

УДК 519.17:004.023

ОБ ОДНОМ ЭВРИСТИЧЕСКОМ КРИТЕРИИ В ЗАДАЧЕ ОПРЕДЕЛЕНИЯ ИЗОМОРФИЗМА ГРАФОВ НА ОСНОВЕ ИНВАРИАНТОВ

^{1,2,3}Хачумов М.В., ³Талалаев А.А., ^{1,2,3}Хачумов В.М.

¹ФИЦ «Информатика и управление» РАН, Москва, e-mail: khmike@inbox.ru;

²ФГАОУ ВО «Российский университет дружбы народов», Москва, e-mail: vmh48@mail.ru;

³ФГБУН «Институт программных систем им. А.К. Айламазяна» РАН, Вельково,
e-mail: arts@arts.botik.ru

Рассматривается эвристический критерий, направленный на преобразование сравниваемых связанных графов для решения задачи определения их изоморфизма. Критерий связан с построением кортежей графов с минимальной суммарной длиной ребер, вычисляемой по определенной метрике, и приводит к получению дополнительных полезных инвариантных характеристик. В качестве атрибутов здесь выступают суммы длин ребер при вершинах оптимизированного кортежа, которые могут быть применены в составе вектора инвариантных характеристик в задаче определения изоморфизма. Ранее подобный критерий был успешно использован в задаче разбиения графа на части с примерно равным числом вершин и с минимальным числом связей между подграфами. В матрице смежности такому варианту графа соответствует максимальная концентрация единиц вдоль главной диагонали, что служит в качестве другой эвристики, которая может быть использована для направленной перестановки строк и столбцов. Выполнен обзор работ, в том числе рассматривающих проблему построения полного инварианта, служащего необходимым и достаточным условием изоморфизма. Приведен пример успешного применения эвристического критерия в задаче определения изоморфизма для двух графов с переименованными вершинами. Рассмотренный подход целесообразно применять в задачах разбиения программ на слабо связанные участки и распознавания образов на основе графовых моделей.

Ключевые слова: граф, изоморфизм, инвариант, эвристический критерий, кортеж, длина ребра, суммарная длина ребер

ON A HEURISTIC CRITERION IN THE PROBLEM OF DETERMINING THE ISOMORPHISM OF GRAPHS BASED ON INVARIANTS

^{1,2,3}Khachumov M.V., ³Talalaev A.A., ^{1,2,3}Khachumov V.M.

¹FRC CSC RAS, Moscow, e-mail: khmike@inbox.ru;

²Peoples' Friendship University of Russia, Moscow, e-mail: vmh48@mail.ru;

³Program Systems Institute of RAS, Veskovo, e-mail: arts@arts.botik.ru

We consider a heuristic criterion aimed at transforming compared connected graphs to solve the problem of determining their isomorphism. The criterion is related to the construction of tuples of graphs with the minimum total length of edges, calculated by a certain metric, and leads to additional useful invariant characteristics. The attributes here are the sums of the lengths of the edges at the vertices of the optimized tuple, which can be used as part of the vector of invariant characteristics in the problem of isomorphism determination. Previously, a similar criterion was successfully used in the problem of splitting a graph into parts with an approximately equal number of vertices and with a minimum number of connections between subgraphs. In the adjacency matrix, this version of the graph corresponds to the maximum concentration of ones along the main diagonal, which serves as another heuristic that can be used for directed permutation of rows and columns. We made a review of papers, including those dealing with the problem of constructing a complete invariant, which serves as a necessary and sufficient condition for isomorphism. An example of successful application of a heuristic criterion in the problem of isomorphism determination for two graphs with renamed vertices is given. It is expedient to apply the considered approach in problems of splitting programs into weakly connected sections and pattern recognition based on graph models.

Keywords: graph, isomorphism, invariant, heuristic criterion, tuple, edge length, total length of edges

Определение изоморфизма графов на основе инвариантов – один из эффективных подходов к распознаванию объектов различной природы. Показано экспериментально, что при определении изоморфизма графов с малым числом вершин высокую эффективность имеет вектор степеней вершин, но в случае восьми вершин и более неизоморфные графы могут иметь идентичные векторы [1]. Достаточно эффективным оказался вектор собственных чисел матрицы смежности графа, что позволяет

вычислять изоморфизм некоторых графов за полиномиальное время. Отмечается, что вектор степеней вершин и характеристики, извлекаемые из матриц смежности, несомненно, являются полезными инвариантами, но не решают в общем случае задачу изоморфизма. В работе [2] в качестве инвариантов предлагается использовать объединения локальных характеристик графа, например плотность, хроматическое число, число Хадwigера, индексы Виннера и Рандича. Менее эффективны

следующие инварианты: определитель матрицы смежности, число компонент связности и диаметр графа. Алгоритм вычисления полного инварианта графа изложен в работе [3]. Предложена система кодирования, которая обеспечивает получение интегрального описателя структуры, инвариантного относительно исходной нумерации её вершин.

Работа направлена на поиск новых инвариантных характеристик, получаемых в результате целенаправленного преобразования графа эвристическим алгоритмом за полиномиальное время. Важно отметить, что проблема распознавания изоморфизма за полиномиальное время является одной из нерешенных проблем современной теории сложности вычислений. Поэтому эффективными являются исследования, опирающиеся на эвристику. Сочетание точных алгоритмов и эвристических, основанных на итерационном уточнении решений, рассмотрено, например, в работе [4]. В статье [5] представлена эвристика для решения проблемы изоморфизма графов методами вероятностного линейного программирования. В настоящей работе выполнены исследования в рамках поиска устойчивых эвристических характеристик, способствующих выявлению изоморфизма.

Материалы и методы исследования

В настоящей работе в качестве исходного материала для решения задачи изоморфизма рассматриваются простые связные графы. Пусть исходные данные представлены в виде графа $G = (V, E)$, где V – множество вершин; E – множество ребер, содержащего n вершин и m ребер.

В качестве метода исследования выступает проверка гипотезы о положительных инвариантных свойствах кортежа графа, характеризуемого минимумом суммарной длины всех связей в кортеже, а в качестве атрибута для определения изоморфизма, – собственно упорядоченный вектор длин, приписываемых вершинам кортежа. Для оценки качества кортежа вводится понятие длины ребра как неотрицательной разности порядковых номеров вершин в кортеже, инцидентных данному ребру. С учетом этого суммарная длина ребер полного графа равна $L_n = \sum_{i=1}^N i(n-i)$, где $N = \frac{n(n-1)}{2}$

есть число ребер графа. Таким образом, к матрице смежности, задающей граф с помеченными вершинами, вектору степеней вершин и другим его инвариантам добавляются новые полезные элементы [6].

Дадим некоторые определения.

Функцию f , относящую к каждому графу G некоторый элемент $f(G)$ из множества M , будем называть инвариантом, если на изоморфных графах ее значения совпадают, т.е. для любых G и G' : $G \cong G' \Rightarrow f(G) = f(G')$.

Инвариант графа, удовлетворяющий необходимому и достаточному условию изоморфизма, т.е. $G \cong G' \Leftrightarrow f(G) = f(G')$, будем называть полным. Если на сравниваемых графах значения функции f отличаются на допустимые величины Δ , т.е. для любых L и L' $G \cong G' \Rightarrow |f(G) - f(G')| \leq \Delta$, то будем называть ее субинвариантом. Такой подход обеспечивает возможность приближенного сравнения графовых моделей с допустимой погрешностью.

Для получения новых элементов, дополняющих вектор инвариантов, рассмотрим эвристический метод построения кортежа графа с требуемыми свойствами. Суть метода состоит в формировании кортежа вершин графа G , в котором вершины с большим числом связей между собой размещаются в кортеже по соседству, а слабо связанные вершины – в отдаленных друг от друга позициях. Алгоритм, приведенный в работе [7], позволяет в цикле строить кортеж графа G с минимальной суммарной длиной ребер. Количество глобальных циклов n определяется числом вершин в графе. Алгоритм в каждом цикле (шаге) с номером r оперирует двумя подграфами D_1^r и D_2^r графа $G = (V, E)$ с выбранными и еще не выбранными вершинами графа соответственно. На каждом r цикле выбирается и отмечается индексом r одна из неотмеченных вершин $a_i \in V$ подграфа D_2^r , связанная непосредственно с одной или несколькими вершинами подграфа D_1^r . Эта вершина переносится из подграфа D_2^r в граф D_1^r и служит очередной вершиной формируемого кортежа. При выборе r -й компоненты кортежа на r -м цикле алгоритма отыскивается такая вершина a_i^r , для которой $\xi_{1i}^r - \xi_{2i}^r = \delta_{i\max}^r$, где ξ_{1i}^r – число дуг, связывающих вершину a_i с D_1^{r-1} а ξ_{2i}^r – число дуг, связывающих a_i с D_2^r . Каждый цикл связан с анализом и оценкой $n - r$ еще не выбранных вершин-кандидатов. Таким образом, в итоге получаем искомый кортеж с минимизированной суммарной длиной связей и соответствующую ему матрицу смежности. Вычислительную сложность алгоритма построения кортежа можно оценить как $n(n-1)$.

Результаты исследования и их обсуждение

Пусть заданы два графа $G_1 = (V_1, E_1)$ и $G_2 = (V_2, E_2)$, причем второй получен путем переименования вершин.

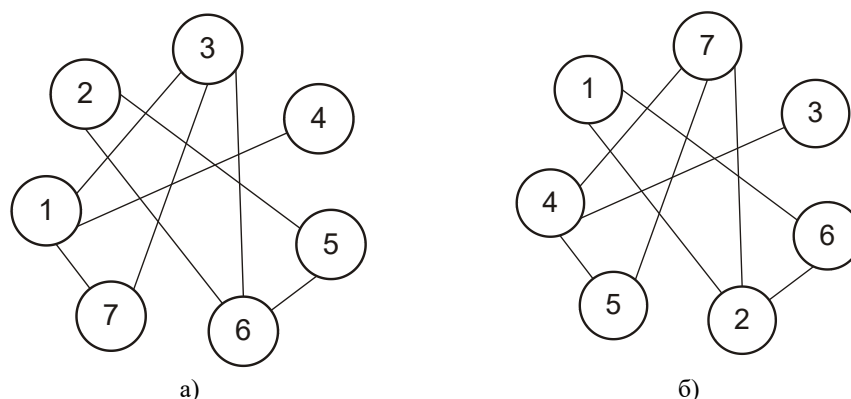


Рис. 1. Графы G_1 (а) и G_2 (б)

Таблица 1

Исходная матрица смежности графа G_1

	1	2	3	4	5	6	7
1	1		1	1			1
2		1			1	1	
3	1		1			1	1
4	1			1			
5		1			1	1	
6		1	1		1	1	
7	1		1				1

Матрица смежности графа G_1 приведена в табл. 1. Заметим, что в этой и последующих матрицах смежности в диагоналях для элементов с одинаковыми индексами содержатся «1», которые не означают наличия петель в графах, а служат исключительно для визуального анализа распределения единиц вокруг диагоналей. Кортеж, постро-

енный по данной матрице смежности, имеет вид, представленный на рис. 2, а. Суммарные длины ребер в кортеже в соответствии с их номерами (11, 7, 9, 3, 4, 8, 10). Общая длина восьми ребер без повторений $L_1 = 26$. Применяя алгоритм построения кортежа с минимальной суммарной длиной ребер, получим результат, приведенный на рис. 2, б.

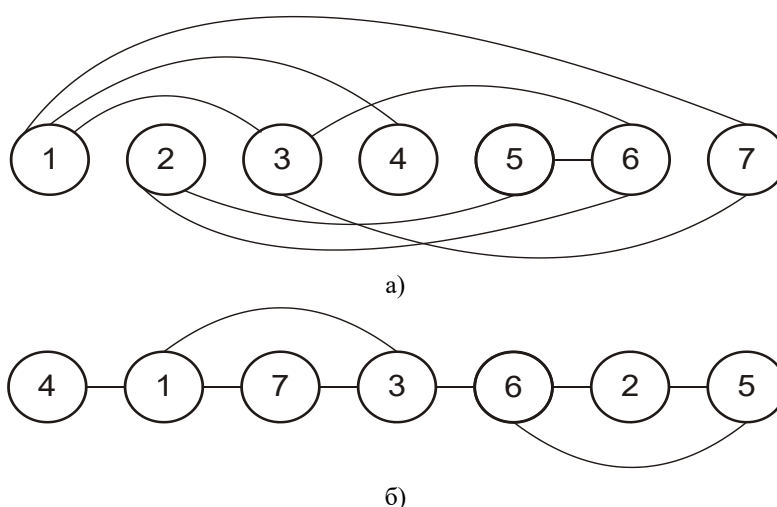


Рис. 2. Кортеж графа G_1 без оптимизации (а) и с минимальной суммарной длиной ребер (б)

Суммарные длины ребер при вершинах графа в оптимизированном кортеже в соответствии с их порядковыми номерами в кортеже (1, 4, 2, 4, 4, 2, 3). Общая длина ребер без повторений $L_1^o = 10$. Матрица смежности графа G_1 после упорядочения представлена в табл. 2.

Видно, что единицы группируются вдоль главной диагонали. Рассмотрим для сравнения и определения возможного

изоморфизма граф $G_2 = (V_2, E_2)$ с другой разметкой вершин (табл. 3).

Результаты его отображения в виде кортежа и упорядоченного кортежа в соответствии с алгоритмом, представлены на рис. 3, а и б.

В табл. 4 приведена преобразованная в соответствии с алгоритмом построения оптимизированного кортежа матрица смежности графа G_2 .

Таблица 2

Преобразованная матрица смежности графа G_1

	4	1	7	3	6	2	5
4	1	1					
1	1	1	1	1			
7		1	1	1			
3		1	1	1	1		
6				1	1	1	1
2					1	1	1
5					1	1	1

Таблица 3

Исходная матрица смежности графа G_2

	1	2	3	4	5	6	7
1	1	1				1	
2	1	1				1	1
3			1	1			
4			1	1	1		1
5				1	1		1
6	1	1				1	
7		1		1	1		1

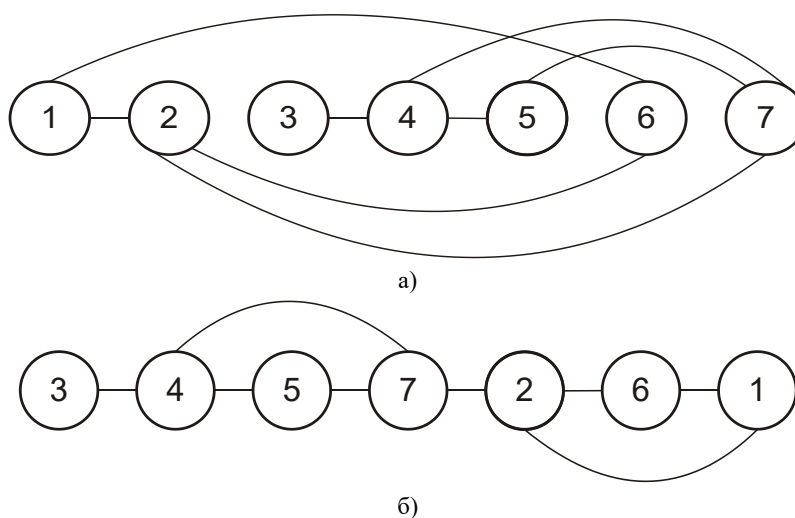


Рис. 3. Кортеж графа G_2 без оптимизации (а) и с минимальной суммарной длиной ребер (б)

Таблица 4

Преобразованная матрица смежности графа G_2

	3	4	5	7	2	6	1
3	1	1					
4	1	1	1	1			
5		1	1	1			
7		1	1	1	1		
2				1	1	1	1
6					1	1	1
1					1	1	1

Длины ребер при вершинах в кортеже (рис. 3, а) в соответствии с их номерами представлены вектором (6, 10, 1, 5, 3, 9, 10). Общая длина L_2 восьми ребер без повторов $L_2 = 22$. Применяя алгоритм построения кортежа, получим граф, представленный на рис. 3, б. Суммарные длины ребер при вершинах графа в оптимизированном кортеже в соответствии с их порядковыми номерами: в кортеже (1, 4, 2, 4, 4, 2, 3). Общая длина ребер в нем без повторов $L_2^o = 10$.

Таким образом, в рассмотренном примере графы G_1 и G_2 в оптимизированных кортежах имеют одинаковые векторы длин ребер (1, 4, 2, 4, 4, 2, 3) и равные показатели $L_1^o = L_2^o = 10$, что в интеграции с другими инвариантами позволяет считать их изоморфными.

Отметим отдельно, что применяемый алгоритм построения кортежа [7] допускает в общем случае наличие нескольких альтернативных вариантов. Поэтому в сочетании с эвристикой может потребоваться ограниченный перебор порядка вершин с целью достижения равенства характеристик оптимизированных кортежей для G_1 и G_2 .

Заключение

В настоящей работе предложено для определения изоморфизма воспользоваться дополнительно эвристическим алгоритмом построения кортежа графа, обладающего свойством минимума суммарной длины ребер по установленной метрике. На примере показано, что переименование вершин в графе $G_1 = (V_1, E_1)$ и его преобразование к графу $G_2 = (V_2, E_2)$ не приводит к изменению величин L_1^o, L_2^o и векторов длин ребер при вершинах оптимизированных кортежей графов. Рассмотренный алгоритм позволяет достаточно быстро строить

такие кортежи, а полученные на его основе числовые характеристики целесообразно использовать в составе векторов инвариантов для определения изоморфизма.

Работа выполнена при финансовой поддержке проекта РФФИ № 20-07-00022 А «Разработка и исследование методов распознавания образов на основе инвариантов к яркостным и геометрическим преобразованиям в системах технического зрения беспилотных летательных аппаратов».

Список литературы

- Орлова П.Н. Исследование алгоритмов решения задачи определения изоморфизма графов. [Электронный ресурс]. URL: https://dspace.tltsu.ru/xmlui/bitstream/handle/123456789/12335/Орлова%20П.Н._ПМИп-1601a.pdf (дата обращения: 18.01.2022).
- Dehmer M., Grabner M., Mowshowitz A., Emmert-Streib F. An efficient heuristic approach to detecting graph isomorphism based on combinations of highly discriminating invariants. *Advances in Computational Mathematics*. 2013. Vol. 39. No. 2. P. 311–325. DOI: 10.1007/s10444-012-9281-0.
- Погребной А.В. Полный инвариант графа и алгоритм его вычисления // Известия Томского политехнического университета. Информационные технологии. 2014. Т. 325. № 5. С. 110–122.
- Stoichev S.D. New exact and heuristic algorithms for graph automorphism group and graph isomorphism. *Journal of Experimental Algorithmics (JEA)*. 2019. Vol. 24. P. 1–27. DOI: 10.1145/3333250.
- Takapoui R., Boyd S. Linear Programming Heuristics for the Graph Isomorphism Problem // ArXiv.2016. [Электронный ресурс]. URL: <https://arxiv.org/pdf/1611.00711.pdf> (дата обращения: 18.01.2022).
- Агроник А.Ю., Кочина Л.В., Хачумов М.В. Имитационное моделирование и анализ технологических процессов сетями Петри // Приборы и системы. Управление, контроль, диагностика. 2017. № 4. С. 19–28.
- Топольский Н.Г., Святенко И.Ю., Трефилов Г.Б., Сатин А.П. Интерактивный оптимизационный метод декомпозиции графов причинно-следственных связей в системах поддержки принятия решений // Интернет-журнал «Технологии техносферной безопасности». 2009. № 5 (27).; URL: <http://agps-2006.narod.ru/ttb/2009-5/11-05-09.ttb.pdf> (дата обращения: 18.01.2022).