Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ
Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра дискретной математики
и информационных технологий

Покрытия множеств

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 421 группы

направления 09.03.01 «Информатика и вычислительная техника»

факультета компьютерных наук и информационных технологий

Потловой Анастасии Юрьевны

Научный руководитель
д. ф.-м.н., профессор ______________ В.А.Молчанов
подпись, дата

Зав. кафедрой
к. ф.-м.н., доцент ______________ Л.Б. Тяпаев
подпись, дата

Саратов 2017

ВВЕДЕНИЕ

Задача о покрытии множества (ЗПМ) системой его подмножеств [1–3] является математической моделью для многих важный приложений, таких как составление расписания [4], планирование сервиса, логического анализа данных, упрощение булевских выражений [3] и т. д. ЗПМ относится к классу NP-полных задач [5], поэтому построение алгоритма нахождения оптимального или приближенного решения является весьма сложной задачей. Однако структура представления реальной задачи предоставляет дополнительную информацию, позволяющую решать ЗПМ довольно больших размерностей (несколько сотен строк и несколько миллионов столбцов) с результатом, отличающимся от оптимального решения примерно на 1 %, за приемлемой времени счета [3].

Целью данной работы является изучение методов решения задачи о покрытии множества.

В настоящей работе исследуются основные свойства задачи о покрытии множеств и рассматриваются основные виды алгоритмов решения поставленной задачи, такие как:

- жадный алгоритм;
- генетический алгоритм;
- точный алгоритм.

Целью практической части данной работы является проведение вычислительного эксперимента, а именно, нахождение решения задачи о покрытии множеств путем программной реализации точного алгоритма на языке C# в среде VISUAL STUDIO 2015 COMMUNITY.

1. Постановка задачи о покрытии множества

Комбинаторная постановка задачи о покрытии множества состоит в следующем. Пусть даны множество $M = \{1, \ldots, m\}$ и набор его подмножеств $M_1, \ldots, M_n$ таких, что $\bigcup_{j=1}^{n} M_j = M$. Совокупность подмножеств M_j, $j \in J \subseteq \{1, \ldots, n\}$, называется *покрытием* множества M, если $\bigcup_{j \in J} M_j = M$. Если каждому M_j приписан вес $c_j \geq 0$, то требуется найти покрытие минимального суммарного веса. Задача называется невзвешенной если все подмножества M_j имеют единичные веса.

Задача о покрытии множества возникает, когда мы стремимся экономно приобрести товары, разложенные по наборам. Мы хотим получить, как минимум, по одному товару каждого вида, покупая при этом как можно меньшее количество наборов. Задача поиска хоть какого-нибудь покрытия множества не представляет сложности, т.к. можно купить все предлагаемые наборы товаров. Но поиск наименьшего покрытия множества позволяет нам минимизировать наши расходы.

2. Методы решения поставленной задачи

Дано множество U из n элементов, и набор подмножеств U, $S = \{S_i, \dots, S_k\}$. Каждому подмножеству S_i сопоставлена некоторая неотрицательная стоимость $c\colon S \rightarrow Q^+$, где Q^+ – множество положительных рациональных, $S' \subseteq S$ является покрытием, если любой элемент из U принадлежит хотя бы одному элементу из S' [1].

Задача о покрытии множествами заключается в нахождении набора подмножеств, покрывающего все множество U и имеющего минимально возможный вес (в случае взвешенной задачи) или минимально возможное число подмножеств (в случае невзвешенной задачи).

Можно представить задачу в матричном виде [2]. Пусть $A = (a_{ij})$ — произвольная матрица размера $m\ x\ n$ с элементами $a_{ij} \in \{0,1\}$ без нулевых строк и столбцов. Будем говорить, что в A строка i покрывается столбцом j, если $a_{ij} = 1$. Подмножество столбцов называется покрытием, если в совокупности они покрывают все строки матрицыA. Пусть каждому столбцу поставлено в соответствие положительное число c_j, называемое весом столбца. Требуется найти покрытие минимального суммарного веса. Вводя переменные x_j, равные 1, если столбец j входит в искомое покрытие, и равные 0 в противном случае, приходим к следующей формулировке задачи о покрытии:

$$\sum_{j=1}^{n} c_j x_j \rightarrow \min,$$

при ограничениях

$$\sum_{j=1}^{n} a_{ij} x_j \geq 1, \mathrm{i} = 1, \dots, m,\ x_j \in \{0,1\}, j = 1, \dots, n.$$

2.1 Обзор алгоритмов решения ЗПМ

Для решения задачи покрытия множеств разработано множество алгоритмов, которые можно разделить на следующие классы: приближенные алгоритмы с априорной оценкой, эвристические алгоритмы, точные алгоритмы [7].

Одним из первых приближенных алгоритмов является жадный алгоритм. Для задачи о покрытии с произвольными весами В. Хватал предложил модификацию жадного алгоритма [9].

Поскольку эвристики носят вероятностный характер, их сложность невозможно оценить априорно. К таким методам можно отнести методы лагранжевой релаксации, генетические алгоритмы, поиск с запретами, алгоритмы муравьиной колонии, нейронные сети.

Примером точных алгоритмов служит метод ветвей и границ.

2.2 Жадный алгоритм

В 1979 году Хватал [8] предложил жадный алгоритм для задачи о покрытии множествами. На каждом новом шаге алгоритма происходит выбор самого эффективного множества, т.е. такого множества, которое покрывает наибольшее количество элементов на данном этапе, а если таких несколько, то выбирается то, которое обладает наименьшей стоимостью.

После чего удаляем покрытые элементы и продолжаем до тех пор, пока не будут покрыты все элементы.

Выбранные в результате работы алгоритма множества являются результатом работы жадного алгоритма для задачи о покрытии множества.

2.3 Генетический алгоритм.

В 1975 году Джон Холланд предложил генетический алгоритм (ГА), основанный на принципах естественного отбора и наследования [10-13]. Рассмотрим применение генетического алгоритма для решения задачи покрытия множеств. На рисунке 2 представлена общая схема работы ГА.

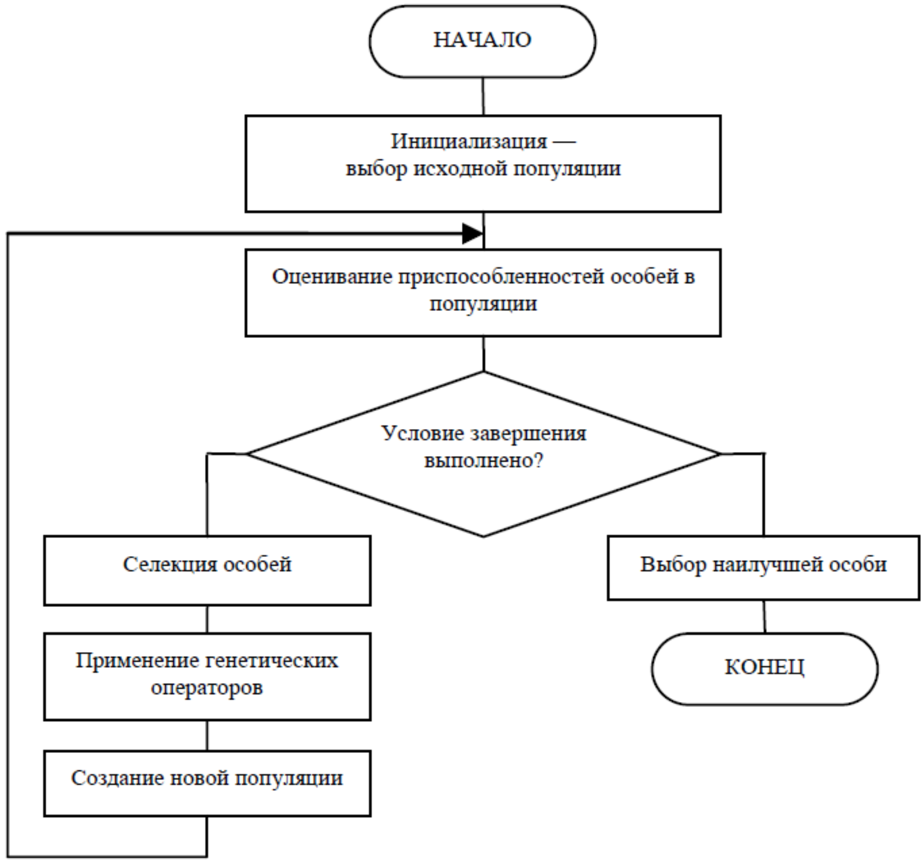

Рисунок 2 – Общая схема ГА

2.4 Точный алгоритм

Для решения общей задачи о покрытии предложено большое число точных алгоритмов, основанных на методах ветвей и границ, отсечения, перебора X-классов и др.

Верхние оценки временной сложности точных методов сопоставимы со сложностью переборных алгоритмов [14-16]. В связи с этим при решении практических задач большую актуальность приобретают эвристические способы сокращения перебора. Значительное внимание уделяется разработке гибридных алгоритмов, в которых переборные схемы комбинируются с отсечениями и различными эвристиками для получения верхних и нижних оценок оптимума целевой функции на вспомогательных подзадачах.

В данной работе рассмотрены последовательный и параллельный алгоритмы, основанные на методе ветвей и границ для задач о покрытии, предложенном в [7].

3. Практическая часть

В ходе выполнения данной работы был разработан программный продукт, позволяющий решить задачу о покрытии множеств методом «Ветвей и границ». В дальнейшем данная программа может быть использована для решения конкретных прикладных задач.

Программа была реализована на языке C# в среде VISUAL STUDIO 2015 COMMUNITY. Код программы приведен в приложении А.

Разработанная программа включает в себя три этапа:

1. Упрощение задачи;
2. Нахождение приближенного решения;
3. Определение оптимального решения;

Приступив к работе с программой, пользователь в первую очередь должен задать размеры матрицы – ввести количество строк и столбцов, Далее необходимо ввести матрицу, для которой планируется найти покрытие. Сделать это можно вводом с клавиатуры, как показано на рисунке 3. Или же загрузив матрицу из файла, выбрав пункт Файл -> Загрузить матрицу из файла, как показано на рисунке 4.

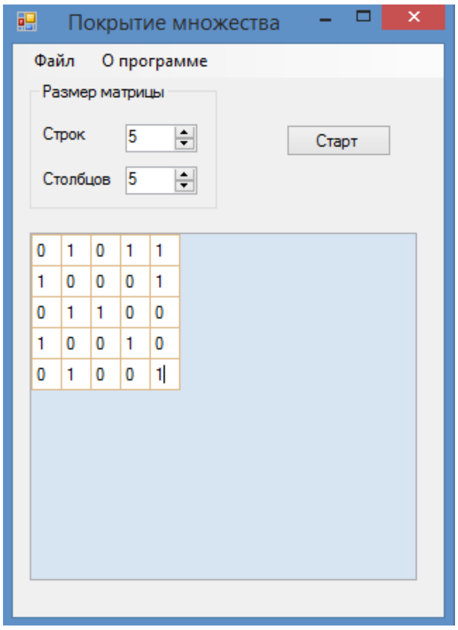

Рисунок 3 – ввод матрицы.

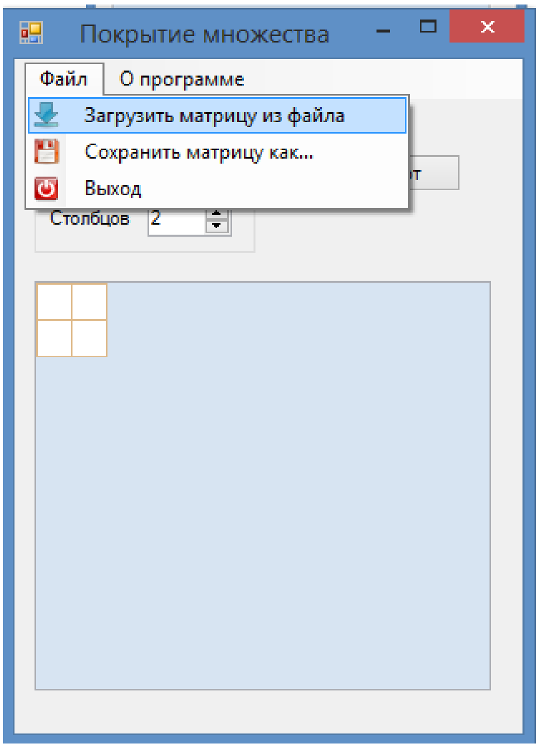

Рисунок 4 – загрузка матрицы из файла.

Далее необходимо нажать на кнопку «Старт», после чего появиться окно, в котором нужно указать стоимость каждого из столбцов матрицы, как показано на рисунке 5.

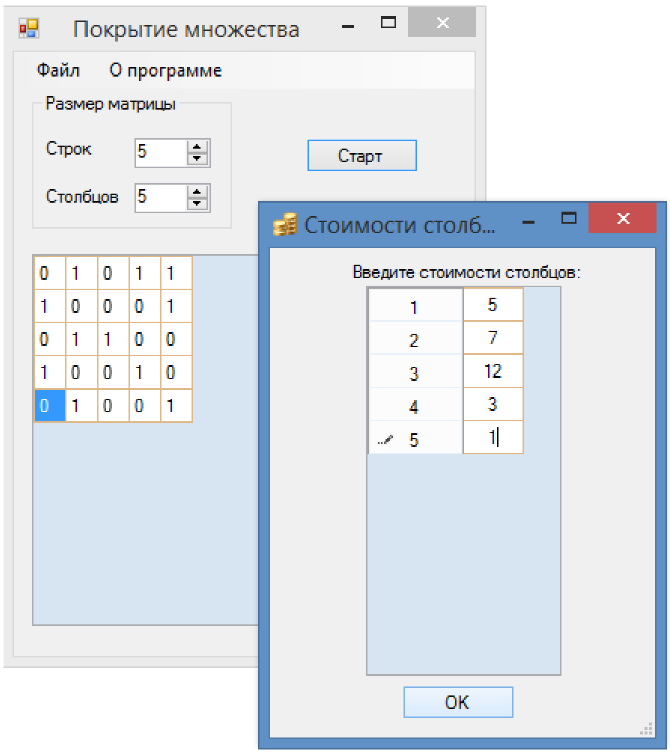

Рисунок 5 – ввод стоимости столбцов матрицы.

Далее, при нажатии на кнопку «ОК» появится окно, в котором будет отображен результат работы программы, как показано на рисунке 6.

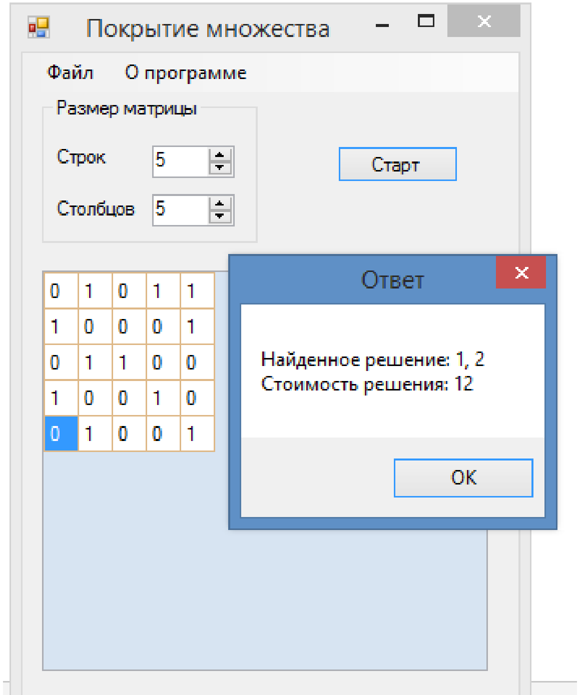

Рисунок 6 – результат работы программы.

ЗАКЛЮЧЕНИЕ

При написании работы изучена специальная литература, включающая в себя статьи и учебники по покрытиям множеств, дискретной математике и теории алгоритмов, описаны теоретические аспекты и раскрыты ключевые понятия исследования.

Разработана компьютерная программа, позволяющая производить вычислительные эксперименты с целью нахождения решения задач о покрытии множеств.

Программа написана на языке C# в среде VISUAL STUDIO 2015 COMMUNITY. Программа оправдывает свое назначение и выполняет операции с различными входными данными, что и являлось целью практической части данной работы.

Таким образом, все поставленные задачи работы в полном объеме решены.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Ceria S. et al. Set Covering Problem // Dell'Amico, F. Maffioli and S. Martello (eds.). Annotated Bibliographies in Combinatorial Optimization. Wiley and Sons, 1997. P. 415–428.

2 Galinier P. and Heztz A. Solution techniques for the large set covering problem // Discrete applied Mathematics. 2007. Vol. 155. Issue 3.

3 Caprara A., Fischetti M., Toth P. Algorithms for the set covering problem // Working Paper, DEIS, University of Bologna. 1998.

4 Capara A. et al. Algorithms for Railway Crew Management // Mathematical Programming. 1997. 79. P. 125–141.

5 Гэри М., Джонсон Д. Вычислительные машины и трудно решаемые задачи: Пер. с англ. 1982.

6 E. Balas and M. Padberg. Set partitioning – a survey. SIAM Review, 18:710-760, 1976

7 Берж К. Теория графов и ее применения. М.: Изд-во иностр. лит., 1962.

8 Герман О. В., Ефремов О. В. Алгоритм решения обобщенной задачи о покрытии / / Экономика и мат. методы. 1998. Т. 34, вып. 4. С. 134-140.

9 Стивен С. Скиена. Алгоритмы. Руководство по разработке. Второе издание. Пер. с англ. 2011. Стр. 631-637

10 Емеличев В. А. Ковалев М. М., Кравцов М. К. Многогранники, графы, оптимизация. М.: Наука, 1981.

11 Еремеев А. В., Заозерская Л. А., Колоколов А. А. Задача о покрытии множества: сложность, алгоритмы, экспериментальные исследования / / Между нар. конф. «Дискретный анализ и исследование операций»: Материалы конф. Новосибирск: Изд-во Ин-та математики, 2000. С. 37-41.

12 Журавлев Ю. И. Теоретико-множественные методы алгебры логики //Проблемы кибернетики. Вып. 8. М.: Физматгиз, 1962. С. 5-44.

13 Агеев А. А. О сложности задач минимизации полиномов от булевых переменных / / Управляемые системы: Сб. науч. тр. Новосибирск: Ин-математики СО АН СССР, 1983. Вып. 23. С. 3-11.

14 Сайко Л. А. Исследование мощности L-накрытий некоторых задач о покрытии / / Дискретная оптимизация и анализ сложных систем: Сб. науч. тр. Новосибирск: ВЦ СО АН СССР, 1989. С. 76-97.
15 Lifschitz V., Pittel B. The worst and the most probable performance of a class of set-covering algorithms / / SIAM J. Cornput. 1983. V. 12, N 2. P. 329-346.
16 Липский В. Комбинаторика для программистов / пер. с пол. В.А. Евстигнеева; под ред. А.П. Ершова. – М.: «Мир», 1988. – С. 1 2 - 108.