ОПРЕДЕЛЕНИЕ ВОЗМОЖНЫХ ЗНАЧЕНИЙ НЕЛИНЕЙНОСТИ БУЛЕВЫХ ФУНКЦИЙ МНОГИХ ПЕРЕМЕННЫХ С ИСПОЛЬЗОВАНИЕМ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ

Д.А. Никитин, К.В. Дьяконов

В данной работе рассматривается возможность решения задачи определения возможных значений нелинейности булевых функций многих переменных с использованием вычислений на кластере. Результаты могут быть полезны для алгебры и криптографии, в частности, могут применяться для выбора узлов замен блочных шифров с заданной величиной нелинейности.

Для достижения устойчивости шифра к линейному и дифференциальному криптоанализу необходимо, чтобы все преобразования бит в ходе шифрования были, во-первых, как можно более нелинейными (то есть далеко отстоящими в смысле Хэмминга от множества аффинных булевых функций), во-вторых, обладали лавинным эффектом. Количественными характеристиками, выражающими устойчивость шифрования к этим методам криптоанализа, являются нелинейность и динамическое расстояние, определенные для процедуры преобразования бит.

При выборе узлов замен полезно знать максимальное значение нелинейности, достижимое при замене группы бит заданной длины. Пока не существует аналитического способа определения возможных значений нелинейности булевых функций от определённого числа переменных N, либо нелинейности конкретной булевой функции. Поэтому приходится вычислять нелинейность, основываясь на его определении, то есть практически перебором: вычисляя расстояние Хэмминга каждой функции от всех линейных функций того же числа переменных. При выборе узла замены, таким образом, приходится осуществлять подобный перебор для всех нелинейных булевых функций N переменных.

Вычислительная сложность данной задачи не позволяет решать её за приемлемое время с использованием персонального компьютера. В частности общее количество булевых функций от N переменных равно 2^{2^N}, а размер одной функции (хранимой в виде её таблицы истинности) – 2^N бит, то есть 2^{N-3} байт. Например, для N=32 размер одной функции – 512 Мбайт. Решение задачи в аналитическом виде также является чрезвычайно трудной проблемой. В настоящее время известны возможные значения нелинейности для функций от менее десяти переменных. В то время как для криптографии интересны функции от 32-х булевых переменных и более. Ввиду большой практической и теоретической значимости, есть потребность в таблице значений нелинейности для различных значений количества переменных.

Для нахождения значений нелинейности наиболее разумным предполагается использовать вычисления на кластере, так как характер описанного алгоритма переборный – и он может быть подвергнут распараллеливанию. Так как даже мощности суперкомпьютера может быть недостаточно при больших N, алгоритм следует оптимизировать. Например, таблицу линейных функций можно сократить вдвое, если учесть, что расстояние Хэмминга до некоторой функции f_2, инверсной к функции f_1, равно $(2^N - \text{расстояние до } f_1)$. Также, в формуле для линейных функций можно отбросить и не учитывать свободный член, так как если он равен 0, то функция не меняется, а если он равен 1, то мы получаем инверсию предыдущей функции. Кроме высокоуровневой оптимизации (на уровне алгоритма) имеет смысл попытаться прибегнуть к оптимизации на среднем и низком уровнях.

Среди ожидаемых результатов работы: таблица значений нелинейности булевых функций для различных значений количества переменных, таблицы истинности функций с большими значениями нелинейности, способы оптимизации вычисления нелинейности для исключения прямого перебора, гипотеза о возможных значениях нелинейности для больших N на основе набранной статистики.