

МЕТОД КОЛИЧЕСТВЕННОГО ОПРЕДЕЛЕНИЯ СХОДСТВА ГРАФОВ
НА ОСНОВЕ СТРУКТУРНЫХ СПЕКТРОВ

В.А. Кохов

В работе предложен подход к количественному определению сходства графов, позволяющий формировать и исследовать различные виды структурного сходства. Он основан на использовании информации о количестве канонических изоморфных вложений фрагментов из заданного набора в исследуемые графы. Рассмотрен метод стратифицированного построения инвариантов графов, используемых при определении сложности и сходства графов. Приведено функциональное наполнение пакета прикладных программ (ППП) SIGMA-GMS, разработанного для IBM PC AT и включающего базовый набор модулей (изоморфизм, канонизация, декомпозиция на неизоморфные фрагменты, построение матриц сходства и др.) для решения задач различения, количественного определения сложности и сходства графов, возникающих в экспертных системах с правдоподобными рассуждениями [1] и при создании интеллектуальных систем автоматизированного проектирования регулярных топологий вычислительных систем и сред [2].

1. Основные определения и обозначения

Пусть $G = (V(G), E(G))$ — обыкновенный граф с множеством вершин $V(G) = \{v_1, v_2, \dots, v_p\}$ и множеством ребер $E(G) =$

$= \{e_1, e_2, \dots, e_q\}$, где $E(G) \subseteq V(G) \times V(G)$. В графе G будем выделять три вида фрагментов:

- частичный граф $H = (V(H), E(H))$ графа G , если $V(H) = V(G)$ и $E(H) \subseteq E(G)$;

- подграф $H = (V(H), E(H))$ графа G , если $V(H) \subseteq V(G)$ и $E(H) = E(G) \cap (V(H) \times V(H))$;

- частичный подграф $H = (V(H), E(H))$ графа G , если $V(H) \subseteq V(G)$ и $E(H) \subseteq E(G) \cap (V(H) \times V(H))$.

Пусть $G = (V(G), E(G))$ и $H = (V(H), E(H))$ - графы, и β - взаимно-однозначное отображение множества $V(G)$ на множество $V(H)$ и γ - взаимно-однозначное отображение $E(G)$ на $E(H)$. Обозначим через φ упорядоченную пару $\langle \beta, \gamma \rangle$ и будем говорить, что φ есть изоморфизм графа G на H , если вершина v инцидентна ребру e в графе G тогда и только тогда, когда вершина βv инцидентна ребру γe в графе H . Если φ существует, то будем говорить, что графы G и H изоморфны и обозначать $G(R)H$. Отношение изоморфизма R между графами является отношением эквивалентности.

Пусть H - фиксированный граф и существует фрагмент f графа G , такой, что $V(f) = \beta V(H)$ и $E(f) = \gamma E(H)$. Будем писать $f \approx \varphi(H)$ и говорить, что φ отображает граф H на часть f графа G , т.е. $\varphi = \langle \beta, \gamma \rangle$ - изоморфизм H на часть f графа G . В соответствии с тремя видами фрагментов f графа G будем различать изоморфизм частичному графу, изоморфизм подграфу и изоморфизм частичному подграфу как три разновидности изоморфного вложения графа H в граф G .

Под изоморфным пересечением по фрагментам графов $A = (V(A), E(A))$ и $B = (V(B), E(B))$ будем понимать граф $MCS(A, B) = (V(MCS), E(MCS))$, который изоморфен некоторому фрагменту графа A и некоторому фрагменту графа B , причем хотя бы один из этих фрагментов максимален, т.е. имеет наибольшее число ребер, для ко-

того еще сохраняются указанные свойства. Через $MCS(A, B)$ обозначим граф, являющийся изоморфным пересечением по подграфам графов A и B , т.е. граф, который изоморфен некоторому подграфу графа A и некоторому подграфу графа B , причем хотя бы один из этих подграфов имеет наибольшее число вершин, для которого еще сохраняются указанные свойства.

Множество обыкновенных графов обозначим через M . Пусть Q - непустое множество элементов с определенной на нем эквивалентностью σ . Тогда функцию IN , заданную на M и принимающую значения в Q , будем называть инвариантом графа, если

$$\forall G_i, G_j \in M [G_i(R)G_j \Rightarrow IN(G_i) (\sigma) IN(G_j)].$$

Инвариант IN будем называть полным, если выполняется условие

$$\forall G_i, G_j \in M [IN(G_i) (\sigma) IN(G_j) \Rightarrow G_i(R)G_j].$$

Пусть $Aut(G)$ - группа графа - совокупность автоморфных подстановок на множестве V графа G , а $|Aut(G)|$ обозначает порядок группы. Понятия и обозначения теории графов, не определяемые далее в тексте, можно найти в [3].

2. Основные подходы к количественному определению сходства графов

Среди разнообразия подходов к определению сходства графов можно выделить три основных подхода. Первый подход связан с использованием различных топологических индексов графа, т.е. числовых инвариантов графа [4,5]. Наибольший интерес представляют топологические индексы, основанные на симметрии окрестностей вершин графа, так как они имеют обычно большую дискриминирующую способность по сравнению с другими топологическими индексами. Методы построения этих индексов используют подход, основанный на теории информации. Пусть X есть некоторое множе-

ство, состоящее из p элементов. Предположим, что по некоторому критерию эквивалентности элементы множества разбиваются на k классов эквивалентности X_i так, что

$$p = \sum_{i=1}^k p_i,$$

где p_i - число элементов подмножества X_i . Тогда величина p_i/p есть вероятность того, что выбранный случайным образом элемент множества X будет находиться в i -м подмножестве. Для количественной оценки информации, содержащейся в одном элементе множества, можно использовать энтропию распределения вероятностей элементов этого множества, которая определяется формулой Шеннона [5]:

$$IC = - \sum_{i=1}^k p_i \cdot \log_2 p_i.$$

Если вместо множества X рассматривать множество вершин графа, то IC представляет собой информационное содержание графа в расчете на одну вершину. Полное информационное содержание, т.е. мера сложности в расчете на один граф, определяется по формуле $TI(G) = p \cdot IC$. Полное информационное содержание графа рассматривается как интегральная количественная мера его структурной неоднородности или разнообразия. Различие методов построения топологических индексов, основанных на симметрии окрестностей, состоит в разнообразии характеристик окрестностей вершин через числовые инварианты графов окрестностей. Все эти методы не обеспечивают различение окрестностей вершин с точностью до изоморфизма графов окрестностей. В работе [6] предложен метод построения топологических индексов, основанных на симметрии окрестностей однотипных фрагментов графа, т.е. вместо множества вершин графа рассматривается множество всех неупорядоченных n -ок вершин графа, где $n = 1, 2, 3, \dots, (p-1)$. Отличительной осо-

бенностью метода является использование точного алгоритма распознавания изоморфизма окружения однотипных фрагментов графа.

Второй подход основан на использовании подграфовой метрики [7,8]. Пусть Γ обозначает пространство графов. Подграфовая метрика $d: \Gamma \times \Gamma \rightarrow N$, $N = \{0, 1, 2, \dots, K\}$, определяется по одной из следующих формул:

$$d1(A, B) = |V(A)| + |V(B)| - 2|V(MCS)|;$$

$$d2(A, B) = |V(A)| + |E(A)| + |V(B)| + |E(B)| - 2(|V(MCS)| + |E(MCS)|);$$

$$d3(A, B) = |E(A)| + |E(B)| - 2|E(MCS)| + ||V(A)| - |V(B)||.$$

Кроме того, используются индексы сходства (MSI) и несходства (MDI) графов:

$$MSI(A, B) = (NAB(MCS)/NAB(A)) \times (NAB(MCS)/NAB(B));$$

$$MDI(A, B) = 1.0 - MSI(A, B),$$

$$\text{где } NAB = |V| + |E|.$$

Отличительной особенностью данного подхода является то, что при его использовании в решении задач структурного поиска в базе графов, возможна организация трех случаев поиска структуры A , сходной с заданной структурой B : $B \subseteq A$; $d(A, B) < K1$; $d(A, MCS(A, B)) < K1$, где $K1$ - целое положительное число.

В работе [9] рассмотрены вопросы алгебраического определения сходства на множествах графов с помощью операций псевдоресечения и псевдообъединения, действующих на множествах графов. Показано, как можно определить эти операции, чтобы они удовлетворяли аксиомам дистрибутивной решетки. При этом учтена возможность установления сходства на пометках вершин графов.

Третий подход связан с использованием фрагментов графа и метрик на числовых последовательностях. Среди методов данного подхода следует выделить два, описанных в работах [10,11]. В первом из них графы представляются последовательностями простых цепей каждой возможной длины от каждой вершины. Суммарная

по вершинам последовательность цепей каждой длины в графе рассматривается как точка в K -мерном евклидовом пространстве. Структуры с близкими координатами будут скапливаться в одной и той же части многомерного пространства. Предполагается, что евклидово расстояние является мерой структурного сходства графов. Более мощный метод предложен в [11]. Он заключается в составлении списка всех связанных фрагментов молекулярного графа, по числу которых затем определяются различные коэффициенты сходства.

Обзор по подходам к упорядочиванию графов на основе различных теоретико-графовых характеристик, задающих разветвленность графов, приведен в [12], а задающих сложность молекулярных графов, в [13].

В данной работе предлагается наиболее общий подход к определению структурного сходства графов, позволяющий на единой основе строить систему количественного определения сложности и сходства графов. В основе предлагаемого подхода используется метод стратифицированного построения и исследования структурных характеристических инвариантов графов [3,6].

3. Структурные спектры и индексы сложности графа

ОПРЕДЕЛЕНИЕ 1. Набор фрагментов $F(G) = \langle f_1, f_2, \dots, f_j, \dots, f_k \rangle$ графа G , состоящий из всех неизоморфных собственных частей графа, называется структурным спектром графа G , а k - длиной спектра.

Пусть $A(G, f)$ обозначает количество всех изоморфных вложений фрагмента $f \in F(G)$ в граф G . Тогда, под числом канонических изоморфных вложений фрагмента f в граф G будем понимать величину $w(G, f)$, определяемую отношением $w(G, f) = A(G, f) / (|Aut(f)|)$.

ОПРЕДЕЛЕНИЕ 2. Структурный спектр графа G , в котором каждый фрагмент f_j повторяется число раз, равное $w_j(G, f_j)$, назы-

вается полным структурным спектром графа G и обозначается через

$$WF(G) = \langle (f_1^1, f_1^2, \dots, f_1^{w_1(G, f_1)}), \dots, (f_k^1, f_k^2, \dots, f_k^{w_k(G, f_k)}) \rangle.$$

С целью сокращения длины записи будем использовать обозначение $WF(G) = \langle w_1 f_1, w_2 f_2, \dots, w_j f_j, \dots, w_k f_k \rangle$, где w_j — число изоморфных канонических вложений фрагмента f_j в граф G .

Если фрагменты $f \in F(G)$ являются связными частями графа G , то структурный (полный структурный) спектр будем называть связным структурным (полным структурным) спектром графа G .

На рис.1 приведены примеры двух пар деревьев и их связные и полные связные структурные спектры.

Пусть $FF = \langle G_1, G_2, \dots, G_i, \dots, G_r \rangle$ обозначает произвольный набор неизоморфных графов.

ОПРЕДЕЛЕНИЕ 3. Структурным спектром $SS(G/FF)$ графа G относительно набора графов FF называется поднабор графов из FF , в который входят те и только те графы $G_i \in FF$, для которых выполняется условие $\exists j [f_j(R)G_i]$, где $f_j \in F(G)$, $j = 1, 2, \dots, k$, $i = 1, 2, \dots, r$. Выполнив перенумерацию графов $G_i \in SS(G/F)$ в естественном порядке, получим

$$SS(G/FF) = \langle G_1, G_2, \dots, G_i, \dots, G_{rr} \rangle,$$

где $rr \leq r$.

ОПРЕДЕЛЕНИЕ 4. Полным структурным спектром $FSS(G/FF)$ графа G относительно набора графов FF называется структурный спектр $SS(G/FF)$ относительно набора графов FF , в котором каждый граф $G_i \in SS(G/FF)$ повторяется число раз, равное $w_i(G, G_i)$, и обозначается через
















Деревья	Полный структурный спектр		
	Структурный спектр	$w_i(T_1, f_i)$	$w_i(T_2, f_i)$
 T_1	 	4	4
		3	3
 T_2		2	3
 T_1	   	6	6
		5	5
		5	5
		3	4
 T_2	   	1	2
		1	1
		2	1

Рис.1. Деревья с изоморфными связными структурными спектрами

$$FSS(G/FF) = \langle (G_1^1, G_1^2, \dots, G_1^{w_1(G, G_1)}), \dots, (G_{rr}^1, G_{rr}^2, \dots, G_{rr}^{w_{rr}(G, G_{rr})}) \rangle.$$

С целью сокращения длины записи будем использовать следующее обозначение

$$FSS(G/FF) = \langle w_1 G_1, w_2 G_2, \dots, w_i G_i, \dots, w_{rr} G_{rr} \rangle.$$

ОПРЕДЕЛЕНИЕ 5. Графы G_1, G_2, \dots, G_r называются эквивалентными (изоморфными) по структурному спектру относительно набора $F = \langle f_1, f_2, \dots, f_r \rangle$, если справедливо

$$\forall f_i \in F[f_i(R) f_1^i(R) f_2^i(R) \dots f_r^i(R)]$$

для любого $i = 1, 2, 3, \dots, r$, где $f_1^i \in F(G_1)$, $f_2^i \in F(G_2), \dots, f_r^i \in F(G_r)$.

ОПРЕДЕЛЕНИЕ 6. Графы G_1, G_2, \dots, G_r называются изоморфными (эквивалентными) по полному структурному спектру относительно набора графов F , если они эквивалентны по структурному спектру относительно набора f и выполняется условие

$$w_i = w_1^i = w_2^i = \dots = w_r^i$$

для любого $i = 1, 2, 3, \dots, r$.

Определим характеристику, с помощью которой будем сравнивать графы по их сложности. Пусть для некоторого графа G построен его полный связный структурный спектр $WF(G)$. Заметим, что $w_1 = p$, $f_1 = K_1$, $w_2 = q$, $f_2 = K_2$. Примем индекс IC фрагмента f_1 за 1, а фрагмента f_2 - за 3. Так как для любого фрагмента ($f_i \in F(G)$), $i \neq 1, 2$, можно определить его полный структурный спектр, а для каждого фрагмента от фрагмента графа G можно построить его полный структурный спектр и т.д., то рекурсивным образом всегда можно вычислить индекс спектральной сложности G .

ОПРЕДЕЛЕНИЕ 7. Индексом спектральной сложности графа G называется число $IC(G)$, определяемое следующим образом

$$IC(G) = w_1 IC(f_1) + w_2 IC(f_2) + \dots + w_k IC(f_k).$$

Индекс спектральной сложности дерева $T(q)$ меняется в следующих границах $IC(P(q)) \leq IC(T(q)) \leq IC(K_{1,q})$. Нетрудно определить вид функций $IC(P(q))$ и $IC(K_{1,k})$:

$$\begin{aligned} IC(P(q)) &= (q+1) + qIC(P(1)) + (q-1)IC(P(2)) + \dots \\ &\dots + (q-j)IC(P(j+1)) + \dots + 3[IC(P(q-2))] + 2[IC(P(q-1))]; \\ IC(K_{1,q}) &= (q+1) + c_q^1 IC(K_{1,1}) + c_q^2 IC(K_{1,2}) + \dots \\ &\dots + c_q^j IC(K_{1,j}) + \dots + c_q^{q-1} IC(K_{1,q-1}). \end{aligned}$$

Т а б л и ц а 1

Индексы спектральной сложности для деревьев с $p < 10$

№	IC	№	IC	№	IC	№	IC	№	IC
0	1	23	4098	32	27698	70	63312	86	197726
1	3	18	4126	47	28470	71	63884	85	198332
2	9	19	5068	44	34366	69	73768	79	223790
3	31	24	6881	43	37060	68	83274	60	226576
4	40	20	8531	33	37430	73	85172	80	226882
5	106	21	14421	34	48040	55	85404	84	259706
6	155	25	4220	46	59563	75	95582	93	260442
7	231	26	6996	45	64379	72	95938	64	296444
8	362	27	9068	35	83545	56	97340	61	298808
9	566	29	9848	36	145636	91	109404	88	350452
10	646	37	11674	48	14408	74	112188	82	351378
13	899	28	12550	49	24180	57	128032	83	397104
11	1033	38	15244	50	32336	77	129172	62	401734
12	1666	30	15506	51	37100	76	146884	63	530596
14	1236	41	16700	66	40740	90	147732	89	622361
15	2006	40	19908	52	48588	87	149692	81	705857
16	2476	39	21398	67	54772	59	150518	64	943361
17	3026	31	21636	53	55232	92	168570	65	1690867
22	3290	42	26210	54	56220	78	189612		

Длины связанных структурных спектров для деревьев с $r < 11$

№	NF	№	NF	№	NF	№	NF	№	NF	№	NF
1	1	35	11	69	16	102	16	135	26	168	24
2	2	36	7	70	17	103	19	136	25	169	25
3	3	37	10	71	18	104	25	137	18	170	25
4	3	38	13	72	18	105	25	138	18	171	27
5	4	39	12	73	19	106	21	139	23	172	17
6	5	40	13	74	17	107	24	140	25	173	25
7	4	41	14	75	15	108	21	141	25	174	25
8	5	42	12	76	20	109	21	142	26	175	18
9	7	43	15	77	19	110	27	143	23	176	15
10	7	44	15	78	22	111	24	144	27	177	22
11	7	45	12	79	19	112	26	145	22	178	26
12	5	46	10	80	13	113	18	146	27	179	27
13	6	47	12	81	15	114	21	147	29	180	25
14	6	48	8	82	20	115	25	148	27	181	25
15	9	49	13	83	20	116	20	149	32	182	24
16	10	50	16	84	16	117	19	150	30	183	25
17	8	51	16	85	22	118	19	151	28	184	28
18	10	52	16	86	18	119	15	152	31	185	22
19	10	53	15	87	21	120	9	153	30	186	18
20	9	54	16	88	19	121	14	154	21	187	24
21	6	55	20	89	14	122	19	155	20	188	31
22	8	56	18	90	19	123	20	156	25	189	29
23	10	57	19	91	18	124	19	157	27	190	24
24	9	58	13	92	15	125	21	158	28	191	32
25	7	59	17	93	19	126	22	159	30	192	30
26	11	60	20	94	18	127	22	160	24	193	26
27	13	61	16	95	9	128	23	161	31	194	27
28	12	62	16	96	15	129	21	162	26	195	24
29	12	63	16	97	19	130	22	163	20	196	21
30	13	64	13	98	20	131	24	164	25	197	25
31	15	65	8	99	20	132	25	165	25	198	16
32	12	66	12	100	18	133	22	166	19	199	22
33	13	67	16	101	21	134	22	167	16	200	26
34	13	68	16								

В табл.1 представлены индексы спектральной сложности для всех деревьев с $p < 10$. Нумерация деревьев соответствует расположению их диаграмм в [15]. В табл.2 для каждого дерева с $p < 11$ указано число фрагментов (NF), составляющих его связный структурный спектр. Из анализа на ЭВМ спектров деревьев следует справедливость следующих утверждений.

УТВЕРЖДЕНИЕ 1. Наименьшие по числу вершин деревья с изоморфными связными структурными спектрами имеют $p = 4$ (рис.1).

УТВЕРЖДЕНИЕ 2. Все деревья с $p < 10$ различаются своими связными структурными спектрами, кроме двух пар деревьев, приведенных на рис.1.

УТВЕРЖДЕНИЕ 3. Связный структурный спектр не является полным инвариантом в классе графов-деревьев.

Заметим, что указанные выше деревья различаются своими структурными спектрами и полными связными структурными спектрами. Если при исследовании деревьев рассматривать их полные связные цепные спектры, т.е. полные связные спектры, построенные относительно набора связных цепей P_1, P_2, \dots, P_{10} , то справедливы следующие утверждения.

УТВЕРЖДЕНИЕ 4. Все деревья с $p < 11$ различаются своими полными связными цепными спектрами, кроме четырех пар деревьев, приведенных на рис.2.

УТВЕРЖДЕНИЕ 5. Все деревья с $p < 11$ различаются своими полными спектрами, построенными относительно набора цепей.

4. Структурные спектральные индексы графа

Пусть $SS(G/F) = \langle f_o^1, f_1^1, \dots, (f_i^1, f_i^2, \dots, f_i^j, \dots, f_i^1), \dots, (f_q^1, f_q^2, \dots, f_q^d) \rangle$ обозначает структурный спектр графа G относительно набора связных графов F , в котором f_i^j обозначает j -й фрагмент графа G с числом ребер i .

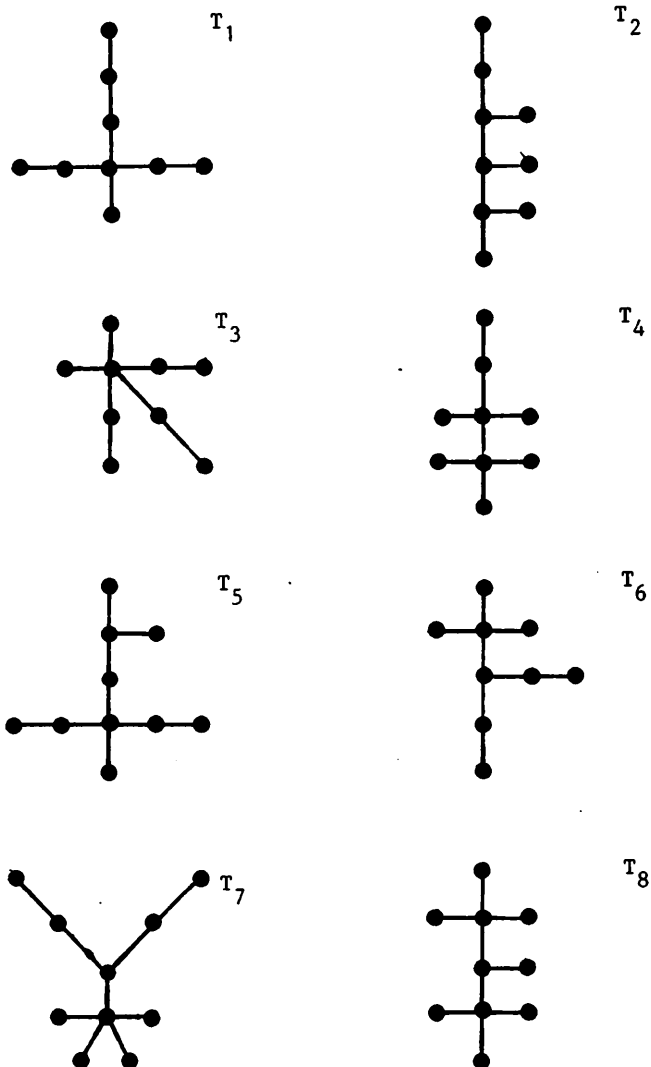


Рис.2. Пары деревьев (T_1, T_2) , (T_3, T_4) , (T_5, T_6) , (T_7, T_8) с изоморфными полными связными цепными спектрами.

ОПРЕДЕЛЕНИЕ 8. Структурным спектральным индексом $SSI(G/F)$ графа G относительно набора графов F называется запись следующего вида

$$SSI(G/F) = z_0 + z_1 + \dots + (z_i^1 + z_i^2 + \dots + z_i^j + \dots + z_i^1) + \dots \\ \dots + (z_q^1 + z_q^2 + \dots + z_q^d),$$

где z_i^j - фиктивная переменная, задающая информацию о наличии в $SS(G/F)$ графа типа j с числом ребер i без учета структуры этого графа, 1 - число типов графов, т.е. число неизоморфных графов, имеющих i ребер.

В качестве примера приведем значения $SS(G/F)$ для четырех деревьев (рис.1):

$$(p=4) \quad SSI(T_1/F) = SSI(T_2/F) = z_0 + z_1 + z_2;$$

$$(p=6) \quad SSI(T_1/F) = SSI(T_2/F) = z_0 + z_1 + z_2 + (z_3^1 + z_3^2) + (z_4^1 + z_4^2).$$

ОПРЕДЕЛЕНИЕ 9. Полным структурным спектральным индексом $FSI(G/F)$ графа G относительно набора графов F называется $SSI(G/F)$ графа G , определенный относительно полного структурного спектра графа G , т.е.

$$FSI(D/F) = w_0^1 z_0 + w_1^1 z_1 + \dots + \langle w_i^1 z_i + w_i^2 z_i + \dots \\ \dots + w_i^j z_i^j + \dots + w_i^1 z_i^1 \rangle + \dots + \langle w_q^1 z_q^1 + w_q^2 z_q^2 + \dots + w_q^d z_q^d \rangle,$$

где w_i^j обозначает число канонических изоморфных вложений графа F_i^j в граф G . Угловые скобки задают закон упорядочения, например по убыванию значений w_i^j . Заметим, что $w_0^1 = p$, $w_1^1 = q$.

В качестве примера приведем значения $FSI(G/F)$ для четырех деревьев (рис.1), вычисленные относительно связанных частей деревьев:

$$FSI(T_1/F) = 4z_0 + 3z_1 + 2z_2,$$

$$FSI(T_2/F) = 4z_0 + 3z_1 + 3z_2;$$

Т а б л и ц а 3

Полные структурные спектральные индексы деревьев с $p < 8$
относительно их связанных фрагментов

№1	№2	№3	Структурный спектральный индекс
1	1	1	$2z_0 + z$
2	2	2	$3z_0 + 2z_1 + z_2$
3	3	3	$4z_0 + 3z_1 + 2z_2$
4	4	4	$4z_0 + 3z_1 + 3z_2$
5	5	5	$5z_0 + 4z_1 + 3z_2 + 2z_3$
6	6	6	$5z_0 + 4z_1 + (2z + z)_3$
7	7	7	$5z_0 + 4z_1 + 6z_2 + 4z_3$
8	8	8	$6z_0 + 5z_1 + 4z_2 + 3z_3 + 2z_4$
9	9	9	$6z_0 + 5z_1 + 5z_2 + (3z + z)_3 + (2z + z)_4$
10	10	10	$6z_0 + 5z_1 + 5z_2 + (4z + z)_3 + (2z + z)_4$
11	12	12	$6z_0 + 5z_1 + 7z_2 + (4z + 3z)_3 + (3z + z)_4$
12	13	13	$6z_0 + 5z_1 + 10z_2 + 10z_3 + 5z_4$
13	11	11	$6z_0 + 5z_1 + 6z_2 + (4z + 2z)_3 + 4z_4$
14	14	14	$7z_0 + 6z_1 + 5z_2 + 4z_3 + 3z_4 + 2z_5$
15	15	15	$7z_0 + 6z_1 + 6z_2 + (4z + z)_3 + (3z + z)_4 + (2z + z)_5$
17	17	17	$7z_0 + 6z_1 + 6z_2 + (5z + z)_3 + (3z + 2z)_4 + (z + z + z)_5$
18	20	20	$7z_0 + 6z_1 + 6z_2 + (6z + z)_3 + (3z + 3z)_4 + 3z_5$
19	21	21	$7z_0 + 6z_1 + 8z_2 + (6z + 4z)_3 + (6z + 2z + z)_4 + (2z + 2z)_5$
20	23	23	$7z_0 + 6z_1 + 10z_2 + (10z + 4z)_3 + (6z + 5z)_4 + (4z + z)_5$
21	24	24	$7z_0 + 6z_1 + 15z_2 + 20z_3 + 15z_4 + 6z_5$
22	18	18	$7z_0 + 6z_1 + 7z_2 + (4z + 2z)_3 + (4z + 2z)_4 + 4z_5$
23	19	19	$7z_0 + 6z_1 + 7z_2 + (6z + 2z)_3 + (5z + 2z)_4 + (2z + z + z)_5$
24	22	22	$7z_0 + 6z_1 + 9z_2 + (6z + 5z)_3 + (9z + z)_4 + (3z + 2z)_5$

$$FSI(T_1/F) = 6z_0 + 5z_1 + 5z_2 + (3z_3^1 + z_3^2) + (2z_4^1 + z_4^2);$$

$$FSI(T_2/F) = 6z_0 + 5z_1 + 5z_2 + (4z_3^1 + z_3^2) + (2z_4^1 + z_4^2).$$

В табл. 3 приведены FSI для деревьев с $p < 8$. В качестве F выбраны связанные фрагменты дерева. С целью сокращения записи вместо обозначения

$$\langle z_i^1 + z_i^2 + \dots + z_i^j + \dots + z_i^1 \rangle$$

использовано обозначение

$$(z + z + \dots + z + \dots + z)_i.$$

Нумерация деревьев №1 соответствует расположению их диаграмм в [15]. Нумерации №2 и №3 задают упорядоченность деревьев, задаваемую ростом значений их структурных спектральных индексов относительно связанных фрагментов ("разветвленность" дерева) и их связанных цепных фрагментов ("линейность" дерева), соответственно.

Полный структурный спектральный индекс графа является инвариантом, характеризующим "сложность" графа через информацию об общем количестве неизоморфных фрагментов графа, количестве типов фрагментов графа, включающих одинаковое число ребер, и количестве возможных способов "вырезки" каждого фрагмента из графа. Выбирая различные стратифицированные наборы F, можно построить широкий спектр характеристик "сложности" графа, с целью проведения сравнительного сходства графов и их упорядочения в соответствии с различными интерпретациями понятия "сложность" графа. Сравнительный анализ SSI для деревьев приводит к справедливости следующих утверждений.

УТВЕРЖДЕНИЕ 6. *Наименьшие по числу вершин деревья с одинаковыми FSI относительно связанных цепей содержат 9 вершин. Существует две пары таких деревьев.*

Деревья, приведенные на рис.2, имеют следующие значения FSI относительно связанных цепей:

$$FSI(T1/P) = FSI(T2/P) = 8z_1 + 10z_2 + 10z_3 + 6z_4 + 2z_5;$$

$$FSI(T3/P) = FSI(T4/P) = 8z_1 + 13z_2 + 12z_3 + 3z_4;$$

$$FSI(T5/P) = FSI(T6/P) = 9z_1 + 12z_2 + 11z_3 + 9z_4 + 4z_5;$$

$$FSI(T7/P) = FSI(T8/P) = 9z_1 + 15z_2 + 12z_3 + 9z_4.$$

УТВЕРЖДЕНИЕ 7. Среди деревьев с $p < 11$ существует четыре пары деревьев с одинаковыми значениями FSI относительно связанных цепей (рис.2).

5. Задачи структурно-топологического различения графов

Пусть ξ - отношение эквивалентности на графах и задан некоторый набор фрагментов $F^* = \langle f_1, f_2, \dots, f_i, \dots, f_r \rangle$, среди которых могут быть изоморфные и несвязные. Обозначим через $P(G)$ множество всех фрагментов графа G . Под задачей структурно-топологического различения (ниже просто различения) графов

$$M1 = \{G_1, G_2, \dots, G_j, \dots, G_k\},$$

где $M1 \subseteq M$, будем понимать задачу определения ξ -эквивалентности графов по набору фрагментов F^* , т.е. задачу в которой анализируется выполнение следующего условия

$$\forall f_i \in F^* [f_i(\xi) f_1^i(\xi) f_2^i(\xi) \dots (\xi) f_k^i], \text{ для любого } i = 1, 2, 3, \dots, r,$$

где $f_1^i \in P(G_1), f_2^i \in P(G_2), \dots, f_r^i \in P(G_r)$. Таким образом, задача различения графов определяется тройкой $\langle M1, F^*, \xi \rangle$.

Классификацию задач различения будем осуществлять на основе следующих признаков: вид используемого отношения ξ -эквивалентности; вид используемого набора F^* ; вид части каждого графа G_j , $j = 1, 2, \dots, l$, которая сравнивается на ξ -эквивалентность с графами набора F^* .

Примерами отношения ξ -эквивалентности графов являются: изоморфизм; 2-изоморфизм; структурный изометризм [16]; структурный изотопизм [16]; гомеоморфизм. В качестве примеров набора графа F^* могут выступать: цепи длины от 1 до p ; звезды $K_{1,n}$, где $n = 1, 2, \dots, q$; все деревья с числом вершин от 1 до p ; циклы длины от 3 до p ; клики мощности от 1 до p ; $(p-1)$ -вершинные подграфы исследуемых графов; $(q-1)$ -реберные фрагменты графов.

Выделим класс задач различения, для которого $\xi = R$. Среди задач данного класса рассмотрим задачу $\langle M1, F^*, R \rangle$, где $M1 = \{G_1, G_2\}$, $M1 \subseteq M$. Тогда при $F^* = P(G_1)$ и условиях $|V(G_1)| = |V(G_2)|$ и $|E(G_1)| = |E(G_2)|$ имеем задачу распознавания изоморфизма графов G_1 и G_2 , а при условиях $|V(G_1)| \leq |V(G_2)|$ и $|E(G_1)| < |E(G_2)|$ имеем задачу распознавания изоморфного вложения G_1 в G_2 , как частные случаи задач различения. Если $F^* \subset P(G_1)$ и $|V(G_1)| < |V(G_2)|$, то имеем задачу определения изоморфного пересечения графов G_1 и G_2 по набору фрагментов F^* .

В работе [3] предложен обобщенный метод решения задач из данного класса, на основе которого построены эффективные алгоритмы. Отличительной особенностью метода, названного методом монотонного расширения частичных решений, является то, что он построен на единой основе для всех видов задач изоморфного вложения и пересечения графов, т.е. распознает изоморфизм подграфа, изоморфизм частичному подграфу, изоморфизм частичному графу и предусматривает возможность модификации путем замены инвариантов и используемого предиката принадлежности с целью адаптации к различным семействам графов и отношений ξ -эквивалентности. Основу метода составляет направленный перебор с применением инвариантов и симметрии для сокращения перебора и использованием свойств монотонности отношений ξ -эквивалентности по включению частей. Отношение ξ -эквивалентности называется монотонным, если из ξ -эквивалентности графа G в граф H следует ξ -эквивалентность вложения любой части графа G в H .

Основным инструментом при решении задач различения является понятие "инвариант" графа. Рассмотрим один из подходов к построению инвариантов графов на основе инвариантов вершин графов.

6. Структурные характеристические инварианты графа, метод их построения и исследования

Орбиту вершины $v \in V$ группы $\text{Aut}(G)$ обозначим через

$$\Theta(\text{Aut}(G), v) = \{v^* : \exists g \in \text{Aut}(G)[g(v^*) = v]\}.$$

Орбиты задают разбиение множества V на классы ξ -эквивалентности; обозначим его через $V^*(G) = \{\theta_1, \theta_2, \dots, \theta_p\}, 1 \leq p$. Вершины v_i и v_j , принадлежащие одной орбите, будем называть ξ -эквивалентными и записывать $v_i(\xi)v_j$.

ОПРЕДЕЛЕНИЕ 10. Функция In , заданная на множестве вершин V графа $G = (V(G), E(G))$ и принимающая значения в Q , называется инвариантом вершины $v \in V$, если $\forall v_i, v_j \in V[v_i(\xi)v_j \Rightarrow \text{In}(v_i)(\sigma)\text{In}(v_j)]$.

Пусть над множеством V с отношением ξ заданы инварианты вершины $v \in V$: $\text{In}_1, \text{In}_2, \dots, \text{In}_K$, принимающие соответственно значения в множествах $Q_1(\sigma_1), Q_2(\sigma_2), \dots, Q_K(\sigma_K)$, тогда функцию $\text{In}(v) = \langle \text{In}_1(v), \text{In}_2(v), \dots, \text{In}_K(v) \rangle$, где в угловых скобках записана упорядоченная последовательность соответствующих инвариантов, будем называть декартовым произведением инвариантов вершины.

Пусть задана функция ϵ над $Q(\sigma)$, принимающая значения в $H(p)$. Тогда функция $h(v) = \epsilon[\text{In}(v)]$ является инвариантом вершины и называется суперпозицией инвариантов In и ϵ . Вершины v_i и v_j графа G называются In -эквивалентными, т.е. $v_i(\text{In})v_j$, если справедливо соотношение $\text{In}(v_i)(\sigma)\text{In}(v_j)$.

Перейдем к рассмотрению вопросов сравнимости In -инвариантов по критерию их "чувствительности", т.е. способности разли-

чать автоморфные вершины графа. Обозначим через $|V(\sigma)|$ мощность множества классов ξ -эквивалентности вершин графа $G \in M$, а через $|V(In)|$ обозначим мощность множества классов In -эквивалентности [17].

ОПРЕДЕЛЕНИЕ 11. Чувствительностью In -инварианта, заданного на V , называется отношение вида

$$\alpha(In, \xi) = \alpha(In) = |V(In)| / |V(\xi)|.$$

Очевидно, что $0 < \alpha \leq 1$. Если $\alpha(In) = 1$, то In -инвариант называется точным или полным для графа G . При этом из условия $In(v_i)(\sigma)In(v_j)$ следует $v_i(\xi)v_j$ и точный инвариант решает задачу различения на множестве V графа G , если эта задача решена в Q с эквивалентностью σ .

ОПРЕДЕЛЕНИЕ 12. In -инвариант называется точным в классе $M^* \subset M$ если справедливо $\forall G \in M^* [\alpha(In) = 1]$.

Будем называть In_1 -инвариант более сильным в классе M^* , чем In_2 -инвариант, если справедливо

$$\forall G \in M^* \forall v_i, v_j \in V [In_1(v_i)(\sigma_1)In_1(v_j) \Rightarrow In_2(v_i)(\sigma_2)(v_j)].$$

Очевидно, что In_1 -инвариант и In_2 -инвариант будут равносильными в классе M^* , если справедливо

$$\begin{aligned} \forall G \in M^* \forall v_i, v_j \in V [In_1(v_i)(\sigma_1)In_1(v_j) \Rightarrow In_2(v_i)(\sigma_2)In_2(v_j) \& \\ \& In_2(v_i)(\sigma_2)In_2(v_j) \Rightarrow In_1(v_i)(\sigma_1)In_1(v_j)]. \end{aligned}$$

С вычислительной точки зрения для равносильных In -инвариантов становится актуальной задача сравнения их по вычислительной и емкостной сложности.

Одним из наиболее общих методов построения In -инвариантов может выступать метод ярусного разбиения множества вершин графа на подмножества относительно вершин $v \in V$ [16, 3]:

$$VR(G, v) = \{V_j(v), j=0, 1, 2, \dots, e(v), v^* \in V_j \Leftrightarrow d(v, v^*) = j\},$$

где $e(v)$ — эксцентриситет вершины v ; $d(v, v^*)$ — расстояние меж-

ду вершинами v, v^* . Подмножества $V_j(v)$ называют ярусами высоты j в ярусном разбиении графа относительно v . Под окружением вершины v высоты j в ярусном разбиении графа будем понимать подграф G^j , порожденный подмножеством вершин $V^j = v \cup V_1 \cup V_2 \cup \dots \cup V_j$. Обозначим его через $G^j \langle v^j \rangle$. Таким образом, выделяя последовательно подграфы $G^j \langle v^j \rangle$, где $j = 1, 2, 3, \dots, e(v)$, и для каждого из них, определив структурный спектр $F(G^j, v)$, в который входят фрагменты, включающие вершину v , приходим к одному из методов стратифицированного построения In -инвариантов с целью исследования вопросов характеристики ξ -эквивалентности вершин графа из заданного класса. Заметим, что спектр $F(G^j, v)$ тоже может быть стратифицирован (например, по возрастанию значений $\text{IC}(f)$).

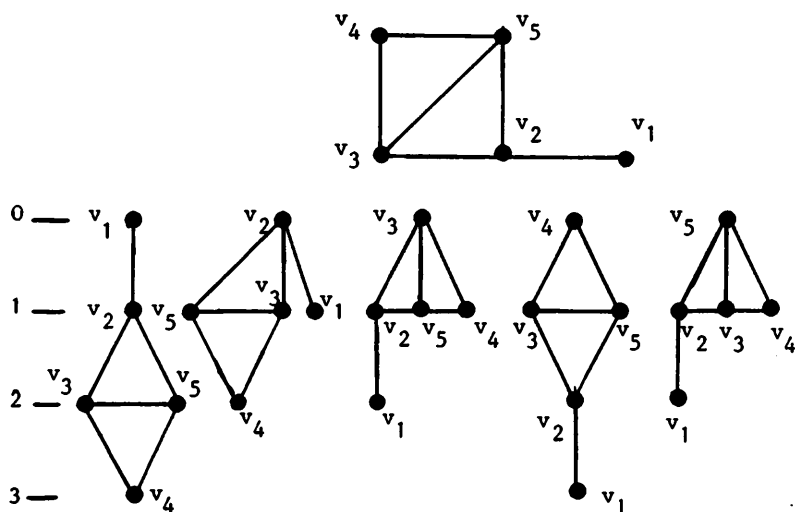


Рис.3

На рис.3 приведен пример графа и его ярусные разложения относительно каждой из вершин.

Пусть $F(G) = \langle f_1, f_2, \dots, f_j, \dots, f_r \rangle$ обозначает набор связанных фрагментов графа G . Рассмотрим матрицу $MF_v(G/F) = ||mf_{ji}||$, где $i = 1, 2, 3, \dots, p$; $j = 1, 2, \dots, r$; mf_{ji} равно числу фрагментов $f_j \in F$, включающих вершину v_i . Упорядочим столбцы матрицы лексикографически по признаку возрастания значений. В результате получим канонический вид матрицы $MF_v(G/F)$, который обозначим через $MF_v^k(G/F)$, если ее рассматривать в совокупности с каноническим видом набора фрагментов $F(G)$ (например, фрагменты представлены своими каноническими матрицами смежности). Каждый столбец матрицы $MF_v^k(G/F)$ представляет собой значение Ip -инварианта соответствующей вершины графа G , а вся матрица - инвариант графа G . Матрицу $MF_v^k(G/F)$ будем называть матрицей канонических изоморфных вложений фрагментов из $F(G)$ в граф G . Она является структурным характеристическим инвариантом графа [6].

Т а б л и ц а 4

Матрица $MF_v^k(G/F)$

F	1	4	2	3	5	1	4	2	3	5	1	4	2	3	5
f_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
f_2	1	2	3	3	3	1	2	3	3	3	1	2	3	3	3
f_3	0	2	4	6	6	2	4	7	8	8	2	5	7	8	10
f_4	0	0	2	2	2	2	4	10	8	8	8	6	10	8	10
f_5	0	0	0	0	0	0	0	4	4	4	4	4	4	4	4
f_6	0	0	1	2	2	1	2	3	3	3	1	2	3	3	3
f_7	0	0	0	0	0	0	0	4	4	4	4	4	4	4	4
f_8	0	1	1	2	2	0	1	11	2	2	0	1	1	2	2
f_9	0	0	1	4	4	1	4	5	5	5	1	4	5	5	5
f_{10}	0	0	0	0	0	0	0	2	2	2	2	2	2	2	2
f_{11}	0	0	0	0	0	0	0	2	2	2	2	2	2	2	2
f_{12}	0	0	0	1	1	0	1	1	1	1	0	1	1	1	1
f_{13}	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
f_{14}	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1

В табл. 4 приведены три матрицы $MF_v^k(G/F)$ для графа G (рис.4), построенные соответственно относительно наборов $F(G^1, v)$, $F(G^2, v)$, $F(G^3, v)$ для каждой вершины графа G .

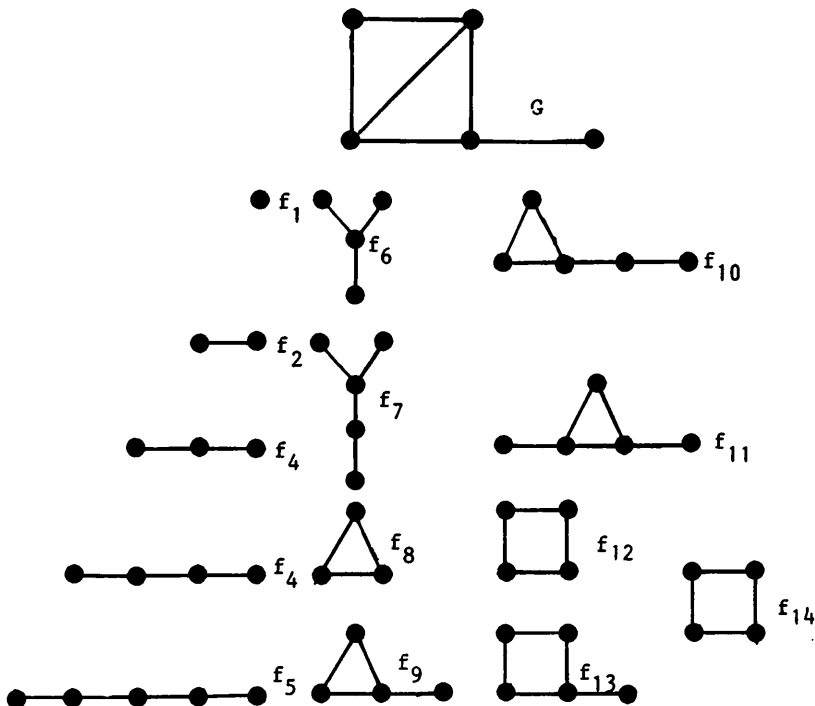


Рис.4

В работе [3] рассмотрен метод построения Ip -инвариантов однотипных фрагментов графа, т.е. вместо ярусного разбиения графа G на подмножества относительно вершин, используется метод ярусного разбиения графа G на подмножества относительно неупорядоченных p -ок вершин графа.

Нетрудно убедиться, что справедливо утверждение, связывающее $FSS(G/F)$ со структурным характеристическим инвариантом графа G относительно набора F .

УТВЕРЖДЕНИЕ 8. Имеем

$$\begin{aligned} \text{FSS}(G/F) &= \langle w_1 f_1, w_2 f_2, \dots, w_k f_k, \dots, w_r f_r \rangle = \\ &= \langle ((\sum_{i=1}^p m f_{i1})/|f_1|) f_1, ((\sum_{i=1}^p m f_{i2})/|f_2|) f_2, \dots \\ &\dots, ((\sum_{i=1}^p m f_{ir})/|f_r|) f_r \rangle, \end{aligned}$$

где $|f_i| = |V(f_i)|$, т.е. $w_k = \sum_{i=1}^p m f_{ik} / |f_k|$.

Заметим, что деревья (рис.2) различаются своими каноническими матрицами $MF_v(G/P)$, где P обозначает набор связанных цепей.

Таким образом, матрица $M_v^k(G/F)$ является более чувствительным инвариантом графа G , чем $\text{FSS}(G/F)$ и позволяет решать как задачу различения вершин графа, так и задачу различения графов.

7. Метод количественного определения сходства вершин, графов и семейств графов

Будем говорить, что сформулирована задача количественного определения сходства графов, если заданы следующие параметры:

- 1) G_1, G_2, \dots, G_k - графы, исследуемые на сходство;
- 2) $\langle f_1, f_2, \dots, f_r \rangle$ - структурный спектр, относительно которого определяется сходство графов;
- 3) $\pi_1, \pi_2, \pi_3, \dots, \pi_1$ - совокупность правил для проведения сравнения графов;
- 4) $d(G_i, G_k)$ - метрика или псевдометрика для графов.

Необходимо построить матрицу коэффициентов попарного сходства графов и провести упорядочение графов относительно графа, выбранного в качестве стандарта.

Предлагаемый метод предназначен для количественного определения сходства между вершинами (однотипными фрагментами) гра-

фа, между графами и между семействами графов. Он базируется на построении характеристических инвариантов вершин графа с использованием различных (желательно стратифицированных) структурных спектров графов. Коэффициенты сходства вычисляются на основе определения чисел канонических изоморфных вложений фрагментов из заданного спектра в исследуемые графы.

1. Сходство вершин графов. Для определения сходства вершин данный метод включает 5 следующих этапов:

1. Построение структурного спектра $SS(G/F)$, отвечающего заданным требованиям исследования вершин на сходство (например, спектр цепей, спектр звезд, спектр ярусных форм, спектр каркасов, спектр деревьев, спектр гомеоморфно-неприводимых деревьев, спектр циклов, спектр клик, спектр подграфов окружений).

2. Представление спектра в каноническом виде $SS^*(G/F)$.

3. Построение матрицы $MF_V(G)$ канонических изоморфных вложений фрагментов из $SS(G/F)$ в исследуемый граф. Столбец j матрицы $MF_V(G)$ представляет собой набор чисел m_{ij} , где m_{ij} - число способов, которыми фрагмент f_i может быть "вырезан" из графа G , если данный фрагмент включает вершину v_j , $i = 1, 2, \dots, r$, $j = 1, 2, 3, \dots, p$.

4. Представление матрицы $MF_V(G)$ в каноническом виде $MF_V^k(G)$.

5. Выбор метрики (Евклида, Хемминга и др.) или псевдометрики и определение коэффициентов сходства вершин на основе вычисления попарных расстояний, определяемых с помощью метрики. Коэффициенты сходства каждой пары вершин можно определить с помощью построения матриц сходства $SM1$ и $SM2$ [11]. Расчет значений элементов матрицы $SM1$ выполняется построчным сравнением каждых двух столбцов матрицы $MF_V^k(G)$ и делением меньшего из двух чисел в одной строке на большее (для $SM2$ делением суммы

Матрица канонических изоморфных еложений фрагментов

Деревья с $p = 6$	Матрица MF(G)						Фрагменты
	1	2	3	4	5	6	
1 	6	6	6	6	6	6	
	5	5	5	5	5	5	
2 	4	5	5	5	10	6	
	3	3	4	3	0	4	
	0	1	1	4	10	2	
3 	2	2	1	0	0	0	
	0	0	0	1	5	0	
4 	0	1	2	3	0	4	
	6	5	5	3	0	4	
5 	6	6	4	3	0	4	
	1	0	1	0	0	0	
6 	2	0	2	0	0	0	
	0	1	0	1	0	0	
	1	1	0	0	0	1	

меньших из двух чисел на сумму больших по всем строкам) с последующим вычислением среднего значения полученных частных.

Требования на построение $SS(G/F)$ формируются с помощью предиката, разрешающего или запрещающего фрагмент графа. Если ограничений нет, то для графа строится его связный структурный спектр.

2. Сходство графов. Отличительной особенностью сравнения графов от сравнения вершин является лишь то, что попарно сравнивать теперь необходимо не вектора, а матрицы $MF_v^k(G)$ исследуемых графов.

С целью упрощения сравнения (с уменьшением в общем случае чувствительности применяемых инвариантов) легко превратить матрицу $MF_v^k(G)$ в вектор (см. утверждение 8) суммированием элементов строки и делением на число вершин фрагмента. Таким образом, вступает в силу выбор правил на проведение сравнения.

В табл. 5 в качестве примера приведена матрица $MF(G)$ для деревьев с $p = 6$, каждый столбец которой связан теперь с характеристикой графа-дерева вместо вершины. В качестве фрагментов, составляющих $SS(G/F)$, разрешены все части дерева, кроме частей с изолированными вершинами. В табл. 6 представлены мат-

Т а б л и ц а 6

Матрицы сходства $SM1$ и $SM2$

	1	2	3	4	5	6	
1	1.000	0.587	0.629	0.327	0.200	0.479	
2	0.800	1.000	0.571	0.533	0.217	0.618	
$SM2 = 3$	0.756	0.756	1.000	0.441	0.217	0.553	$= SM1$
4	0.500	0.636	0.636	1.000	0.330	0.578	
5	0.263	0.309	0.309	0.469	1.000	0.280	
6	0.600	0.714	0.756	0.756	0.358	1.000	

Т а б л и ц а 7

Матрицы сходства SM1 и SM2

	1	2	3	4	5	6	
1	1.000	0.737	0.728	0.397	0.126	0.528	
2	0.839	1.000	0.546	0.545	0.214	0.721	
SM2 = 3	0.781	0.742	1.000	0.477	0.1888	0.515	= SM1
4	0.545	0.655	0.655	1.000	0.340	0.685	
5	0.250	0.312	0.312	0.500	0.100	0.267	
6	0.636	0.759	0.759	0.800	0.393	1.000	

рицы сходства SM1 и SM2. Если ограничиться только рассмотрением связанных цепных фрагментов в структурном спектре $SS(G/F)$, то матрицы сходства будут иметь вид, представленный в табл.7.

3. Сходство семейств графов. Если просуммировать в табл.5 элементы одной и той же строки матрицы $MF(G)$ и результат разделить на количество графов, то мы получим некоторую усредненную характеристику для семейства из 6 графов. Таким образом, мы можем построить новую матрицу для проведения сравнения на сходство различных семейств графов по заданному набору фрагментов структурного спектра исследуемых графов.

8. Функциональное наполнение ППП SIGMA-GMS

ППП представляет собой инструментальное программное средство для исследования и накопления данных о структурном сходстве граф-моделей систем и процессов. Под граф-моделью понимаем формальную систему следующего вида: $GMS = \langle V, E, f, WV, WE \rangle$, где V - множество вершин; E - множество ребер; WV - вес вершин; $f: E \Rightarrow B(V)$, где $B(V)$ - булеан множества V ; WE - вес ребер. Частными случаями граф-моделей являются гиперграфы, мультиграфы, графы, деревья.

Основная проблемная ориентация ППП - экспертные системы по химической информатике и экспертные системы по регулярным топологиям вычислительных сетей и сред.

ППП включает 8 подсистем, открытых к расширению.

ПС-1 - подсистема "Ввод, вывод и визуализация" предназначена для ввода-вывода граф-моделей, преобразования их стандартных представлений, прорисовкой граф-моделей по безкоординатному представлению с учетом вложения их симметрии в симметрию плоскости и возможностью использования библиотеки шаблонов.

ПС-2 - подсистема "Идентификация и изоморфизм" предназначена для построения канонических матриц смежности и инцидентности, распознавания изоморфизма и выделения неизоморфных граф-моделей в заданном наборе.

ПС-3 - подсистема "Изоморфное вложение и характеристичный анализ" предназначена для определения вложения и поиска заданных фрагментов, являющихся запрещенными или разрешенными фрагментами граф-модели.

ПС-4 - подсистема "Симметрия" предназначена для определения орбит, орбит стабилизатора и орбит фиксатора группы автоморфизмов граф-моделей. Строится порождающее множество и все автоморфизмы группы автоморфизмов граф-моделей [3].

ПС-5 - подсистема "Декомпозиция на фрагменты" предназначена для декомпозиции граф-моделей на неизоморфные фрагменты, отвечающие заданным ограничениям.

ПС-6 - подсистема "Сходство на основе спектров и спектральных индексов" предназначена для определения сходства на основе определения канонических изоморфных (гомеоморфных) вложений фрагментов и использования метрик и построения матриц сходства $SM1$ и $SM2$.

ПС-7 - подсистема "Сходство на основе максимальных изоморфных пересечений" предназначена для определения максималь-

ных изоморфных пересечений пары или семейства граф-моделей и определения расстояний d_1, d_2, d_3 между граф-моделями.

ПС-8 - подсистема "Сходство на основе индексов сложности" предназначена для вычисления цепного индекса сложности граф-моделей:

$$CL1(G) = ((p*q)/(p+q)) \sum_{i < j} (L(i,j)),$$

где $L(i,j)$ - число различных простых цепей из v_i в v_j . На основе обобщенной модели сложности граф-моделей с использованием единого теоретико-информационного подхода вычисляются индексы информационного содержания граф-моделей, связанные с симметрией (Ip-эквивалентностью) окружения однотипных фрагментов. Вычисляются индексы спектральной сложности граф-модели.

ППП разработан для IBM PC AT, языки программирования C++ и MS FORTRAN версии 5.0 и выше.

З а к л ю ч е н и е

Отличительной особенностью предложенного выше метода количественного определения сходства является то, что он позволяет стратифицированно формировать различные понятия сходства (структурной эквивалентности или толерантности) в зависимости от выбора Ip-инвариантов. Определение количественной меры сходства между графами (вершинами или однотипными фрагментами графа, семействами графов) позволяет решать задачу упорядочивания (частичного упорядочивания) графов относительно некоторого выбранного стандартного графа, в соответствии с выделенным набором свойств для сравнения.

В заключении работы выделим несколько сфер, стимулирующих разработку теории сходства графов и эффективных инструментальных программных средств для количественного определения сходства графов. Так, например, для создания интеллектуальных экспертных систем с правдоподобными рассуждениями необходима раз-

работка развитых теорий и инструментальных программных средств для определения сходства структурированных нечисловых объектов (графов, мультиграфов и др.) [1]. Сдерживающим фактором при создании интеллектуальных систем автоматизированного проектирования регулярных топологий ВС и сред является неразработанность методов определения сходства граф-моделей регулярных топологий [2]. Разработка экспертных систем в области химической информатики, связанных с изучением механизмов реакций и количественных корреляций вида структура-активность, основана на допущении наличия у однотипных молекул сходных свойств их молекулярных графов [5].

В теории и практике ситуационного управления требуются эффективные алгоритмы выполнения базовых операций, среди которых наиболее сложной и трудоемкой является операция пересечения (сходства) семантических сетей [18].

Наряду с динамическими задачами, описываемыми дифференциальными или разностными уравнениями, часто возникают задачи, где во времени меняется сама структура системы. Примерами задач такого рода являются: динамические процессы "жизни" административной структуры; перестройки, движущиеся "снизу-вверх", в организации систем связи, снабжения и т.п.; динамические процессы изменения номенклатуры товаров или изделий определенной категорий; динамические процессы перестройки архитектур ВС и сред. Совокупность методов описания и изучения таких систем в [19] названа графодинамикой. При проведении исследования в графодинамике вводится в рассмотрение понятие "расстояние" между графами. Это расстояние может отражать интуитивное представление о близости или сходстве графов. Определение понятия сходства графов помогает ввести представление об устойчивости графовых траекторий по отношению к малым возмущениям и о монотонности в смысле этого сходства процессов в графодинамике.

Л и т е р а т у р а

1. ГЕРГЕЙ Т., ФИНН В.К. Об интеллектуальных системах //Экспертные системы: состояние и перспективы.-М.: Наука,1991. - С. 9-20 (ИППИ РАН).
2. АРТАМОНОВ Г.Т., ТЮРИН В.Д. Топология сетей ЭВМ и мно - гопроцессорных систем. - М.: Радио и связь, 1991. - 248 с.
3. НЕЧЕПУРЕНКО М.И., ПОПКОВ В.К., КОХОВ В.А. и др. Алго - ритмы и программы решения задач на графах и сетях. - Новоси - бирск: Наука, Сиб. отд-ние, 1990. - 515 с.
4. BALABAN A.T., MOTOC I., BONCHEV D., MEKENY - AN O. Topological Indices for Structure-Activity Cor - relation //Topics Curr.Chem.- 1983.-Vol.114.-P. 21-55.
5. Химические приложения топологии и теории графов: Пер.с англ./Под ред.Р.Кинга. - М.: Мир, 1987. - 560 с.
6. КОХОВ В.А. Методы и программное обеспечение для срав - нительного анализа граф-моделей информационных систем //Мате - риалы международной НТК "Проблемы функционирования информаци - онных сетей". - Новосибирск, 1991. - С. 191-198.
7. JOHNSON M., NAIM M., NICHOLSON V., TSAI C. Unique mathematical features of the Substructure Met - ric Approach to Qvantitative Molecular Similarity Analysis //Studies in Physical and Theoretical Che - mistry. - 1987.-Vol. 51. -P. 219-225.
8. МЕТЕЛЬСКИЙ Н.Н., КОРНИЕНКО Н.М. Метризация одного клас - са графов //ДАН БССР. - 1979. - Т.23. - С. 5-7.
9. КУЗНЕЦОВ С.О., ФИНН В.К. Распространение процедур экс - пертных систем типа DCM на графы //Техническая кибернетика. - 1988. - №5. -С. 5-15.
10. RANDIC M. Characterization of Atoms, Molecu - les, and Classes of Molecules Based on Path Enumeration //Math.Chem. - 1979.-Vol. 7. - P. 5-64.
11. БЕРЦ С., ХЕРНДОН У. Подобие в графах и молекулах /Ис - кусственный интеллект: применение в химии: Пер. с англ. /Под ред. Т.Пирса, Б.Хони. - М.: Мир, 1988. - С.199-206.
12. ROUVRAY D.H. The Challenge of Chracterizing Branching in Molecular species //Discrete Applied Mathematics. - 1988.-Vol. 19.-P. 317-338.

13. BONCHEV D., POLANSKY O. On the Topological Complexity of Chemical Systems //Studies in Physical and Theoretical Chemistry. - 1987.-Vol. 51. - P.126-158.
14. KOKHOV V.A. Methods, Algorithms and Programs for Analysis of Molecular Graph Similarity //The Fourth Japan-USSR Symposium on Computer Chemistry. October 28-30, 1991. Toyohashi University of Technology, Japan. - 1991. -P. 53-54.
15. Харари Ф. Теория графов. - М.: Мир, 1973. - 300 с.
16. SKOROBOGATOV V.A., DOBRYNIN A.A. Metric Analysis of Graphs //Communication in Mathematical Chemistry. - 1988.-N 23. - P.105-151.
17. ПЕТРЕНЮК А.Я. Применение инвариантов в комбинаторных исследованиях //Вопросы кибернетики. - М.: Изд-во АН СССР, 1973. - С. 129-135.
18. ПОСПЕЛОВ Д.А. Ситуационное управление: теория и практика. - М.: Наука, 1986.
19. АЙЗЕРМАН М.А., ГУСЕВ Л.А., ПЕТРОВ С.В. и др. Динамический подход к анализу структур, описываемых графами (основы графодинамики) //Исследования по теории структур. - М.: Наука, 1988. - С. 5-76.

Поступила в ред.-изд.отд.

12 августа 1993 года