

**ЭФФЕКТИВНОСТЬ ПРИМЕНЕНИЯ АЛГОРИТМОВ ПОИСКА В ГЛУБИНУ И ПОИСКА В  
ШИРИНУ ДЛЯ НАХОЖДЕНИЯ h-КРАТНЫХ ПОКРЫТИЙ МАТРИЦЫ**

А.В. Ямшанов<sup>1</sup>

Научный руководитель: профессор, д.т.н. А.Е. Янковская<sup>1,2,3,4</sup>

<sup>1</sup>Томский государственный университет систем управления и радиоэлектроники,

Россия, г. Томск, пр. Ленина, 40, 634034

<sup>2</sup>Томский государственный архитектурно-строительный университет,

Россия, г. Томск, ул. Соляная, 2, 634003

<sup>3</sup>Национальный исследовательский Томский государственный университет,

Россия, г. Томск, пр. Ленина, 36, 634050

<sup>4</sup>Национальный исследовательский Томский политехнический университет,

Россия, г. Томск, пр. Ленина, 30, 634050

E-mail: [yav@keva.tusur.ru](mailto:yav@keva.tusur.ru)

**EFFICIENCY OF DEPTH-FIRST AND BREADTH-FIRST APPROACHES  
FOR SOLVING OF MATRIX COVERAGE PROBLEM**

A.V. Yamshanov<sup>1</sup>

Scientific Supervisor: Prof., Dr. A.E. Yankovskaya<sup>1,2,3,4</sup>

<sup>1</sup>Tomsk State University of Control Systems and Radioelectronics, Russia, Tomsk, Lenin str., 40, 634034

<sup>2</sup>Tomsk State University of Architecture and Building, Russia, Tomsk, Solyanaya sq., 2, 634003

<sup>3</sup>National Research Tomsk State University, Russia, Tomsk, Lenin str., 36, 634050

<sup>4</sup>National Research Tomsk Polytechnic University, Russia, Tomsk, Lenin str., 30, 634050

E-mail: [yav@keva.tusur.ru](mailto:yav@keva.tusur.ru)

***Abstract.** Depth-first and breadth-first approaches for solving of matrix coverage problem are briefly presented. Test environment and test approach are described. The comparison of both algorithms performance are given and discussed. Plans for future investigations are proposed.*

Одной из возникающих при применении тестовых методов распознавания, предложенных А.Е. Янковской [1], является задача поиска столбцовых покрытий матрицы. При этом если задача поиска однократных покрытий является достаточно хорошо изученной проблемой [2-4], то количество работ, связанных с построением h-кратных (многократных [5]) покрытий, крайне ограничено [6].

Поиск в глубину является основным подходом к нахождению матричных покрытий, описанным в большинстве публикаций. Подход заключается в обходе дерева поиска с одновременным его построением и поиском всех кратчайших неповторяющихся путей, ведущих от корня дерева поиска к листьям. Большинство описанных алгоритмов являются "жадными", то есть для каждой вершины дерева поиска в первую очередь выбирается столбец, имеющий наибольший потенциальный вклад на данном этапе построения. Чаще всего для его оценки используется количество ненулевых значений в рассматриваемом столбце. Поиск в глубину хорошо изучен, но область его применения в основном ограничена поиском однократных покрытий и неизвестны критерии его применимости для поиска h-

кратных покрытий. Преимуществом поиска в глубину является быстрое получение первого решения (возможно не являющегося минимальным покрытием); а также множество разработанных эвристик, позволяющих отсечь большое количество ветвей дерева поиска и тем самым существенно сократить временные затраты. Недостатком является неочевидное распараллеливание, например, первоначально выбранный нами подход к распараллеливанию алгоритма оказался крайне неэффективным из-за частой обработки путей, которые могли быть отсечены; а также частое попадание алгоритма в локальные минимумы при поиске  $h$ -кратных покрытий, что негативно сказывается на его эффективности.

Поиск в ширину является менее изученным. В нашей реализации используется специальный генератор потенциальных покрытий, последовательно возвращающего все возможные комбинации битовых последовательностей для возрастающего количества признаков. Затем выполняется проверка полученных масок на покрытие ими всех строк БМИ для заданного  $h$ . Преимуществом является простота распараллеливания с крайне эффективной программной реализацией, в том числе с применением технологии GPGPU. Количество перераспределений памяти минимально, что существенно повышает производительность программной реализации, особенно при использовании параллельных вычислений. Основным недостатком является его недостаточная исследованность, поскольку при поиске однократных покрытий алгоритм является менее эффективным. Кроме того, первое покрытие может быть получено в самом конце работы программы.

Для исследования эффективности применения обоих подходов был собран следующий стенд: четырехядерный процессор Intel(R) Core(TM) i7-3770@3.40GHz с Hyper-Threading, 16GiB оперативной памяти DDR3, операционная система Ubuntu Linux 4.8.0, компилятор GCC 4.8.5 с флагами компиляции "-std=c++11 -pthread -O2". Исследуемой характеристикой является время выполнения программы, необходимое для нахождения заданного количества покрытий. Для исключения влияния тестовой среды (шума) тестирование производится сразу после загрузки операционной системы в текстовом режиме с минимально возможным набором запущенных фоновых приложений. Тестовые запуски программы, выполняющиеся менее минуты, запускаются несколько раз, а в качестве результата берется среднее время выполнения программы. Для валидации результатов тестовые данные включают в себя исходные данные и эталонные параметры (распределение покрытий по количеству признаков) заданного количества безызбыточных покрытий, существующих для исходной матрицы. Для всех наборов исходных данных проверяется, что каждая строка из набора строк на выходе программы является покрытием для исходной матрицы и параметры найденных покрытий соответствуют эталонным параметрам. В качестве исходных данных используются тестовые данные, соответствующие реальным данным из области кардиологии. При этом сохранено количество признаков и выявленные закономерности [1], увеличено количество объектов в обучающей выборке и добавлен небольшой шум (менее 1%). Результаты тестирования представлены в табл. 1 и проиллюстрированы на рис. 1.

Полученные в ходе тестирования результаты являются достаточно контр интуитивными, поскольку применение поиска в ширину показывает на 2-3 порядка лучшие результаты. Во-первых, это обусловлено различиями в задачах поиска однократных и  $h$ -кратных покрытий: примененные при поиске в глубину эвристики наследованы от поиска однократных покрытий и не адекватны для  $h$ -кратных покрытий; специальные эвристики для поиска  $h$ -кратных покрытий еще не разработаны. Во-вторых,

поиск в ширину позволяет выполнить более эффективную программную реализацию, что также отражается на результате.

Таблица 1

Время выполнения программы в зависимости от количества объектов

Количество объектов в обучающей выборке	128	256	384	512
Алгоритм	Время выполнения программы, секунды			
Поиск в глубину	17,188	57,177	-	156,077
Параллельный поиск в глубину	3,504	10,401	36,504	21,149
Поиск в ширину	0,036	0,063	0,224	0,215
Параллельный поиск в ширину	0,014	0,026	0,065	0,058

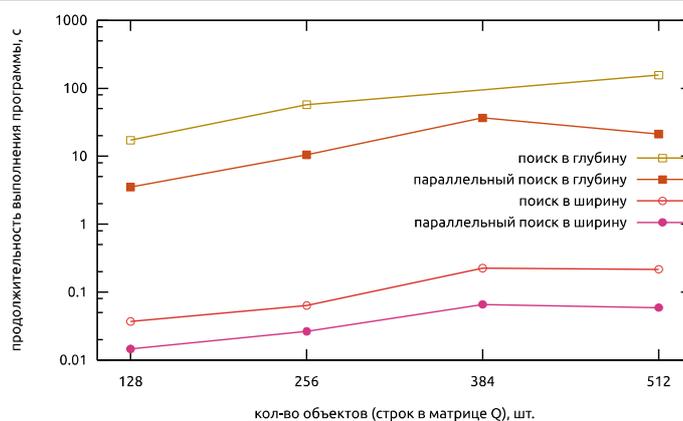


Рис. 1. Время выполнения программы в зависимости от количества объектов

Дальнейшие исследования направлены на программную реализацию поиска в ширину с применением технологии GPGPU и исследование её производительности. Кроме того, целесообразным видится применение генетических алгоритмов.

Работа выполнена при финансовой поддержке гранта РФФИ (проект № 16-07-00859а).

#### СПИСОК ЛИТЕРАТУРЫ

1. Янковская А.Е. Логические тесты и средства когнитивной графики. LAP LAMBERT Academic Publishing, 2011, – с.92
2. Янковская А.Е., Китлер С.В. Принятие решений на основе параллельных алгоритмов тестового распознавания образов // Искусственный интеллект. Украина, Донецк: ИПШ "Наука і освіта". — 2010. — № 3. — С. 151-159.
3. Кудрявцев В.Б., Андреев А.Е. Тестовое распознавание // Фундаментальная и прикладная математика 15.4 (2009): 67-99.
4. Дюкова Е.В., Прокофьев П.А. Об асимптотически оптимальном перечислении неприводимых покрытий булевой матрицы // ПДМ. – 2014. – № 1. –С. 96–105.
5. Агибалов Г.П. Нахождение оптимальных многократных покрытий множеств // Труды / Сиб. физ. - техн. ин-т при Том. ун-те. Томск. – 1966. – Вып. 48. – С. 79–86.
6. Янковская А.Е. Принятие решений, устойчивых к ошибкам измерения значений признаков в интеллектуальных системах // Искусственный интеллект. Интеллектуальные системы (ИИ-2009). Материалы X Международной научно-технической конференции. — Таганрог: Изд-во ТТИ ЮФУ, 2009. — С. 127-130.