умножая данные из таблиц на $31{,}536 \cdot 10^{12}$ для MIPS years, на $31{,}536 \cdot 10^{15}$ для 1 ГГц CPU years ($\approx 10^9$ — средняя производительность одного процессора Pentium) и на 10^{12} для максимальной полученной производительности по LINPACK в TFlop/s.

Результаты интерполяции и экстраполяции данных представлены на рис. 1 и 2. Видно, что рост производительности суперкомпьютеров отстаёт от роста трудоёмкости алгоритма NSF при увеличении размера его входных данных.

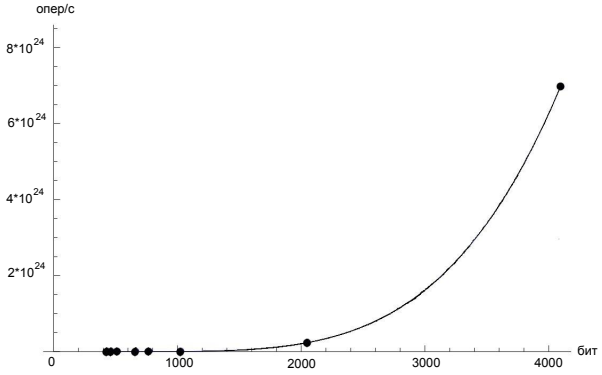

Рис. 1. Рост трудоёмкости алгоритма NSF в зависимости от размера модуля

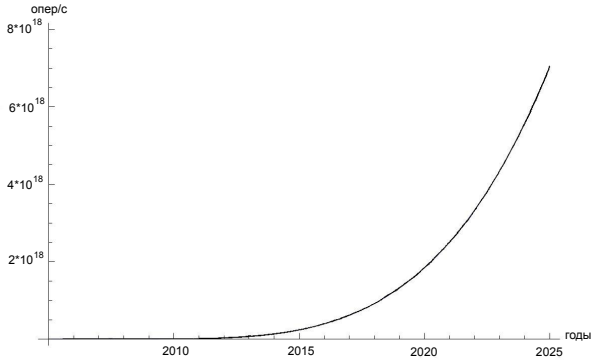

Рис. 2. Предполагаемый рост производительности суперкомпьютеров

Экстраполяция данных трудоёмкости алгоритма NFS при возрастании модуля факторизации и производительности будущих суперкомпьютеров показывает отсутствие динамичности метода NFS.

Так, трудоёмкость факторизации 2048-битного модуля (значение модуля выбрано как далёкое от последних рекордных) стремится к 10^{23} опер/с, в то время как производительность перспективного суперкомпьютера предположительно будет составлять в 2025 г. от 10^{18} до 10^{19} опер/с.

Напрашивается вывод, что примерно с 768-битного модуля метод решета числового поля теряет свою динамичность.

ЛИТЕРАТУРА

1. Информационное сообщение о рекордных факторизациях различных модулей RSA. `http://www.crypto-world.com/FactorRecords.html`.
2. *Глухов М. М., Круглов И. А., Пичкур А. Б., Черемушкин А. В.* Введение в теоретико-числовые методы криптографии. М.: Лань, 2011. 395 с.
3. *Василенко О. Н.* Теоретико-числовые алгоритмы в криптографии. М.: МЦНМО, 2006. 333 с.
4. Официальная страница Top500. `http://www.top500.org/`.

УДК 519.254

ТОЧНОЕ ВЫЧИСЛЕНИЕ РАСПРЕДЕЛЕНИЙ С ПОМОЩЬЮ ЦЕПЕЙ МАРКОВА

А. М. Зубков, М. В. Филина

Для построения статистических критериев с заданными вероятностями ошибок требуется знание распределений используемых в этих критериях статистик. Как правило, точные формулы слишком громоздки с вычислительной точки зрения. Поэто-

му вместо точных формул в качестве приближений часто используют формулы из соответствующих предельных теорем, относящихся к случаям, когда объём выборки неограниченно возрастает. Зависимость точности таких формул от объёма выборки обычно оценивается лишь по порядку. Поэтому представляют интерес методы точного вычисления распределений статистик для выборок умеренного объёма.

В нескольких работах авторов было показано, что в различных комбинаторно-вероятностных схемах можно строить предназначенные для реализации на ЭВМ алгоритмы точного вычисления распределений различных статистик с помощью аппарата неоднородных по времени цепей Маркова.

Так, в [1, 2] предложен способ вычисления распределений так называемых разделимых статистик в полиномиальной схеме, имеющих вид $\zeta = \sum\limits_{k=1}^{N} f_k(\nu_k)$, где $\nu_1, \ldots, \nu_N$ — числа появлений исходов $1, \ldots, N$ с вероятностями $a_1, \ldots, a_N$ в T независимых испытаниях, а $f_1(x), \ldots, f_N(x)$ — заданные функции. Примерами разделимых статистик являются статистика Пирсона $\chi = \sum\limits_{k=1}^{N} (\nu_k - Tp_k)^2/Tp_k$, где $p_1, \ldots, p_N$ — гипотетические вероятности исходов, а также число μ_0 исходов, не появившихся ни разу, число μ_r исходов, появившихся ровно $r \geqslant 1$ раз, и т. п.

Способ основан на том, что последовательность $\sigma_n = \sum\limits_{k=1}^{n} \nu_k$, $n = 0, 1, \ldots, N$, образует (неоднородную по времени) цепь Маркова с переходными вероятностями

$$\mathbf{P}\{\sigma_n = s+m \,|\, \sigma_{n-1} = s\} = C_{T-s}^{m} \left(\frac{a_n}{a_n + a_{n+1} + \ldots + a_N}\right)^m \left(\frac{a_{n+1} + \ldots + a_N}{a_n + a_{n+1} + \ldots + a_N}\right)^{T-s-m},$$

а последовательность сумм $\zeta_n = \sum\limits_{k=1}^{n} f_k(\nu_k)$, $n = 0, 1, \ldots, N$, является аддитивным функционалом от траектории цепи $\{\sigma_n\}$. Поэтому последовательность $\Sigma_n = (\sigma_n, \zeta_n)$, $n = 0, 1, \ldots, N$, тоже является дискретной цепью Маркова с переходными вероятностями

$$\mathbf{P}\{\Sigma_n = (s + m, z + f_n(m)) \,|\, \Sigma_n = (s, z)\} = \mathbf{P}\{\sigma_n = s + m \,|\, \sigma_{n-1} = s\}.$$

Это позволяет рекуррентно находить распределения $\Sigma_1, \Sigma_2, \ldots, \Sigma_N$ обычным умножением векторов распределений на матрицы переходных вероятностей. Распределение ζ_N совпадает с искомым распределением разделимой статистики ζ.

В [4, 5] приведены результаты численного исследования распределений статистики Пирсона для схем с числом исходов до нескольких сотен и числом испытаний до нескольких тысяч. Эксперименты обнаружили неожиданные отличия точных распределений от их обычно используемых аппроксимаций. Игнорирование этих отличий может приводить к принятию неправильных решений из-за некорректной интерпретации результатов статистической обработки данных.

В [3] описан более сложный способ построения цепей Маркова, позволяющих находить распределения чисел связных компонент и циклических точек в графах случайных равновероятных отображений множества из N элементов и в графах итераций двух таких отображений. Результаты вычислений для $N \leqslant 2000$ показали, что при таких N точность приближений, полученных с помощью предельных теорем, невелика.

ЛИТЕРАТУРА

1. *Зубков А. М.* Рекуррентные формулы для распределений функционалов от дискретных случайных величин // Обозр. прикл. промышл. математики. 1996. Т. 3. Вып. 4. С. 567–573.
2. *Зубков А. М.* Методы расчета распределений сумм случайных величин // Труды по дискретной математике. М.: Физматлит, 2002. Т. 5. С. 51–60.
3. *Зубков А. М.* Вычисление распределений чисел компонент и циклических точек случайного отображения // Математические вопросы криптографии. 2011. Т. 1. № 2. С. 5–18.
4. *Filina M. V. and Zubkov A. M.* Exact computation of Pearson statistics distribution and some experimental results // Austrian J. Statistics. 2008. V. 37. No. 1. P. 129–135.
5. *Filina M. V. and Zubkov A. M.* Tail properties of Pearson statistics distributions. // Austrian J. Statistics. 2011. V. 40. No. 1&2. P. 47–54.

УДК 004.432.2

ВЫЧИСЛЕНИЕ СОВЕРШЕННОЙ НЕЛИНЕЙНОСТИ БУЛЕВОЙ ФУНКЦИИ НА ГРАФИЧЕСКОМ ПРОЦЕССОРЕ

А. В. Медведев

Для криптографических приложений представляют интерес функции с высокими показателями «нелинейности», потому что они труднее поддаются анализу. Существует несколько характеристик нелинейности функции, одна из них — совершенная нелинейность [1], которая показывает, насколько функция удалена от класса функций с линейной структурой. Разработан алгоритм вычисления совершенной нелинейности произвольной булевой функции для параллельной реализации на видеокартах NVIDIA, поддерживающих технологию CUDA [2].

Обозначим через $P_2(n)$ множество всех булевых функций от $n \geqslant 1$ переменных и X — область их определения, $X = \{0,1\}^n$.

Определение 1. Для функции $f \in P_2(n)$ и набора $a \in X$ функция $f'_a(x) = f(x) \oplus f(x \oplus a)$ называется *производной функции f по направлению a.*

Определение 2. Говорят, что булева функция f *имеет линейную структуру,* если существует вектор $a \in X \setminus \{0^n\}$, что $f'_a = \mathrm{const}$, т. е. $f = f_a$ либо $\neg f = f_a$. Множество всех функций в $P_2(n)$, имеющих линейную структуру, обозначается $LS(n)$.

Определение 3. Число $CN_f = \mathrm{d}(f, LS(n)) = \min\limits_{g \in LS(n)} \mathrm{d}(f,g)$ называется *совершенной нелинейностью* функции f. Здесь $\mathrm{d}(f,g)$ — расстояние между функциями f и g, равное количеству наборов, на которых они различаются.

Существует ряд способов вычисления CN_f; наиболее подходящим для параллельной реализации оказался следующий:

$$CN_f = \min_{a \in X \setminus \{0^n\}} \min(\mathrm{w}(f'_a), 2^n - \mathrm{w}(f'_a))/2,$$

где $\mathrm{w}(.)$ — вес булевой функции. Основную трудность здесь представляет вычисление функции $f(x \oplus a)$. Проблема состоит в том, что CUDA предполагает запуск большого числа нитей параллельно [3], но их совокупная память ограничена (3 Гб для Tesla C2050), поэтому невозможно хранить значения $f(x \oplus a)$ для каждой нити, и вес производной вычисляется «по частям» (алгоритм 1).