

Bounds on Capacity of Deletion Channel

Siva Theja Maguluri and Bruce Hajek
University of Illinois Urbana Champaign



1 Background

A binary symmetric channel where each bit is deleted independently with some probability is very well understood. A slight variation of this is a binary deletion channel, where each bit is deleted with certain deletion probability, 'd'. This differs from the erasure channel in that in erasure channel, one knows which bits are deleted. Deletion channel is a special case of a channel with Synchronization errors. Shannon's theorem for these channels was established in 1967 by Dobrushin[1]. Turns out that this channel is very difficult to understand. There is no known method to compute the capacity of this seemingly simple channel.

2 Lower Bound

One way to find a lower bound on the capacity is to restrict our attention to (symmetric) first order Markov inputs. When the input is a symmetric Markov process with parameter γ , output is also a Markov process with parameter $\frac{1-\gamma}{1+d(1-2\gamma)}$. For an input sequence X and output sequence Y,

$$\Pr(Y | X) = \#S(X, Y) d^{n-m} (1-d)^m$$

Where, #S(X,Y) is the number of ways in which Y can be covered by X. A long X sequence is generated according to Markov process, and a deletion channel is simulated to get an output sequence and #S(X,Y) is calculated using dynamic programming. Then, using the following (remains to be proven)

$$H(Y | X) = \lim_{n \rightarrow \infty} \frac{1}{n} \Pr(Y^n | X^n)$$

mutual information is calculated. This is done for various values of γ and the optimum value is chosen.

3 Upper Bound

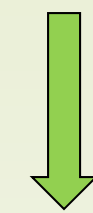
If there is a genie which gives information about which bits are deleted, one gets an erasure channel. Thus, capacity of an erasure channel is an obvious upper bound on the capacity of a deletion channel. If instead, the genie adds markers after every n bits, and the markers are not deleted, then the capacity of this channel is an upper bound on the capacity of deletion channel. This new channel however is a discrete memoryless channel with 2^n inputs and $(2^{n+1}-1)$ outputs, and it's capacity can be computed using Blahut-Arimoto Algorithm. This was calculated using $n=11$.

5 Conclusion

Simple computer based methods give better bounds on capacity of deletion channel

- All the lower bounds so far have focused on first order Markov inputs. Using higher order Markov inputs might improve the lower bound
- The ratio of current upper bound and lower bound is large for large d. So, we need better bounds for large deletion probability.

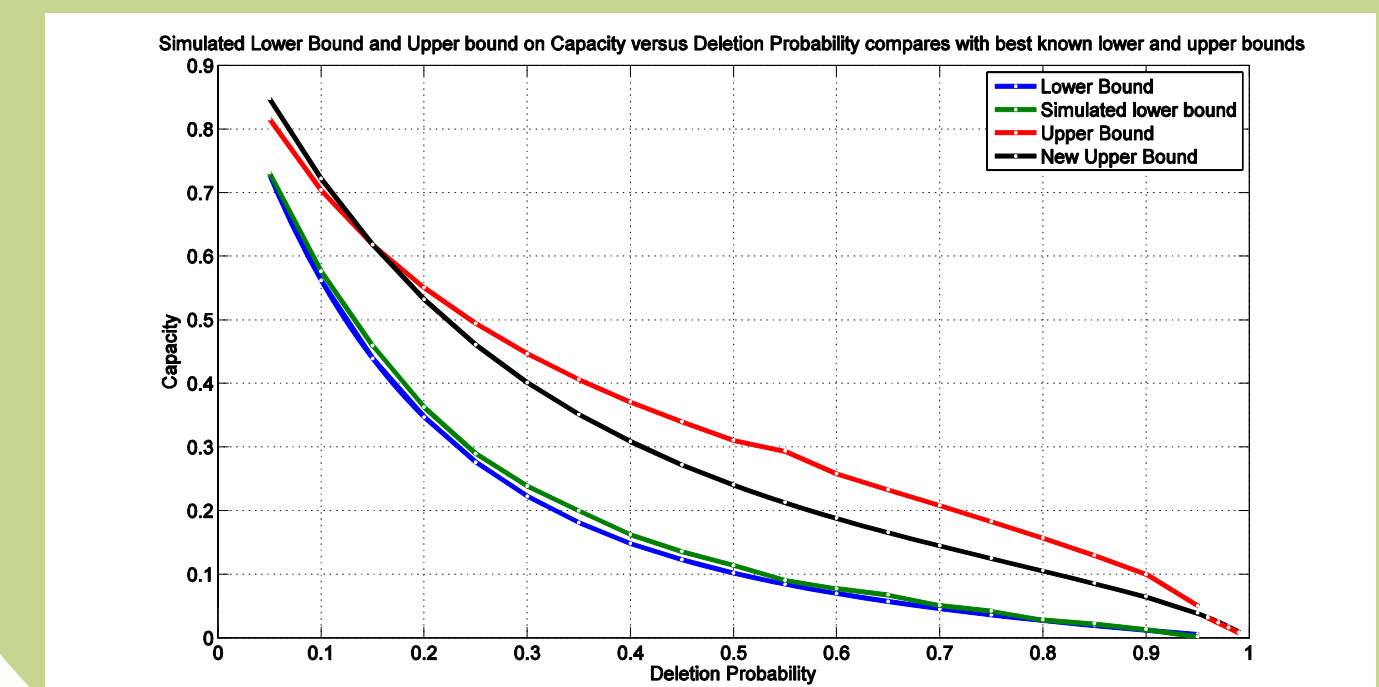
100110101111001101



100111011

4 Results

Calculated upper bound and simulated lower bounds are shown comparing with the best known upper and lower bounds so far.



References

1. R.L.Dobrushin. Shannon's theorem for Channels with Synchronization Errors. Problems of Information Transmission, 3(4):11-26,1967
2. M. Mitzenmacher. A survey of results for deletion channels and related synchronization channels. Summary appeared in SWAT 2008

