

УДК 621.038

СЖАТИЕ ДВОИЧНЫХ КОДОВ НА ОСНОВЕ БИНОМИАЛЬНЫХ ЧИСЕЛ

Чередниченко В. Б., ст. преподаватель
Национальный университет внутренних дел,
филиал в г. Сумы. E-mail <chered_ukr@ukr.net>

В разнообразных электронных системах нередко применяется сжатие информации. Для сжатия равновесных кодов ранее были предложены простые алгоритмы, которые несложно реализовать аппаратными средствами. При этом исходные комбинации сначала превращаются в двоичные биномиальные коды, а затем последние преобразуются в номера.

Для расчета среднего числа тактов преобразования равновесных кодов в номера была предложена формула:

$$\bar{N} = \left(n - \frac{k(n-k)(n+2)}{(k+1)(n-k+1)} \right) + \frac{C_n^k - 1}{2} \quad (1)$$

где n – длина кодовых комбинаций,
 k – число единиц в комбинации.

Максимум среднего количества тактов преобразования кода в номер \bar{N}_{\max} находится в точке $k = n/2$ для четных значений n , а для нечетных n имеется два одинаковых максимума в точках $k = n/2 + 1/2$ и $k = n/2 - 1/2$. Два минимума этой функции находятся в точках $k=1$ и $k=n-1$. Первым слагаемым в (1) можно пренебречь при $k \geq 3$.

Когда в коде неодинаковы вероятности $p_0, p_1, \dots, p_i, \dots, p_N$ появления различных кодовых комбинаций $b_0, b_1, \dots, b_i, \dots, b_N$, тогда среднее количество тактов нумерации \bar{N} при

различной вероятности появления кодовых комбинаций равно:

$$\bar{N} = \sum_{i=0}^N p_i b_i = p_0 0 + p_1 1 + \dots + p_i b_i + \dots + p_N N \quad (2)$$

Коэффициент уплотнения равновесных кодов равен:

$$S = n / \log_2 C_n^k \quad (3)$$

Этот коэффициент имеет максимальную величину при двух крайних значениях $k = 1$ и $k = n - 1$. Минимальный коэффициент уплотнения имеет место при $k = n/2$ для четных значений n , а для нечетных n минимум находится в двух точках $k = n/2 + 1/2$ и $k = n/2 - 1/2$.

Таким образом, при использовании описанных алгоритмов максимальное сжатие получается при минимальных затратах времени на преобразование.

Предлагается использовать данный метод для сжатия любых двоичных кодов длиной n . Для этого каждую из исходных кодовых комбинаций можно считать равновесной с количеством единиц k в ней. Тогда все сжатые комбинации будут содержать полученный после преобразования код и дополнительные данные о количестве единиц в каждой исходной последовательности $q = \log_2 n$. При этом коэффициент уплотнения уменьшится:

$$S = n / \log_2 C_n^k + \log_2 n \quad (4)$$

Проведенные расчеты показывают, что в окрестности точки $k = n/2$ имеется интервал, где сжатия не происходит. Для $n=16$ его относительная ширина равна $0,5n$, а для $n=128$ она уменьшается до $0,28n$. Тогда при количестве единиц в исходной комбинации, соответствующем этому интервалу удлинения, целесообразно «пропускать» исходной код без обработки. Это улучшает характеристики сжатия и значительно уменьшает суммарное время преобразования.