

МЕТОД СЖАТИЯ ДАННЫХ С НЕРАВНОМЕРНЫМ РАЗБИТИЕМ НА БЛОКИ

**В. А. Лужецкий, д.т.н., профессор,
Адил Ашрафул Хок, аспирант
Винницкий национальный технический университет**

Для сжатия данных широко используются методы, учитывающие статистику сообщений, и словарные методы. Однако известные их недостатки стремятся устранить как различными усовершенствованиями, так и поиском принципиально новых подходов к трактовке данных, подлежащих сжатию.

Один из таких подходов заключается в рассмотрении данных, независимо от их фактического содержания, как чисел. При этом файл данных, который необходимо сжать, рассматривается как М-представление нескольких сверхбольших целых положительных чисел одинаковой разрядности (1024 и больше разрядов) линейной формой Фибоначчи. Именно свойство М-представления и обеспечивает сжатие данных.

Экспериментальные исследования метода сжатия с использованием М-представления, проведенные авторами, показали, что коэффициент сжатия существенно зависит от разрядности блоков данных и их содержания. Блоки данных, которым соответствуют числа, сравнимые с числами Фибоначчи, сжимаются с высоким коэффициентом сжатия. Если же число, соответствующее блоку данных, значительно отличается от ближайшего числа Фибоначчи, то такой блок мало сжимается.

Исходя из этого, в докладе рассматривается метод,

который предусматривает разбиение файла данных на блоки различной длины таким образом, чтобы числа, которые соответствуют их содержанию, имели М-представления, обеспечивающие наперед заданный коэффициент сжатия.

Пусть k - желаемый коэффициент сжатия, n - разрядность блока, тогда разрядность М-представления должна быть

$$n^* = n(1 - k). \quad (1)$$

В предлагаемом методе предусматривается разрядность блока от n_{\min} до n_{\max} .

Сначала выбирается блок разрядности n_{\min} и для него вычисляются координаты и индекс М-представления. Если полученная разрядность М-представления не удовлетворяет условию (1), то разрядность блока увеличивается на единицу, и вновь вычисляются координаты и индекс М-представления.

Такие действия выполняются до тех пор, пока либо не будет получено М-представление необходимой разрядности, либо разрядность блока не превысит n_{\max} . В последнем случае выбирается тот вариант М-представления, который имеет наименьшую разрядность из всех рассмотренных до этого. Таким образом, разрядность блока данных адаптируется под заданный коэффициент сжатия.

Неравномерность разбиения файла данных на блоки обеспечивает увеличение коэффициента сжатия по сравнению с разбиением на блоки одинаковой разрядности, но приводит к увеличению времени сжатия, тогда как время восстановления исходного файла остается одинаковым.