Researcher finds optimal fix-free codes

April 3 2009

Dr. Serap Savari

(PhysOrg.com) -- More than 50 years after David Huffman developed Huffman coding, an entropy encoding algorithm used for lossless data compression in computer science and information theory, an electrical and computer engineering faculty member at Texas A&M University has discovered a way to construct the most efficient fix-free codes.

Huffman coding uses a variable-length code table for choosing the representation for each symbol, resulting in a prefix code (that is, the bit string representing some particular symbol is never a prefix of the bit string representing any other symbol) that expresses the most common characters using shorter strings of bits than are used for less common

source symbols. Huffman was able to design the most efficient compression method ofthis type since no other mapping of individual source symbols to unique strings of bits will produce a smaller average output size when the actual symbol frequencies agree with those used to create the code.

Dr. Serap Savari, an associate professor in the Department of Electrical and Computer Engineering at Texas A&M, has developed the first approach to finding the optimal fix-free code, variable length codes in which no codeword is the prefix or suffix of another codeword.

“My method of finding optimal fix-free codes is computationally demanding, but no one has solved the problem before even though it was first posed in 1990,” Savari said. “Earlier algorithms produced good fix-free codes in a reasonably (time) efficient way, but without the guarantee of optimality.”

While there are numerous applications for fix-free codes, the most important applications have been in communications. Fix-free codes have been investigated for joint source-channel coding and have been applied within the video standards H.263+ and MPEG-4 because their property of efficient decoding in both the forward and backward directions assists with error resilience. They are also interesting for problems in information retrieval such as searching for patterns directly in compressed text. Savari is uncertain how her discovery will impact these and other applications of fix-free codes, but hopes that her work will be used by researchers and people implementing practical systems..

“My work is like Huffman’s in that it is basic research that is motivated by practically important problems and which contributes to the theory of data compression,” she said.

Savari has already been invited to discuss her findings at numerous

seminars throughout the United States, including Stanford University, the University of Illinois-Urbana- Champaign, The University of California, Berkeley, the University of California, San Diego, Caltech, the University of Southern California and possibly MIT in the fall.

Provided by Texas A&M University