Dynamic Huffman Coding

Eduardo Sany Laber, Ruy Luiz Milidiu, Artur Alves Pessoa.
Deparamento de Informática, PUC-RJ, Brasil
{laber,milidiu,artur}@inf.puc-rio.br

A static Huffman code for a given source alphabet $\{a_1, \ldots, a_n\}$ is a prefix code generated by Huffman's algorithm. Such a code minimizes $\sum_{i=1}^{n} w_i l_i$ among all prefix codes, where w_i is the weight of a_i and l_i is the number of bits used to encode a_i.

For data communication purposes, the initial parsing required by the static Huffman algorithm represents a big disadvantage. That happens because the data must be transmitted on-line. As soon as the symbol arrives to the transmitter, it must be encoded and transmitted to the receptor. For these situations, adaptive Huffman codes have been largely used. This method determines the mapping from symbol alphabets to codewords based upon a running estimate of the alphabet symbol weights. The code is adaptive, just changing to remain optimal for the current estimates.

Two methods are presented in the literature for implementing dynamic Huffman coding. The first one is the FGK algorithm [1] and the second one is the Λ algorithm [2]. In [2], Vitter proved that the total number of bits D_t transmitted by the FGK algorithm for a message with t symbols is bounded below by $S_t - n + 1$, where S_t is the number of bits required by the static Huffman method and bounded above by $2S_t + t - 4n + 2$. Furthermore, he conjectured that D_t is bounded above by $S_t + O(t)$. We present an amortized analysis to prove this conjecture by showing that $D_t \leq S_t + 2t - 2k - \lceil \log \min(k+1, n) \rceil$, where k is the number of distinct symbols in the message [3]. We also present an example where $D_t = S_t + 2t - 2k - 3\lfloor (t-k)/k \rfloor - \lceil \log(k+1) \rceil$, showing that the proposed bound is asymptotically tight. These results explain the good performance of FGK observed by some authors through practical experiments.

[1] Knuth, D.E., *Dynamic Huffman Coding*, **Journal of Algorithms 6,2(1985), 163-180.**

[2] Vitter, J. S., *Design and Analysis of Dynamic Huffman codes*, **Journal of ACM 34, 4(1987).**

[3] Milidiú, R.L., and Laber, E.S., and Pessoa, A. *Improved Analysis of FGK Algorithm*, **Monografias em Ciência da Computação 39/97, Departamento de Informática, PUC-RJ, 1997. Submitted to Journal of the Algoritms.**