УДК 621.396.6

ПАРХОМЕНКО А.А., аспирант (УкрГАЖТ)

Исследование субэкспоненциального алгоритма SAT- задач

В статье были опубликованы результаты экспериментального исследования предложенного субэкспоненциального алгоритма, что позволяет точно решать SAT- задачи большой размерности.
Ключевые слова: *SAT- задачи, булева функция, солвер, метод Монте – Карло.*

Введение

В последние голы уделяется довольно много внимания вопросам разработки систем автоматической трансляции в SAT комбинаторных задач из различных областей. Большинство таких систем предназначены для преобразования в SAT различных вариантов задачи удовлетворения ограничений (Constraint Satisfaction Problem, CSP) и др [1 - 3]. Ряд комбинаторных проблем естественным образом можно рассматривать с позиции обращения эффективно вычислимых дискретных функции. Таковыми, например, являются проблемы криптоанализа. Формальное описание задачи в этом случае выглядит следующим образом [5]. Рассматривается всюду определенная эффективно вычислимая дискретная функция

$$f : \{0,1\}^* \rightarrow \{0,1\}^*, \quad (1)$$

заданная некоторой программой A (f). На самом деле программа A (f) задает семейство функций вида

$$fn : \{0,1\}^* \rightarrow \{0,1\}^*, n \in \mathrm{N}. \quad (2)$$

Требуется, зная $y \in Range$ (fn), найти такой $x \in \{0,1\}$, что fn (x) = y [7]. Возможность применения SAT – подхода к решению этой задачи основана на следующем факте. Мы можем представить процесс преобразования алгоритмов A (f) слова ″x″ в слова ″y″ как последовательность булевой формулы. При этом используется идея Кука по пропозициональному кодирования алгоритмов, а также идея Кинга по символьному исполнению программ. Данная общая схема была реализована в виде программного комплекса Transalg. Комплекс Transalg представляет собой систему автоматической трансляции процедурных описаний дискретных функции в системе булевых уравнений и, в конечном счете, в SAT- задачи [3-5]. В качестве языка описания в Transalg используется проблемно – ориентированный язык ТА, имеющий С- подобный синтаксис. При трансляции ТА- описаний используются стандартные техники теории компиляции. Результатом компиляции ТА-программы является не машинный код, а множество булевых формул, которые можно рассматривать как систему булевых уравнений. От данного множества формул делается переход к SAT- задаче при помощи преобразований Цейтина. Также одной из значимых достижений в исследовании комбинаторных проблем можно считать прогресс в решении систем логических (булевых) уравнений большой размерности. На сегодняшний день удается решать системы, содержащие сотни тысяч булевых переменных и уравнений (булевых ограничений) [8]. К булевым уравнениям эффективно сводятся многочисленные комбинаторные задачи [6]. Процедура перехода от исходной постановки к системе булевых уравнений называется пропозициональным кодированием. Практически значимые задачи, которые имеет смысл сводиться к булевым уравнениям, возникают в таких областях, как синтез и верификация схем в микроэлектронике, исследование безопасности коммуникационных протоколов, обоснование корректности программ, криптоанализ, а также при исследовании свойств динамических дискретно – автоматных моделей [6 - 8].

В настоящее время широкое применение для решения SAT- задач находит применение метод Монте – Карло в параллельных вычислительных системах. Распараллеливание SAT- задачи является результатом выделение в множестве булевых переменных исходной конъюктивной нормальной формы некоторого подмножества, называемого декомпозиционным множеством. Для декомпозиционных множеств можно естественным образом определить ряд параметров, характеризующих ″″качество″ декомпозиции. Для оценки этих параметров предлагается использовать вычислительную схему метода Монте – Карло [1 - 4]. В частности, данный метод применен для поиска декомпозиционного множества с найменьшим прогнозным временем решения исходной задачи. Реализована параллельная MPI- программа, с помощью которой на вычислительном кластере был получен прогноз времени решения задачи логического

криптоанализа шифра Bivium. Успешно осуществлен логический криптоанализ нескольких ослабленных версий шифра Bivium, проведено сравнение реального времени криптоанализа с прогнозами [1]. Использование подходов и алгоритмов искусственного интеллекта (ИИ) позволяет решать многие прикладные задачи, такие, как задачи теории расписаний, задачи проектирования экспертных систем и систем поддержки принятий решений, доказательство теорем, задачи тестирования электронных схем, обработка изображений. Одной из важных задач ИИ является задача удовлетворения ограничений (constraint satisfaction problem). К сожалению, большинство интересных задач ИИ являются NP- трудными и решение их в худшем случае может требовать перебора экспоненциального числа решений. Многие практические задачи содержат огромное число переменных и/или ограничений, что создает сложности при попытки решения этих задач с помощью современных решателей [6 - 8]. Перспективными декомпозиционными подходами, использующими структуру разреженных графов, описывающих задачи ИИ, являются графовые декомпозиционные методы, интерес к которым возрос в последнее время, что обусловлено результатами Arnborg et al., доказавших, что ряд NP- трудных задач, поставленных в монадической логике второго порядка, могут быть решены за полиномиальное время с помощью методов динамического программирования на графах, описывающих структуру задачи, с ограниченной древовидной шириной. К графовым декомпозиционным подходам относится класс локальных элиминационных алгоритмов (ЛЭА) вычисления информации, включающий локальные алгоритмы декомпозиции, алгоритмы несериального динамического программирования (НСДП), алгоритмы сегментной элиминации, метод древовидной декомпозиции. Поскольку задача верификации может быть эффективно сведена к задаче выполнимости (satisfiability problem, SAT), которая решается эвристически специальными программными средствами (SAT- solverami), возникла необходимость в исследовании задачи выполнимости на предмет поиска эвристики, которая была бы эффективнее уже существующих. Эффективность здесь понимается в смысле оценки временной сложности в худшем случае. Для этого было выполнено сведение задачи SAT к комбинаторной задаче на игровом поле. Существует широко распространенный метод такого сведения: вначале по булевой формуле строится булева схема из конечного набора функциональных элементов, после этого данная схема ″рисуется″ на игровом поле рассматриваемой задачи: для каждого функционального элемента строится конфигурация, представляющая его в терминах игры [5 - 7]. Построение такой конфигурации может проводиться как вручную, так и автоматически. Также одной из актуальных областей применения SAT – подхода есть связь с задачами комбинаторной оптимизации. Имеются мощные коммерческие пакеты решения оптимизационных задач из семейства 0-1 – целочисленного линейного программирования (ЦЛП). Соответствующие алгоритмы комбинируют технику ветвей, границ и отсечений с методами решения задач линейного программирования над полем рациональных чисел. Однако данные методы подходят далеко не ко всем задачам комбинаторной оптимизации. Современные решатели ЦЛП- задач, как правило, не работают с нелинейными и невыпуклыми целевыми функциями. Между тем и такие задачи допускают эффективную сводимость к булевым уравнениям и, в конечном счете, к SAT- задачам [6 - 8].

Формализация SAT-задачи и ее решение

Рассмотренную булеву функцию $f(x_1, x_2, .., x_n)$ представляем в конъюнктивной форме записи

$$f(x_1, x_2, .., x_n) = (x_1^{\sigma_{11}} \vee x_2^{\sigma_{12}} \vee ... \vee x_n^{\sigma_{1n}}) \wedge ...$$
$$... \wedge (x_1^{\sigma_{m1}} \vee x_2^{\sigma_{m2}} \vee ... \vee x_n^{\sigma_{mn}}),$$

где

$$x_i^{\sigma} = \begin{cases} x_i, -при\ \sigma = 1 \\ \overline{x_i} - при\ \sigma = 0 \end{cases} \quad (3)$$

Операции $\vee$, $\wedge$ являются булевыми и моделируют простейшие логические высказывания: $\vee$ - "″ИЛИ"; $\wedge$ - "И". Для любого двоичного набора $x = (x_1, x_2, .., x_n)$ функция принимает одно из двух возможных значений: единицу или ноль. Задача "выполнимость" заключается в ответе на вопрос: существует ли набор значений переменных $\tilde{o} = (x_1, x_2, .., x_n)$, обращающий функцию *f* в единицу. В работе предложен алгоритм решения SAT- задачи субэкспоненциальной сложности, но в статье не приведены экспериментальные данные алгоритма, которые характеризовали бы возможности его применения, поэтому целью экспериментального исследования алгоритма является оценка его временной сложности. В процессе эксперимента была снята зависимость числа элементарных операций (математическое ожидание) от числа дизъюнктов от 10 до 40 с шагом 10 при фиксированных значениях n = 4, 6, 12, и зависимость для среднего квадратичного отклонения (СКО) [1 - 5].

При исследовании создавались случайные булевы функции, в которых переменные в дизъюнктах генерировались по равномерному закону распределения с заданным числом переменных в каждом дизъюнкте. В процессе работы программы находились наборы выполнимости заданной функции, а также вычислялось математическое ожидание и среднеквадратическое отклонение количества операций и времени выполнения, затраченное алгоритмом на поиск набора выполнимости булевой функции. На каждую точку в графиках генерировалось не менее 50 булевых функций, и результаты получены с доверительной вероятностью 0,95 они представлены на графиках (Рис. 1. – Рис. 6.).

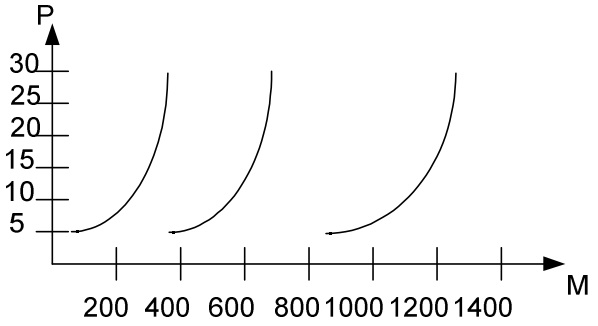

Рис. 1. Зависимость математического ожидания числа операций от количества переменных для квазиэкспоненциального метода:
P– количество переменных; M– математическое ожидание числа операций

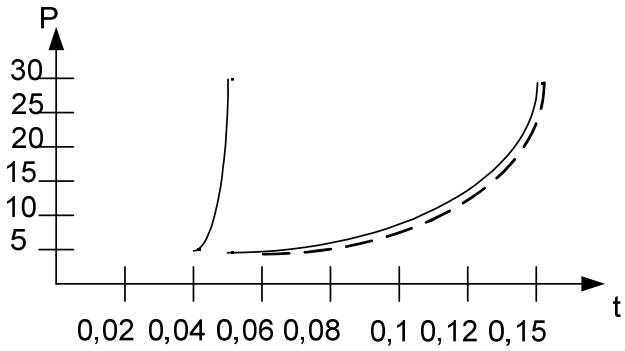

Рис. 2. Зависимость математического ожидания времени выполнения от количества переменных для квазиэкспоненциального метода:
P – количество переменных; t – время выполнения

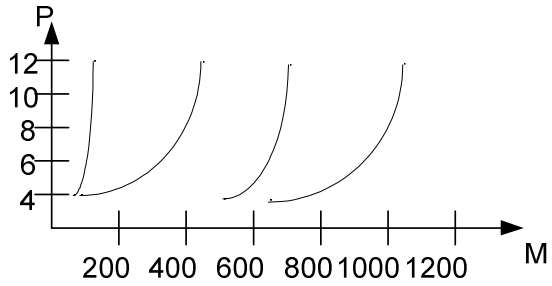

Рис. 3. Зависимость математического ожидания числа операций от количества переменных для квазиэкспоненциального метода:
P – количество переменных; t – время выполнения

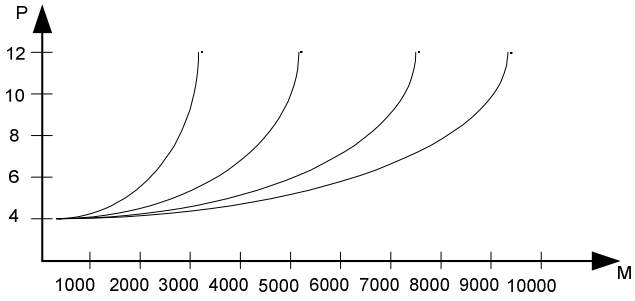

Рис. 4 . Зависимость математического ожидания от числа операций от количества переменных для эвристического метода: P – количество переменных; M- математическое ожидание числа операций

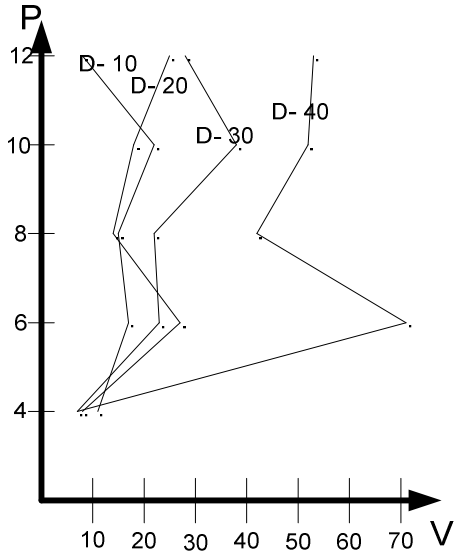

Рис. 5. Зависимость среднеквадратичного отклонения числа операций от числа переменных для квазиэкспоненциального метода: D – количество дизъюнктов; P – количество переменных;
V – среднеквадратичное отклонение числа операций

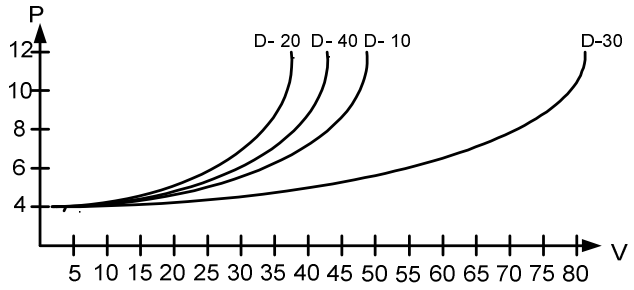

Рис. 6. Зависимость среднеквадратичного отклонения числа операций от количества переменных для эвристического метода: D – количество дизъюнктов; P – количество переменных; V – среднеквадратичное отклонение числа операций

Заключение

В результате экспериментального исследования предложен алгоритм субэкспоненциальной сложности, который позволяет точно решать SAT- задачи достаточно большой размерности и может быть использован в промышленных солверах для решения SAT задач.

Литература

1. Michael R. Garey and David S. Johnson . Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freman, 1979.
2. Листровой С.В. О классе NP и NP- полных задачах.// Электронное моделирование, 2011, т. 33, №1, с 31 – 45.
3. Ахо А., Ульман Д., Хопкрофт Д. Структуры данных и алгоритмы. М.: Вильямс. 2000. 384 с.
4. Заикин О.С., Семенов А.А. Технология крупноблочного параллелизма в SAT-задачах // Проблемы управления. 2008. №1. 43-50.
5. Metropolis N., Ulam S. The Monte Carlo method // J. of the American Statistical Association. 1949. 44, N 247. 335–341.
6. Гришагин В.А., Свистунов А.Н. Параллельное программирование на основе MPI. Изд-во Нижегородского гос. ун-та им. Н.И. Лобачевского, 2005.
7. Een N., Sorensson N. Translating Pseudo-Boolean Constraints into SAT // Journal on Satisfiability, Boolean Modeling and Computation. 2006. Vol. 2. P. 1-25.
8. Посыпки М.А., Заикин О.С., Беспалов Д.В., Семенов А.А. Решение задач криптоанализа поточних шифров в распределенных вычислительных средах // Труды ИСАРАН. 2009. № 46. С. 119-137.

Пархоменко О.О. Дослідження субекспоненціального алгоритму SAT - задач. У статті були опубліковані результати експериментального дослідження запропонованого субекспоненціального алгоритму, що дозволяє точно вирішувати SAT- задачі великий розмірності.

Ключові слова: SAT- задачі, булева функція, солвер, метод Монте – Карло.

Parkhomenko O.O. Investigation subexponential algorithm SAT- tasks. In the article has been published results of the experimental research that considered a subexponential algorithm that allow us to resolve multidimensional SAT tasks with high precision.

Key words: SAT- tasks, Boolean function, solver, Monte - Carlo method.

Рецензент д.т.н., профессор Мойсеенко В.И. (УкрГАЖТ)

Поступила 23.09.2014г.