' > 0$ ,  $N_k = 2N' + 1$ ,  $N_i > N' + N_{ij}$ ,  $N_j > N' + N_{ij}$ ,  $N_{ij} = 1$ ,  $n = 0$ . Неравенство  $R'_{ik} + R'_{jk} < R'_{ij}$  в этом случае равносильно неравенству  $N(N+3)/(N-1) < N_k$ , которое выполняется, например, при  $N = 8$  и  $N_k = 13$ . Отсюда следует, что расстояние  $R'_{ij}$  не удовлетворяет условию треугольника, т.е. не является метрикой.

**Утверждение 3:** относительное (нормированное) расстояние  $R_{ij}^X$  между молекулярными графами  $G_i$  и  $G_j$  является метрикой, а для расстояния  $R_{ij}^V$  может не выполняться условие 1 метрики.

Действительно,  $R_{ij} = r_{ij}/s_{ij} = 1 - N_{ij}/s_{ij}$  и

1.  $R_{ii} = 0$ .

Но для  $i \neq j$  может быть  $R_{ij}^V = 0$  (см. рис. 2), а  $R_{ij}^X = 0$  тогда и только, когда графы изоморфны.

$$2. R_{ij} = R_{ji}.$$

3. Неравенство треугольника  $(R_{ik} + R_{kj}) \geq R_{ij}$  равносильно неравенству

$$z = (N_{ik}/s_{ik} + N_{kj}/s_{kj} - N_{ij}/s_{ij}) \leq 1 \quad \text{или} \quad z = \left( \frac{n_{ik} + n}{n_i + n_k + h} + \frac{n_{jk} + n}{n_j + n_k + h} - \frac{n_{ij} + n}{n_i + n_j + h} \right) \leq 1.$$

Легко видеть, что

$$1) z \leq \left( \frac{n_{ik} + n}{n_i + h} + \frac{n_{jk} + n}{n_j + h} - \frac{n_{ij} + n}{n_i + n_j + h} \right) = z_1 = z|_{n_k=0},$$

$$2) z_1 \leq \left( \frac{n_{ik} + n}{n_i + t} + \frac{n_{jk} + n}{n_j + t} - \frac{n}{n_i + n_j + t} \right) = z_2 = z_1|_{n_{ij}=0}, \quad \text{где } t = h|_{n_{ij}=0} = n_{ik} + n_{jk} + n,$$

$$\text{так как } \frac{n}{n_i + n_j + t} \leq \frac{n_{ij} + n}{n_i + n_j + h}.$$

Отсюда следует, что при выполнении неравенства  $z_2 \leq 1$  выполняется и неравенство  $z \leq 1$ .

Из неравенства  $z_2 \leq 1$  имеем

$$(n_{ik} + n)(n_j + t) + (n_{jk} + n)(n_i + t) \leq (n_i + t)(n_j + t) + C, \quad \text{где } C = \frac{n(n_i + t)(n_j + t)}{n_i + n_j + t};$$

$$n_{ik}n_j + n_{ik}t + nn_j + nt + n_{jk}n_i + n_{jk}t + nn_i + nt \leq n_in_j + n_it + tn_j + t^2 + C,$$

$$t(t + n) + n_i(t - n_{ik}) + n_j(t - n_{jk}) \leq n_it + n_jt + n_in_j + t^2 + C,$$

$$tn \leq n_in_j + n_in_{ik} + n_jn_{jk} + C,$$

$$tn(n_i + n_j + t) \leq D + n(n_i + t)(n_j + t), \quad \text{где } D = (n_in_j + n_in_{ik} + n_jn_{jk})(n_i + n_j + t); \quad 0 \leq D + nn_in_j.$$

Последнее неравенство в области определения переменных выполняется. Следовательно, выполняется неравенство треугольника, и функция расстояния  $R_{ij}^X$  является метрикой.

Из утверждений 1 и 3 непосредственно следует

**утверждение 4:** функции расстояний между графами  $r_{ij}^{VX}$  и  $R_{ij}^{VX}$  являются метриками.

Если задана функция расстояния между молекулярными графами, то для исследуемой выборки соединений можно найти матрицу расстояний и построить полный взвешенный граф  $G_L(V, X)$ , вершины которого соответствуют соединениям выборки, а длины ребер равны расстояниям между молекулярными графами соединений. Взвешенный граф выборки используется для ее таксономии (разбиения) на классы структурно подобных соединений при прогнозе свойств соединений по их структуре в компьютерной системе анализа и классификации биологически активных химических соединений (ВАСС), разработанной в Институте математики СО РАН [2, 6] и при восстановлении структуры соединения по его ИК спектру в НИОХ СО РАН [3, 5].

В алгоритме таксономии в системе ВАСС используется понятие отделимости вершин ребра взвешенного графа выборки, построенного по функции  $R_{ij}^{VX}$ . Отделимость вершин  $v$  и  $u$  ребра  $x = (v, u)$  определяется как величина  $f_x = f(v, u) = \max(l_x/l_v, l_x/l_u)$ , где  $l_x = l(v, u)$  — длина ребра  $x$ ,  $l_v$  (или  $l_u$ ) — наименьшая положительная длина ребра, инцидентного вершине  $v$  (или  $u$ ). Исходный взвешенный граф выборки  $G_L$  может быть преобразован во взвешенный граф  $G_f$ , в котором длины ребер заменены на отделимости их вершин.

Легко видеть, что даже если в графе  $G_L$  для любой тройки вершин выполнены все условия метрики, то в графе  $G_f$  для некоторых троек вершин может не выполняться неравенство треугольника. Например, пусть в графе  $G_L$  есть тройка вершин  $(v, u, w)$ , соединяющие их ребра

имеют длины  $l(v, u) = 5$ ,  $l(v, w) = l(u, w) = 3$  и кратчайшие ребра, инцидентные вершинам, имеют длины  $l(v, v_1) = 1$ ,  $l(u, u_1) = l(w, w_1) = 3$ . Тогда отделимости ребер равны  $f(v, u) = 5$ ,  $f(v, w) = 3$ ,  $f(u, w) = 1$  и  $f(v, w) + f(w, u) < f(v, u)$ .

Аналогичный пример можно построить и для  $\lambda$ -расстояния [ 8 ].

В алгоритме таксономии системы ВАСС разделение выборки на таксоны производится по кратчайшему остову графа  $G_L$ , поэтому неметричность функции отделимости в этом случае не существенна.

#### ЗАКЛЮЧЕНИЕ

Как показано выше, применение функций расстояния, для которых не выполняются условия метрики, может приводить к неоднозначности результатов исследований. Однако на практике приходится использовать функции расстояния, для которых не всегда выполняется условие 1. Это обусловлено требованием высокой скорости (малого времени и памяти) обработки данных о структурах, что приводит к применению приближенных описаний структур соединений, например, векторному описанию. При этом обычно для выполнения условия 1 из выборки устраняют по одному соединению каждой пары, для которой это условие нарушается, или задают для этой пары расстояние, не превышающее минимального положительного расстояния между графами этой выборки.

Расстояние  $r_{ij}^V$  для оценки структурного подобия графов применять не всегда целесообразно. Например, для структурно разных графов  $G_i$  и  $G_j$ , не имеющих общего подграфа ( $p_{ij} = 0$ ), расстояние  $r_{ij}^V$  меньше расстояния  $r_{ik}^V$  между структурно подобными графами  $G_i$  и  $G_k$ , имеющими общий подграф ( $p_{ik} > 0$ ), в случае, если  $p_j < p_k - 2p_{ik}$ . При этом для относительного расстояния выполняется  $R_{ij}^V = 1 > R_{ik}^V = 1 - p_{ik}/s_{ij}$ . Поэтому для оценки структурного подобия графов рекомендуется использовать относительное расстояние  $R_{ij}^V$ .

Два молекулярных графа не содержат общего подграфа, если их вершины не имеют общих меток. Например, пусть для выборки соединений необходимо найти такие общие подграфы молекулярных графов, которые содержат только вершины, входящие в циклы. Тогда вершинам циклов в графах приписывают метки, отличные от меток в ациклических подграфах. В этом случае некоторые молекулярные графы соединений выборки (например, бензол и гексан) не имеют общих подграфов и расстояние между ними максимально.

Значения функций расстояния  $r_{ij}^X$ ,  $R_{ij}^X$  могут быть максимальными даже при наличии в графах общих подграфов. Это возможно в случае, когда графы имеют общий подграф, состоящий из изолированных вершин, т.е. не имеют общих ребер. Например, молекулярные графы этилового спирта  $C_2H_5OH$  и двуокиси углерода  $CO_2$  имеют общий подграф, состоящий только из двух вершин С и О.

Из рассмотренных примеров и проведенного анализа метрических свойств функций расстояний между молекулярными графами следует целесообразность применения в компьютерных исследованиях выборок химических соединений относительной функции расстояния  $R_{ij}^{VX}$  или функции  $r_{ij}^{VX}$ .

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант № 05-01-00816.

## СПИСОК ЛИТЕРАТУРЫ

1. Стьюпер Э., Брюгге У., Джурс П. Машинный анализ связи химической структуры и биологической активности. – М.: Мир, 1982.
2. Макаров Л.И. // Журн. структур. химии. – 1998. – **39**, № 1. – С. 115 – 125.
3. Пиоттух-Пелецкий В.Н., Дерендяев Б.Г., Молодцов С.Г., Богданова Т.Ф. // Там же. – 1997. – **38**, № 4. – С. 785 – 794.
4. Varmuza K., Penchev P.N., Scsibrany H. // J. Chem. Inf. Comput. Sci. – 1998. – **38**. – P. 420 – 427.
5. Дерендяев Б.Г., Макаров Л.И., Богданова Т.Ф., Пиоттух-Пелецкий В.Н. // Журн. структур. химии. – 2001. – **42**, № 2. – С. 325 – 337.
6. Makarov L.I. // Commun. Math. Comput. Chem. – 2003. – **49**. – P. 171 – 178.
7. Johnson M., Naim M., Nicholson V., Tsai C.-C. In: Graph Theory and Topology in Chemistry / Eds. R.B. King and D.H. Rouvray. – Amsterdam: Elsevier, 1987. – P. 219 – 225.
8. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: Изд-во Ин-та математики, 1999.