On Exact Sequence Reconstruction Over a Stochastic t-Error Channel

Vivian Papadopoulou[1,2], V. Arvind Rameshwar[3], and Antonia Wachter-Zeh[1]

[1]Technical University of Munich, Institute for Communications Engineering, Munich, DE
[2]Imperial College London, Department of Electrical and Electronic Engineering, London, UK
[3]India Urban Data Exchange, Indian Institute of Science, Bengaluru, IN

Abstract

In this work, we consider a modification of the classical sequence reconstruction problem of Levenshtein (2001), wherein an uncoded binary sequence $\mathbf{u}$ of length k is transmitted across a substitution channel that introduces exactly t errors. We assume that $\mathbf{u}$ is transmitted several times across the channel, and that the channel introduces t-error patterns that are drawn uniformly at random, without replacement. We are interested in obtaining estimates of the expected number of transmissions E for $\mathbf{u}$ to be reconstructed exactly by the decoder. As a by-product of our analysis, we obtain estimates of the expected number of transmissions for exact reconstruction when the errors are drawn, potentially with repetitions, from the t-Hamming sphere around $\mathbf{u}$. Such a problem arises naturally in the setting of communication with (noiseless) feedback, where the transmitter obtains the potentially erroneous vectors received by the decoder as feedback and is allowed to transmit $\mathbf{u}$ repeatedly until successful reconstruction. Our approach is to analyze the expected number of transmissions E^{Maj} when the decoder employs the simpler, but potentially sub-optimal, majority-vote rule, and thereby obtain estimates of E. First, using a numerical sampling-based approach, we show empirically that the expected number of transmissions E^{Maj} for successful reconstruction is strictly less, in this average-case setting, for selected k, t values, than the number of transmissions E_{L} in the worst-case setting of Levenshtein (2001). Next, we present upper bounds on E^{Maj}, via a recursive algorithm, which, for small values of t, enables us to obtain analytical upper bounds on E^{Maj}, and thereby on E. Further, for the special cases when $t = 1$ and $t = k - 1$, we show that $E = E_{\text{L}}$, indicating that even in the average-case setting, one cannot reconstruct the input sequence with fewer transmissions than in the worst-case, for these parameters.