

УДК 519.1:004.42

**ПРОГРАММНЫЕ СРЕДСТВА
ДЛЯ ПОСТРОЕНИЯ И ИССЛЕДОВАНИЯ МОДЕЛЕЙ
СТРУКТУРНОЙ СЛОЖНОСТИ ОРГРАФОВ**

А.А. Незнанов (*aneznanov@hse.ru*)
Государственный университет –
Высшая Школа Экономики, Москва

Ю.В. Старичкова (*starichkovayv@mpei.ru*)
Московский Энергетический Институт
(Технический Университет), Москва

Рассмотрены оригинальные программные средства, реализующие построение и анализ системы моделей структурной сложности орграфов. Данные средства реализованы в виде подсистемы АСНИ «*Graph Model Workshop*» и нашли применение при исследовании сложности графовых моделей различных систем.

Введение

Исследуется сложность в классе ориентированных графов (орграфов). Актуальность исследования сложности в классе орграфов обусловлена тем, что данный класс графов имеет большое научное значение и широкое практическое применение. В области представления знаний в виде орграфа представлена внутренняя структура ситуации как совокупности взаимосвязанных признаков, влияющих друг на друга. Методология когнитивного моделирования предназначена для анализа и принятия решений в плохо определенных ситуациях. Основана на моделировании субъективных представлений экспертов о ситуации и включает: методологию структуризации ситуации: модель представления знаний эксперта в виде орграфа. В области интеллектуального анализа данных (*Data Mining*) [??], орграфы отражают свойства информации, и ее взаимосвязь. Распознавание образов – модель предметной области представляется в виде семантической сети (орграфа). Структурная лингвистика - графические модели были для решения задачи поиска и обработки естественного языка,

В докладе рассматривается реализация методов построения и анализа индексов и вектор-индексов структурной спектральной сложности орграфов, а также моделей сложности из класса b -моделей.

1. Индексы структурной сложности орграфов в базисах ориентированных цепных фрагментов

Рассмотрим задачу нахождения индекса, вектор-индекса СС и полного структурного спектра (ПСС) орграфа G в базисе произвольных фрагментов. Вектор-индекс, индекс структурной сложности и ПСС орграфа G в базисе произвольных фрагментов [Кохов, 2002]:

$$VISS(G/\bar{B}) = (\bar{f}_0 \cdot w_0 \cdot ISSC((\bar{F}_0)), \bar{f}_1 \cdot w_1 \cdot ISSC((\bar{F}_1)), \dots, \bar{f}_m \cdot w_m \cdot ISSC((\bar{F}_m))),$$

$$ISSC(G/\bar{B}) = \sum_{i=1}^m \bar{f}_m \cdot w_m \cdot ISSC((\bar{F}_m)),$$

$$WF(G/\bar{B}) = (w_0 \bar{F}_0, w_1 \bar{F}_1, \dots, w_m \bar{F}_m),$$



где G – орграф, $\bar{F}_0 - \bar{F}_m$ – элементы базиса, w_i – количество канонических изоморфных вложений фрагментов \bar{F}_i в орграф G , относительно которого характеризуется сложность графа. В качестве параметров построения индексов, вектор-индексов, ПСС необходимо задать сложность хотя бы некоторых базовых фрагментов минимального размера, сложность фрагментов большей длины может вычисляться рекурсивно.

Для разработки алгоритмов и дальнейших исследований в качестве мер сложности выбраны вектор-индексы, индексы СС и ПСС в базисе простых путей ($ISSC(G/P)$), полупутей ($ISSC(G/PP)$), контуров ($ISSC(G/C)$), полуконтуров ($ISSC(G/CC)$) и ориентированных цепных фрагментов (ОЦФ) ($ISSC(G_1/OCF)$).

ОЦФ – связный орграф, состоящий либо из одной вершины (ОЦФ длины 0, наименьший элемент базиса), либо из большего числа вершин, причем две из них имеют степень 1, а остальные – 2 (длина ОЦФ равна числу дуг в нём). Отметим, что полустепени исхода и захода могут быть любыми. Использование ОЦФ в качестве элементов базиса позволяет увеличить дискриминирующую способность индексов СС при сохранении вычислительной сложности алгоритмов построения индексов.

В таблице 1 приведен пример базовых ОЦФ и их значений сложности.

Таблица 1. Минимальные ОЦФ и их стандартная сложность

Диаграммы минимальных ОЦФ	Стандартные значения сложности ОЦФ
	$ISSC(F_0/OCF) = 1$
	$ISSC(F_1/OCF) = 3$

	$ISSC(F_2/OCF) = 9$
	$ISSC(F_3/OCF) = 10$
	$ISSC(F_4/OCF) = 11$

Таблица 2 содержит пример вычисления значений индексов СС в базисах путей, полупутей, контуров, полуконтуров и ОЦФ для орграфов G_1 и G_2 .

Таблица 2. Примеры вычисления $ISSC$ в различных базисах

Орграф G_1	Орграф G_2	Значения индексов
		$ISSC(G_1/P) = 86, ISSC(G_2/P) = 86$
		$ISSC(G_1/PP) = 242, ISSC(G_2/PP) = 242$
		$ISSC(G_1/C) = 133, ISSC(G_2/C) = 16$
		$ISSC(G_1/CC) = 234, ISSC(G_2/CC) = 234$
		$ISSC(G_1/OCF) = 129, ISSC(G_2/OCF) = 128$

Отметим, что хотя путь является ОЦФ, его отдельное рассмотрение полезно как для сравнительного анализа, так и для выделения самого простого алгоритма построения индекса СС.

Корректность работы алгоритмов подтверждена результатами вычислительных экспериментов на различных семействах ориентированных орграфов.

Были исследованы все орграфы до 6 вершин включительно (1540421), и несколько более узких классов (например, 20278544 бесконтурных и 4664216 планарных бесконтурных орграфов до 8 вершин) и несколько представительных семейств с числом вершин до 1000 (более 10000 орграфов). В качестве базисов для исследования были выбраны базисы связанных путей, полупутей, контуров, полуконтуров и ОЦФ всех длин.

Проведена классификация исследованных классов и семейств орграфов с целью разбиения множеств орграфов с одинаковым числом вершин на классы эквивалентности, на основе значения индексов СС в различных базисах. На рис. 1. и рис. 2. приведены примеры графиков чувствительности (отношения числа классов к числу орграфов) для индексов СС орграфов и планарных орграфов в базисах путей, полупутей, контуров, полуконтуров, ОЦФ. Наиболее точную классификацию удалось получить, используя значение индексов СС в базисе ОЦФ, при длине элементов базиса равной числу вершин в орграфе.

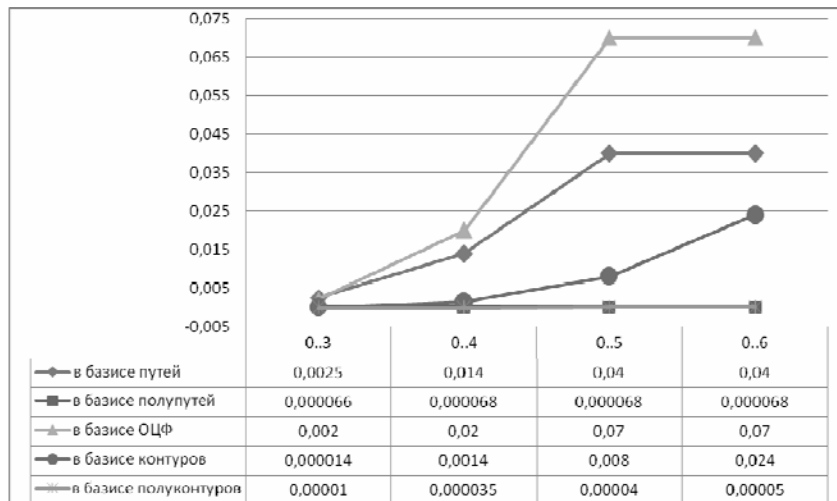


Рис. 1. График чувствительности различения индексов сложности для орграфов на 6 вершинах

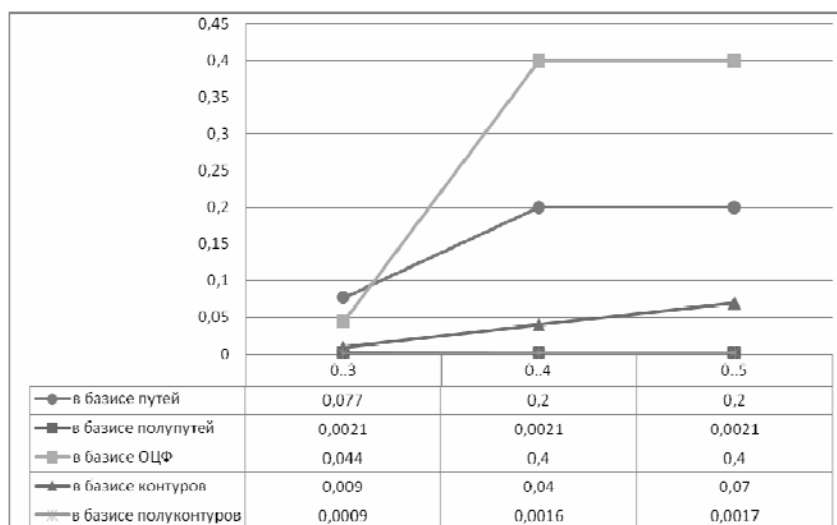


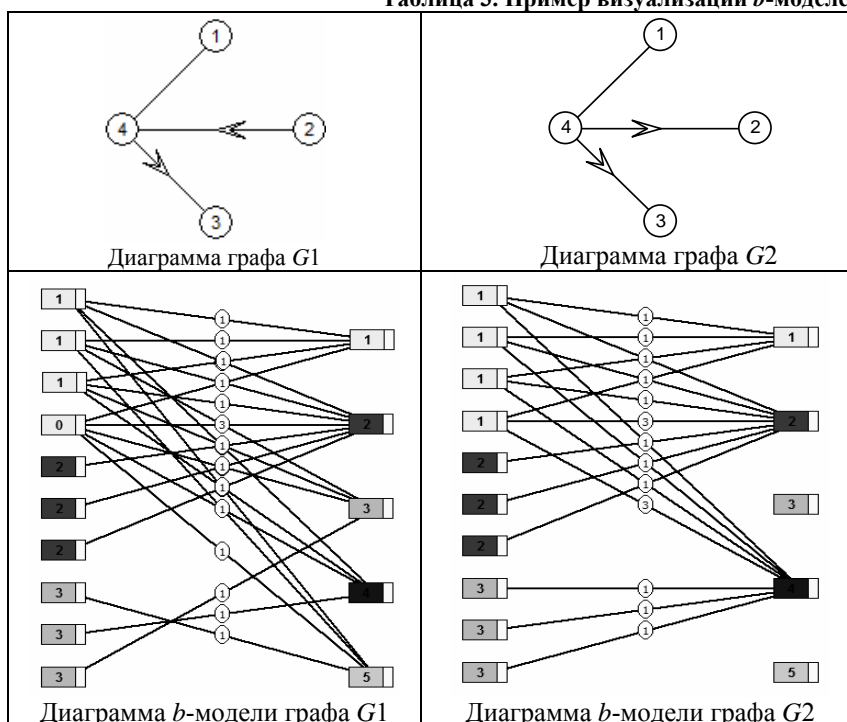
Рис. 2. График чувствительности различения индексов сложности для планарных орграфов на 7 вершинах

2. Структурные модели сложности орграфов

Кроме моделей СС в виде индексов и вектор-индексов, реализованы и более сложные структурные модели сложности – *b*-модели в базисе ОЦФ.

Данный класс моделей СС впервые предложен Коховым В.А. [Кохов, 2002]. *b*-модель представляет собой двудольный граф со структурными весами на вершинах и рёбрах [Незнанов, 2005]. Левая доля – элементы базиса (то есть типы фрагментов орграфа); правая доля – помеченные фрагменты; ребро проводится в случае, если помеченный фрагмент, задаваемый весом вершины левой доли вкладывается во фрагмент, тип которого совпадает с типом, задаваемым весом вершины правой доли. Весы рёбер отражают либо число вложений, либо относительные вклады помеченных фрагментов в общую сложность орграфа. В сравнении с индексами и вектор-индексами СС *b*-модели имеют намного большую дискриминирующую способность, позволяют решать новые классы задач, но требуют больше памяти и имеют большую сложность построения. В таблице 3 приведен пример графов с эквивалентными значениями индекса СС $ISSC(G_1/OCF) = ISSC(G_2/OCF) = 43$, но имеющие различные *b*-модели. Прорисовка диаграмм выполнена идентичным методом для визуального сравнения.

Таблица 3. Пример визуализации *b*-моделей



3. Программная реализация расширения построения индексов и вектор-индексов сложности

Программный комплекс «Сложность орграфов в ориентированных базисах» (далее – *DCDB*) предназначен для вычисления ПСС, вектор-индексов, индексов СС в базисе путей, полупутей, контуров, полуконтуров и ОЦФ орграфов с особой обработкой некоторых их подклассов (например, планарных орграфов). Он реализован в виде набора расширений АСНИ «*Graph Model Workshop*» (*GMW*).

Комплекс создан в среде *Borland Developer Studio 2007* на языке *Delphi*. Объём авторского исходного кода *DCDB* – более 100 КБ, число строк исходного кода основных алгоритмов – 1288, всего компилируемых строк – 3170, объём машинного кода – 1391 КБ. Построенные модели структурной сложности хранятся в виде набора таблиц в базе данных результатов экспериментов *GMW*.

Параметризация построения индексов и вектор-индексов (рис. 3):

- тип базиса для вычисления индекса, вектор-индекса СС;
- длина максимального элемента базиса (для конструктивно перечисляемых базисов) или набор ОЦФ (для выбора базиса, задаваемого пользователем поэлементно);
- значения сложности минимальных элементов базиса (для конструктивно перечисляемых базисов с автоматическим расчётом сложности остальных элементов) или всех элементов.

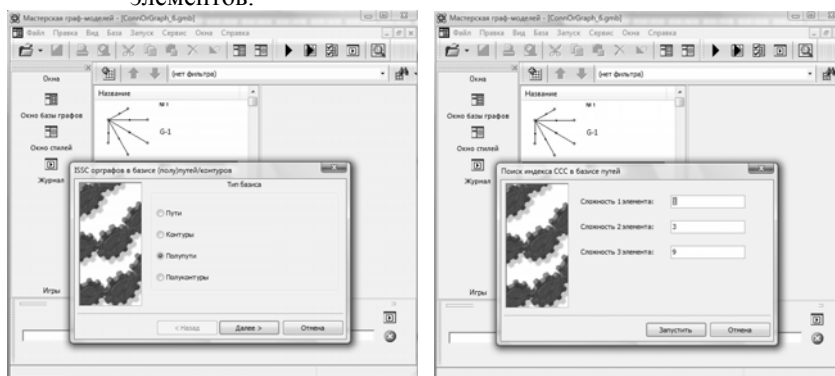


Рис.3. Интерфейс расширения АСНИ «*Graph Model Workshop*»

Технологические ограничения подсистемы (обработка орграфов с числом вершин до 32500, размер фрагмента до 255 вершин) несущественны на фоне высокой временной и емкостной сложности используемых алгоритмов.

Исследована сложность орграфов с использованием индексов, вектор-индексов и *b*-моделей. Расширена функциональность АСНИ «*Graph Model Workshop*» программным комплексом «*DCDB*». Получены оценки чувствительности индексов СС, вектор-индексов, ПСС и *b*-моделей в базисах ОЦФ для планарных, бесконтурных, планарных бесконтурных и семейств орграфов.

Список литературы

- [**Кохов, 2002**] Кохов В.А. Концептуальные и математические модели сложности графов. // М. : Издательство МЭИ (ТУ), 2002.
- [**Незнанов, 2005**] Незнанов А.А. Методы и программные средства для различения расположения фрагментов графовых моделей систем : дис. кан. тех. наук., М. : МЭИ (ТУ), 2005.
- [**Frey, 1998**] В. J. Frey, N. Lawrence and C. M. Bishop. Markovian interence in belief networks. Presented at Learning, Machines that Learn or what some people call Snowbird, Snowbird, Utah, April, 1998.
- [**Heckerman, 1997**] Heckerman D. "Bayesian Networks for Data Mining". Data Mining and Knowledge Discovery. - 1997. - № 1. - P. 79-119.