

УДК 681.3

В.В. Валяев, ЗАО «Белгран», инженер-программист (e-mail: funkybee@yandex.ru)

Э.И. Ватутин, канд. техн. наук, доцент, кафедра вычислительной техники, ЮЗГУ (e-mail: evatutin@rambler.ru)

МЕТОД ОПРЕДЕЛЕНИЯ ИЗОМОРФИЗМА ГРАФОВ ОБЩЕГО ВИДА ЗА ПОЛИНОМИАЛЬНОЕ ВРЕМЯ

Предложены метод и алгоритм определения изоморфизма графов с асимптотической временной сложностью $O(n^4)$ и затратами памяти $O(n^3)$. Метод основан на обходе графа группой окрестностей (волн) с определением полустепеней захода и исхода вершин на стыке окрестностей. За счет использования операций сортировки метод нечувствителен к порядку рассмотрения вершин в процессе построения окрестностей и построчного формирования инварианта, что позволяет считать получаемый инвариант предположительно полным.

Ключевые слова: теория графов, изоморфизм графов общего вида, полиномиально вычисляемый инвариант графа

Одной из широко известных задач в области теории графов является задача выяснения факта изоморфизма двух неориентированных графов $G_1 = \langle A_1, V_1 \rangle$ и $G_2 = \langle A_2, V_2 \rangle$ [1, 2], где $A_i = \{a_1^{(i)}, a_2^{(i)}, \dots, a_n^{(i)}\}$ – множество вершин, $V_i = \{v_1^{(i)}, v_2^{(i)}, \dots, v_m^{(i)}\}$ – множество дуг графа G_i , $\forall v_j^{(i)} = (a_{k_1}^{(i)}, a_{k_2}^{(i)}) \in V_i : (a_{k_1}^{(i)} \in A_i) \wedge (a_{k_2}^{(i)} \in A_i)$, $j = \overline{1, m}$, $i = \overline{1, 2}$. Графы считаются изоморфными ($G_1 \simeq G_2$) в том случае, если возможно отыскание взаимно-

однозначного соответствия $F = \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & \dots & a_n^{(1)} \\ a_{i_1}^{(2)} & a_{i_2}^{(2)} & \dots & a_{i_n}^{(2)} \end{pmatrix}$ между множествами

вершин, называемого также подстановкой изоморфизма, переводящего номера вершин $a_k^{(1)}$ в $a_{i_k}^{(2)}$ (по сути F является биективным бинарным отношением между множествами A_1 и A_2) с получением идентичного графа с совпадающей матрицей смежности M . Фактически в случае наличия изоморфизма граф G_2 может быть получен из графа G_1 путем перенумерации его вершин (рис. 1).

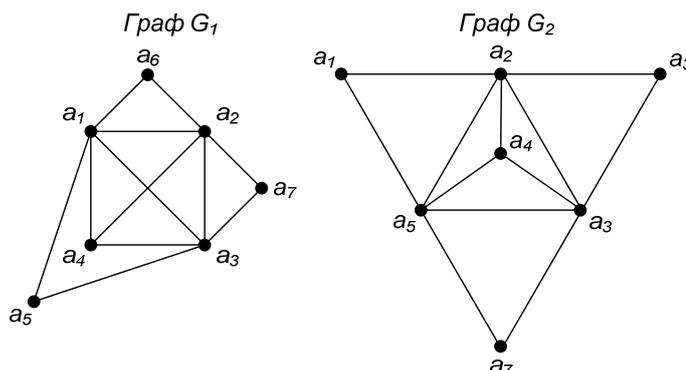


Рис. 1. Пример изоморфизма графов $G_1 \simeq G_2$ и одна из возможных подстановок изоморфизма $F = \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & a_4^{(1)} & a_5^{(1)} & a_6^{(1)} & a_7^{(1)} \\ a_2^{(2)} & a_5^{(2)} & a_6^{(2)} & a_4^{(2)} & a_3^{(2)} & a_1^{(2)} & a_7^{(2)} \end{pmatrix}$

Бинарное отношение изоморфизма графов фактически является отношением эквивалентности, что позволяет разбить множество графов с заданным числом вершин на классы эквивалентности [9]. Графы, входящие в каждый из классов эквивалентности, попарно изоморфны.

Задача изоморфизма графов примечательна тем, что за более чем полвека исследований не доказана ее принадлежность ни к классу полиномиально разрешимых P , ни к классу NP (при условии $P \neq NP$ [3]), что по-прежнему оставляет попытки ее исследования актуальными. Известны лишь частные случаи ее решения. Так, например, задача поиска изоморфного вхождения подграфа в граф общего вида относится к классу NP , а ряд частных случаев (для планарных графов, деревьев [4], R -выражений [5, 6]) может быть эффективно решен с использованием полиномиальных алгоритмов. В общем случае не известен полиномиальный алгоритм решения поставленной задачи общего вида, но и не доказано, что он не существует.

Решение задачи полным перебором [7] сводится к формированию $n!$ перестановок вершин графа G_1 (фактически, к перестановке строк и столбцов для матрицы смежности M_1) с последующим ее сопоставлением с матрицей смежности M_2 на предмет их совпадения. Каждое подобное сравнение матриц смежности требует n^2 действий, а подход в общем характеризуется асимптотической временной сложностью порядка $O(n^2 \cdot n!)$, что существенно ограничивает его применение на практике.

Большинство попыток полиномиального решения задачи сводится к отысканию полного инварианта графа [1, 8], вычислимого за полиномиальное время. При этом под полным инвариантом понимается некая числовая характеристика либо множество характеристик графа, совпадения которых необходимо и достаточно для установления факта изоморфизма. На данный момент подобный полный инвариант либо неизвестен, либо трудновычислим (примерами подобных инвариантов являются хроматическое число, мини- и максикоды матрицы смежности и др.), либо не является полным (например, совпадение отсортированных векторов степеней вершин является необходимым, но не достаточным условием изоморфизма). В работе предложен метод вычисления предположительно полного инварианта для неориентированных графов общего вида за полиномиальное время.

Суть метода сводится к следующему. На первом шаге формируется начальная окрестность q_0 , из которой начинается дальнейшее движение по вершинам графа. На каждом последующем шаге алгоритма для вершин текущей окрестности q_i определяется множество смежных с ней вершин q_{i+1} ,

для которых существуют дуги (a_{j_k}, a_{l_m}) , причем $(a_{j_k} \in q_i) \wedge (a_{l_m} \in q_{i+1})$ и вершины из множества q_{i+1} еще не были рассмотрены: $q_{i+1} \cap \bigcup_{j=0}^i q_j = \emptyset$.

Формирование окрестностей продолжается до тех пор, пока это возможно, т.е. пока из текущей окрестности можно попасть хотя бы в одну еще не рассмотренную вершину: $\exists a_{l_m} \wedge \exists (a_{j_k}, a_{l_m}) : (a_{j_k} \in q_i) \wedge \left(a_{l_m} \notin \bigcup_{p=0}^i q_p \right)$.

Схематично процесс формирования окрестностей приведен на рис. 2.

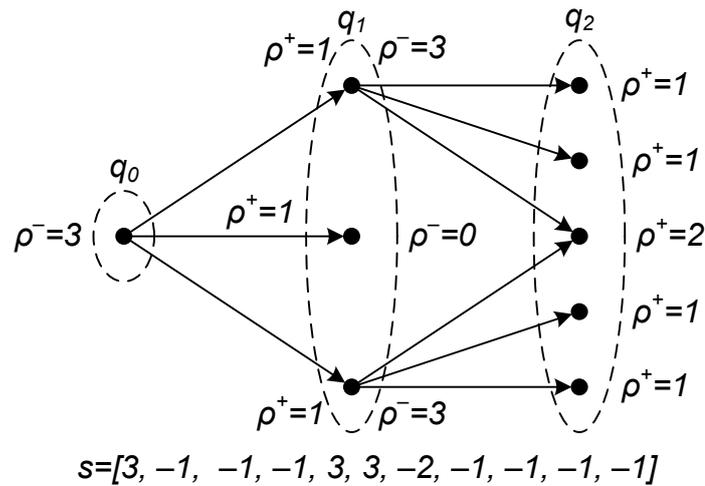


Рис. 2. Иллюстрация к процессу формирования окрестностей

Далее производится рассмотрение пар окрестностей (q_i, q_{i+1}) . Для каждой пары определяется множество связывающих дуг $V_{(q_i, q_{i+1})} = \{v_{j_1}, v_{j_2}, \dots, v_{j_r}\}$, таких что $\forall v_{j_k} = (a_{j_k}^{(нач)}, a_{j_k}^{(кон)}) \in V_{(q_i, q_{i+1})} : (a_{j_k}^{(нач)} \in q_i) \wedge (a_{j_k}^{(кон)} \in q_{i+1})$. Отталкиваясь от дуг данного множества $V_{(q_i, q_{i+1})}$ для всех вершин множества q_i рассчитываются оценки $\rho^+(a_{j_k}^{(нач)})$, а для вершин множества q_{i+1} – оценки $\rho^-(a_{j_k}^{(кон)})$, фактически представляющие собой полустепени исхода и захода для вершин множеств q_i и q_{i+1} соответственно, рассчитанные для ориентированного двудольного графа $G_{(q_i, q_{i+1})} = \langle q_i \cup q_{i+1}, V_{(q_i, q_{i+1})} \rangle \subseteq G$. Полученные оценки группируются в вектора $P_{(q_i, q_{i+1})}^- = [\rho^-(a_{j_1}^{(нач)}), \rho^-(a_{j_2}^{(нач)}), \dots, \rho^-(a_{j_r}^{(нач)})]$ и $P_{(q_i, q_{i+1})}^+ = [\rho^+(a_{j_1}^{(кон)}), \rho^+(a_{j_2}^{(кон)}), \dots, \rho^+(a_{j_r}^{(кон)})]$. Для приведенного выше примера

$$P_{(q_0, q_1)}^- = [3], \quad P_{(q_0, q_1)}^+ = [1, 1, 1];$$

$$P_{(q_1, q_2)}^- = [3, 0, 3], \quad P_{(q_1, q_2)}^+ = [1, 1, 2, 1, 1].$$

Далее из полученных векторов исключаются нулевые оценки и производится сортировка оценок в каждом из них (тип сортировки – по возрастанию или убыванию – не играет роли, в данном случае произведена сортировка по убыванию). Для приведенного выше примера

$$P_{(q_0, q_1)}^{-\prime} = [3], \quad P_{(q_0, q_1)}^{+\prime} = [1, 1, 1];$$

$$P_{(q_1, q_2)}^{-\prime} = [3, 3], \quad P_{(q_1, q_2)}^{+\prime} = [2, 1, 1, 1, 1].$$

Далее производится формирование строки инварианта графа путем конкатенации полученных векторов, причем оценки векторов $P_{(q_i, q_{i+1})}^{+\prime}$ берутся с обратным знаком. Для приведенного выше примера

$$S(q_0) = P_{(q_0, q_1)}^{-\prime} \odot P_{(q_0, q_1)}^{+\prime} \odot P_{(q_0, q_1)}^{-\prime} \odot P_{(q_0, q_1)}^{+\prime} = [3 | -1 -1 -1 | 3 3 | -2 -1 -1 -1],$$

где « \odot » – обозначение операции конкатенации, а символы «|» ограничивают отдельные вектора и добавлены для наглядности.

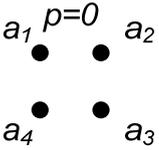
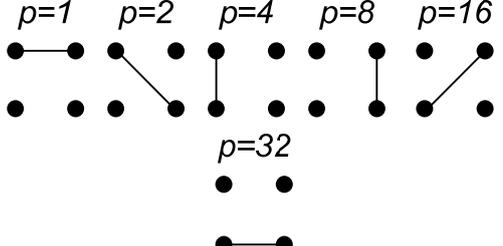
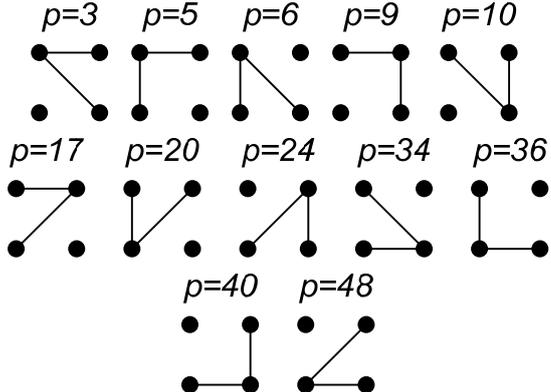
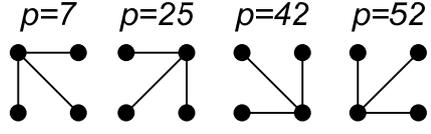
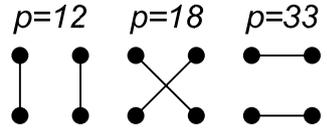
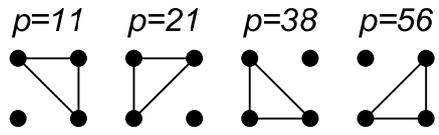
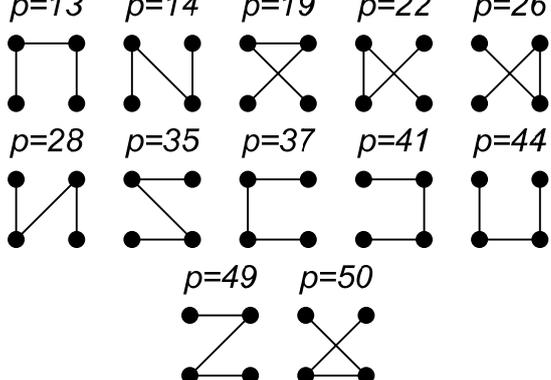
Выбирая различные наполнения начальной окрестности $q_0^{(i)}$ можно получать различные строки инварианта $S(q_0^{(i)})$. Отсортированная в лексикографическом порядке совокупность строк $[S(q_0^{(1)}), S(q_0^{(2)}), \dots, S(q_0^{(t)})]$ образует искомый предположительно полный инвариант.

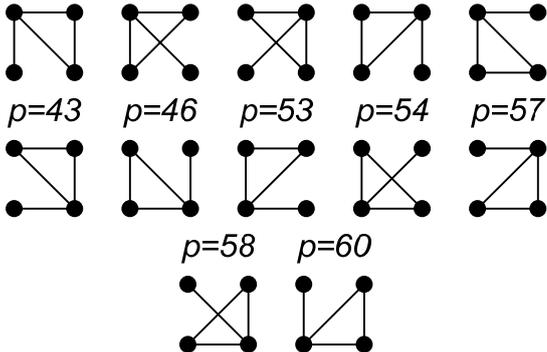
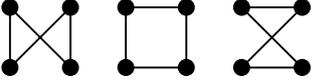
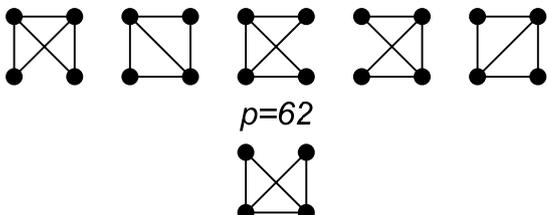
Формирование множества начальных окрестностей $\{q_0^{(1)}, q_0^{(2)}, \dots, q_0^{(t)}\}$ в простейшем случае осуществляется путем помещения каждой из вершин $a_i \in A$ графа G в отдельную начальную окрестность: $q_0^{(1)} = \{a_0\}$, $q_0^{(2)} = \{a_1\}$, $q_0^{(3)} = \{a_2\}$ и т.д. Более трудоемким является формирование начальных окрестностей из всех возможных пар $(\{a_0, a_1\}, \{a_0, a_2\}, \dots, \{a_0, a_n\}, \{a_1, a_2\}, \{a_1, a_3\}, \dots, \{a_1, a_n\}, \dots, \{a_{n-1}, a_n\})$, троек, четверок и т.д. вершин графа.

В таблице в качестве примера приведено разбиение множества всех графов из 4 вершин на классы эквивалентности [9], полученное с использованием предложенного инварианта. В качестве начальных значений окрестностей использована простейшая стратегия поочередного помещения в них вершин графа по одной в каждую окрестность.

Таблица. Множество классов изоморфизма для графов из 4 вершин, полученное с использованием предложенного инварианта (p – код матрицы смежности, полученный путем построчного выписывания элементов верхней треугольной подматрицы и перевода полученного двоичного числа в десятичную форму)

Класс изоморфизма	Примеры графов	Инвариант
-------------------	----------------	-----------

1		$\begin{pmatrix} - \\ - \\ - \\ - \end{pmatrix}$
2		$\begin{pmatrix} - \\ - \\ 1 & -1 \\ 1 & -1 \end{pmatrix}$
3		$\begin{pmatrix} - \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 2 & -1 & -1 & \end{pmatrix}$
4		$\begin{pmatrix} 1 & -1 & 2 & -1 & -1 \\ 1 & -1 & 2 & -1 & -1 \\ 1 & -1 & 2 & -1 & -1 \\ 3 & -1 & -1 & -1 & \end{pmatrix}$
5		$\begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}$
6		$\begin{pmatrix} - \\ 2 & -1 & -1 \\ 2 & -1 & -1 \\ 2 & -1 & -1 \end{pmatrix}$
7		$\begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 2 & -1 & -1 & 1 & -1 \\ 2 & -1 & -1 & 1 & -1 \end{pmatrix}$

8		$\begin{pmatrix} 1 & -1 & 2 & -1 & -1 \\ 2 & -1 & -1 & 1 & -1 \\ 2 & -1 & -1 & 1 & -1 \\ 3 & -1 & -1 & -1 & \end{pmatrix}$
9		$\begin{pmatrix} 2 & -1 & -1 & 1 & 1 & -2 \\ 2 & -1 & -1 & 1 & 1 & -2 \\ 2 & -1 & -1 & 1 & 1 & -2 \\ 2 & -1 & -1 & 1 & 1 & -2 \end{pmatrix}$
10		$\begin{pmatrix} 2 & -1 & -1 & 1 & 1 & -2 \\ 2 & -1 & -1 & 1 & 1 & -2 \\ 3 & -1 & -1 & -1 & & \\ 3 & -1 & -1 & -1 & & \end{pmatrix}$
11		$\begin{pmatrix} 3 & -1 & -1 & -1 \\ 3 & -1 & -1 & -1 \\ 3 & -1 & -1 & -1 \\ 3 & -1 & -1 & -1 \end{pmatrix}$

Из приведенного пример видно, что первый столбец инварианта представляет собой отсортированный вектор степеней вершин графа (при условии, что символы « \leftrightarrow » заменяются на нули).

Для формирования множества начальных окрестностей в простейшем случае необходимо n шагов; число получаемых окрестностей не превышает диаметр графа $d(G) \leq n$; вычисление оценок ρ^+ и ρ^- требует не более $|q_i| \cdot |q_{i+1}| < n^2$ действий; сортировка векторов P^+ и P^- выполняется за время порядка $O(|q_i|)$ и $O(|q_{i+1}|)$ при использовании сортировки распределением [10], каждое из которых не превосходит $O(n)$; итоговая сортировка строк инварианта требует $O(n \log n)$ действий. Таким образом, временные затраты на вычисление инварианта составляют $O(n \cdot \alpha_1 n \cdot (\alpha_2 n^2 + \alpha_3 n + \alpha_4 n) + \alpha_5 n \log n) \simeq O(n^4)$, где $\alpha_i, i = \overline{1, 5}$ – некоторые коэффициенты. Затраты памяти определяются затратами на хранение инварианта и составляют в простейшем случае величину $O(n \cdot d(G) \cdot (|\bar{q}_i| + |\bar{q}_{i+1}|)) \simeq O(n^3)$.

В перспективе дальнейших исследований необходимо исследование вопроса о том, является ли предложенный инвариант действительно полным для графов с большим числом вершин. В пределах графов, входящих в состав класса эквивалентности, инварианты скорее всего должны совпадать. Если при этом не будет наблюдаться совпадения инвариантов для графов из различных классов эквивалентности, инвариант можно эмпирически считать полным.

Список литературы

1. Зыков А.А. Основы теории графов. М.: Наука, 1986. 384 с.
2. http://ru.wikipedia.org/wiki/Изоморфизм_графов
3. http://ru.wikipedia.org/wiki/Равенство_классов_P_и_NP
4. http://en.wikipedia.org/wiki/Graph_isomorphism_problem
5. Ватулин Э.И., Зотов И.В., Титов В.С. Выявление изоморфных вхождений R -выражений при построении множества сечений параллельных алгоритмов логического управления // Информационно-измерительные и управляющие системы. № 11, Т. 7. М.: «Радиотехника», 2009. С. 49–56.
6. Ватулин Э.И., Зотов И.В., Титов В.С. Алгоритм и устройство выявления изоморфных вхождений R -выражений при построении множества сечений параллельных алгоритмов логического управления // Известия вузов. Приборостроение. 2009. Т. 52, № 2. С. 37–45.
7. Курейчик В.М., Глушань В.М., Щербаков Л.И. Комбинаторные аппаратные модели и алгоритмы в САПР. М.: Радио и связь, 1990. 216 с.
8. http://ru.wikipedia.org/wiki/Инвариант_графа
9. http://ru.wikipedia.org/wiki/Класс_эквивалентности
10. http://algotlist.manual.ru/sort/radix_sort.php