УДК 004.942: 62-932.4

А.Л. Становский, д.т.н., профессор,
О.С. Савельева, к.т.н., доцент,
М.Л. Герганов, к.т.н.,
Е.Ю. Лебедева, магистр,
Одесский национальный политехнический университет.

ПРОБЛЕМА РАСПОЗНАВАНИЯ ИЗОМОРФИЗМА ГРАФОВ И ОБРАТНАЯ ЗАДАЧА СТРУКТУРНОЙ НАДЕЖНОСТИ

О.Л. Становський, О.С. Савєльєва, М.Л. Герганов, О.Ю. Лєбедєва. **Проблема розпізнання ізоморфізму графів та зворотна задача структурної надійності.** Показано, що розв'язок задачі, зворотної визначенню структурної надійності складного об'єкта з резервуванням за допомогою інформаційної морфологічної моделі, породжує інваріант, який з високим ступенем ймовірності визначає граф з точністю до ізоморфізму. Запропоновано використовувати такий інваріант при моделюванні ізоморфних перетворень органічних молекул.

A.L. Stanovsky, O.S. Saveljeva, M.L. Gerganov, E.Ju. Lebedeva. **The counts isomorphism improvement problem and structural reliability return problem.** The solution of a task, which is return to the determination of difficult object structural reliability with reservation by means of information morphological model, generates an invariant, with high degree of probability defining count to within of isomorphism is shown. It is offered to use such invariant when modeling isomorphic transformations of organic molecules.

Введение. Как известно, изоморфизм – наличие структурного подобия у разных объектов. Изоморфизм всегда задает отношение эквивалентности на классе множеств со структурой. Многие практические задачи – от конструирования новых машин на макроуровне до создания новых молекул на микро – приводят к необходимости распознавания изоморфизма и изоморфного вложения сложных структур, заданных в виде графов. Любой абстрактный граф идентичен, или, как говорят математики, изоморфен некоторому геометрическому графу. При изображении геометрических графов имеется большая свобода в размещении вершин и в выборе формы соединяющих их ребер. Поэтому может оказаться, что один и тот же граф представляется различными чертежами, разглядывая которые далеко не сразу можно осознать, что они являются изображениями одного и того же графа.

1. Постановка проблемы. Вопрос о том, изоморфны ли два данных графа, в общем случае оказывается очень сложным [1]. С содержательной точки зрения изоморфизм графов структур означает тождественность функционирования самих структур, что допускает в некоторых случаях замену одной структуры другой, ей изоморфной. Проблема установления изоморфизма состоит в нахождении наиболее эффективного алгоритма, который распознает, изоморфны или нет рассматриваемые графы. Но, к сожалению, до сих пор даже не установлено, принадлежит она к классу P проблем, для которых существует алгоритм полиномиальной сложности, или является NP-полной проблемой. Поэтому, для того чтобы выяснить, изоморфны ли два графа имеющие n вершин, в общем случае приходится выполнять $n!$ попарных сравнений, а для распознавания изоморфного вложения графа **H**, имеющего m вершин, в граф **G**, у которого n вершин, необходимо провести $C_n^m \cdot m!$ сравнений. Уже при об изоморфизме таким методом полного перебора становится недостижимым: уже при $n = 20$ перебор всех $n!$ вариантов потребовал бы 40лет машинного времени.

Подобная ситуация, естественно, направила исследователей на путь поиска такого ***инварианта*** (числа или системы чисел), который бы, с одной стороны, легко вычислялся по заданному графу, а с другой – обладал ***свойством полноты***, т.е. определял граф однозначно с точностью до изоморфизма.

К сожалению, воображаемая простота такого подхода сталкивается с проблемами определения как самих инвариантов (поиск обязательно должен содержать доказательство адекватности), так и методов их относительно быстрого вычисления, которое возможно выполнить лишь для графов определенных типов с соблюдением соответствующих ограничений.

2. Анализ последних достижений и публикаций. Пусть **f** – функция, относящая каждому графу **G** некоторый элемент **f(G)** из множества **M** произвольной природы. Эту функцию мы будем называть ***инвариантом***, если на изоморфных графах ее значения совпадают, т.е. для любых **G** и **G'** из изоморфности графов **G** и **G'** следует **f(G) = f(G')**.

Рассмотрим несколько наиболее важных инвариантов графа [2]. Прежде всего, это *количество вершин* $n(\mathbf{G})$ и *ребер* $m(\mathbf{G})$, *вектор степеней вершин* $\mathbf{s(G)} = (s_1, s_2, ..., s_n)$, а также *плотность* **k(G)** – количество вершин клики графа **G**, *неплотность* **e(G)** – наибольшее количество попарно несмежных вершин графа, *хроматическое число* **g(G)**; *количество компонент связности* **K(G)**, *число Хадвигера* **h(G)**. Любая функция от элементов a_{ij} матрицы, не меняющаяся ни при каких перестановках рядов матрицы смежности, является инвариантом графа **G**.

Известно, что графы изоморфны тогда и только тогда, когда их *матрицы смежностей* получаются друг из друга одинаковыми перестановками строк и столбцов или когда их матрицы инцидентности получаются друг из друга произвольными перестановками строк и столбцов [3]. Однако, к сожалению, матрица смежностей не является инвариантом графа: при переходе от одной нумерации его вершин к другой она претерпевает перестановку рядов, состоящую из некоторой перестановки строк и точно такой же перестановки столбцов.

Инвариант **f** называется ***полным***, если для любых графов **G** и **G'** из равенства **f(G) = f(G')** следует изоморфизм графов **G** и **G'**. К сожалению, из перечисленных инвариантов ни один не является полным. В процессе развития теории графов не было нехватки в гипотезах полноты инварианта, использовали даже эвристические методы [4], но все эти гипотезы рано или поздно опровергались конкретными примерами. Последовательно воссоздавая [5] или разрушая [6] сравниваемые графы, в некоторых случаях удается распознать искомый изоморфизм, однако и здесь попытки получить детеминированный инвариант ограничены доказательной базой и объемом машинного времени.

3. Целью работы является создание стохастического инварианта для распознавания изоморфизма графов, основанного на методе изъятия элементов, который бы дополнял существующие методы и позволял в реальном времени, например, конструирования или управления поддержать решение о наличии искомого изоморфизма.

4. Основной материал. Для нахождения такого инварианта предложен следующий подход. Пусть имеется два графа $\mathbf{G}_1$ и $\mathbf{G}_2$, изоморфность которых требуется распознать (рис. 1).

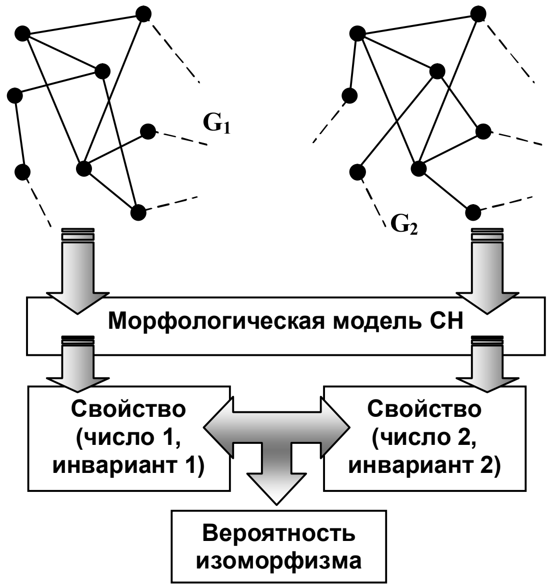

Рис. 1. Схема процесса распознавания изоморфности графов с помощью морфологической модели структурной надежности систем с резервированием.

Построим интеллектуальную морфологическую компьютерную модель (ИММ) структурной надежности (СН) некоторой сложной системы с резервированием (ССР), изоморфную поочередно сравниваемым графам, возможности которой позволяли бы подвергать ССР испытаниям на СН [7] в соответствии с приведенной к неравнозначным элементам стандартной формулой «с выбыванием» для вычисления вероятности безотказной работы (ВБР) таких систем:

$$P(t) = \sum_{i=0}^{m} \left\{ C_n^{m-i} \cdot \prod_{j=1}^{n-i} P_{ej}(t) \cdot \left[1 - \prod_{j=1}^{n-i} P_{ej}(t) \right]^i \right\}, \quad (1)$$

где $P(t)$ – ВБР системы; $P_{ej}(t)$ – ВБР элементов системы, m – количество резервных элементов.

Действительно, из структуры (1) непосредственно вытекает, что оценка ВБР ССР для любого момента времени t требует последовательного определения (расчета) ВБР исходной системы, ВБР системы с одним отказавшим элементом, ВБР системы с двумя отказавшими элементами и т.д. вплоть до ВБР системы с m отказавшими элементами из n, после чего система считается неработоспособной в целом. Отсюда вытекают требуемые свойства ИММ: она должна, с одной стороны, «сама себя повреждать», воссоздавая поитерационно в режиме реального времени состояние объекта с точки зрения удаления отказавших элементов, а с другой, – также быстро рассчитывать ВБР объекта на каждой итерации. Такая модель опирается на следующую научную гипотезу.

Гипотеза: эксперимент над информационной морфологической моделью ССР, алгоритм которого сответствует формуле вычисления ВБР этой системы по (1), является адекватным с точки зрения прогноза ВБР реальной ССР, а инварианты модели и объекта в виде значений ВБР, полученных в результате их испытаний, отличаются ε-неразличимостью и обладают Δ-инвариантностью.

Если Гипотеза верна, то имеет место такое обратное следствие.

Следствие: *если межу двумя исходными состояниями ССР существует изоморфизм, что свидетельствует об инвариантности их графов, то при изоморфном разрушении этих объектов будет наблюдаться подобие между статистическими оценками их работоспособности, независимо от физического типа объектов и задач, которые эти объекты решают, и наоборот, – если статистические оценки их работоспособности ε-неразличимы и Δ-инвариантны то это свидетельствует об инвариантности их исходных графов с точностью до вероятности оценок ВБР систем.*

Таким образом, в качестве инварианта графов, о котором шла речь выше, предлагается рассчитанное с помощью ИММ численное значение ВБР. Для этого была предложена энтропийная формула, значения вероятностей P_i для которых определяются экспериментами на ИММ:

$$\text{ÂÁÐ} = -\sum_{i=R_1+1}^{R_2-1}\left[\text{Ð}_i \log_2 \text{Ð}_i + \left(1-\text{Ð}_i\right)\log_2\left(1-\text{Ð}_i\right)\right], \qquad (2)$$

где R_1 – количество отказавших элементов, при которых ВБР испытуемой системы *еще* равна 1 (полный детерминизм), R_2 – количество отказавших элементов, при которых ВБР ССР *уже* равна 0,5 (полный хаос).

Примером практического применения такого инварианта является использование его в качестве дескриптора химической структуры, – числа, которое однозначно характеризует структуру органического соединения (например, лекарства) с точки зрения некоторого важного свойства [8]. В настоящее время выбор дескриптора – долгая и сложная задача: сначала группу соединений с известной структурой и известными значениями физиологической активности, полученными экспериментально, делят на две части – тренировочный и тестовый наборы. В этих наборах числа, характеризующие активность, уже сопоставлены конкретной структуре. Далее выбирают дескрипторы (обычно, их сотни!) и строят математическую модель активности для соединений из тренировочного набора. Правильность полученной математической модели проверяют на тестовом наборе структур. Сначала вычисляют инварианты для каждой из тест-структур, с помощью модели рассчитывают значение активности и сравнивают их с уже известными экспериментальными значениями. Применение предложенного метода и разработанной для него ИММ позволило в 7 – 10 раз сократить время, расходуемое на выявление изоморфных структур в медицинской химии.

Выводы. Решение задачи, обратной определению структурной надежности сложного объекта с резервированием с помощью информационной морфологической модели, порождает инвариант, который с высокой степенью вероятности определяет граф с точностью до изоморфизма. Использование такого инварианта при моделировании преобразований органических молекул в медицинской химии позволило значительно сократить время, расходуемое на выявление изоморфных структур.

Литература

1. Земляченко, В.Н. Проблема изоморфизма графов / В.Н. Земляченко, Н.М. Корнеенко, Р.И. Тышкевич // Теория сложности вычислений, I. Записки научных семинаров ЛОМИ. – 1982. – Т. 118. – С. 83 – 158.

2. Зыков, А.А. Основы теории графов. М.: Наука, 1987. – 384 с.

3. Лекции по теории графов / В.А. Емеличев, О.И. Мельников, В.И. Сарванов, Р.И. Тышкевич. – М.: Наука, 1990. – 384 с.

4. Гудман, С. Введение в разработку и анализ алгоритмов / С. Гудман, С. Хидетниеми. – М.: Мир, 1981. – 368 с.

5. Рейнгольд, Э. Комбинаторные алгоритмы. Теория и практика / Э. Рейнгольд, Ю. Нивергельт, Н. Део. – М.: Мир, 1980. – 476 с.

6. Татт, У. Теория графов. – М.: Мир, 1988. – 424 с.

7. Становський, О.Л. Критерії відмовостійкості складних технічних систем / О.Л. Становський, О.С. Савєльєва // Наукові вісті інституту менеджменту та економіки «Галицька академія». – Івано-Франківськ, 2007. – № 1(11). – С. 104 – 107.