# MAP Detection in Noisy Channels with Synchronization Errors (Including the Insertion/Deletion Channel)

Wei Zeng, Aleksandar Kavčić
DEAS, Harvard University, Cambridge, MA, 02138
{wzeng, kavcic}@deas.harvard.edu

*Abstract —* **In this paper, we consider the problem of minimizing the symbol error probability when detecting a sequence of symbols transmitted through channels with synchronization errors (or, timing uncertainties). The channels may also be corrupted by additive white Gaussian noise (AWGN) and intersymbol interference (ISI). We build a special trellis to jointly characterize the ISI channel and the timing uncertainty, then propose a soft-output detection algorithm that performs maximum *a posteriori probability* detection on each input symbol as well as the timing uncertainty of each received sample.**

Reliable transmission of information over channels with synchronization errors is important in many real applications, such as communication systems and data storage systems. Due to the timing uncertainty of the receiver, the received sequence has a *random* length. One extreme case of such channels is the "insertion/deletion channel", in which a transmitted symbol may be skipped (deleted) or repeated (inserted), say, a symbol is sampled twice by the receiver.

Denote the channel input sequence $x_1, x_2, \ldots, x_n$ as $x_1^n$. We assume that the source is a stationary, discrete-time, first order Markov source, i.e., we have for any integer $m \geq 0$

$$P(X_{t+1}|X_{t-m}^t) = P(X_{t+1}|X_t). \tag{1}$$

To make the analysis simple, we also assume that $X_t$ is binary.

Let $T$ be the symbol interval. An ideal receiver would sample the waveform at times $0, T, 2T, 3T, \ldots$. However, because of the timing errors, the receiver samples the waveform at times $(0+\varepsilon_0)T$, $(1+\varepsilon_1)T$, ..., $(k+\varepsilon_k)T$, and so on. We assume that variable $\varepsilon_k$ forms a first order Markov process, and its value is determined by uniformly quantizing the symbol interval T to $Q$ levels, i.e.,

$$P(\varepsilon_k|\varepsilon_{k-1}) = \begin{cases} \delta & \text{if} \quad \varepsilon_k = \varepsilon_{k-1} \pm \frac{1}{Q} \\ 1 - 2\cdot\delta & \text{if} \quad \varepsilon_k = \varepsilon_{k-1} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

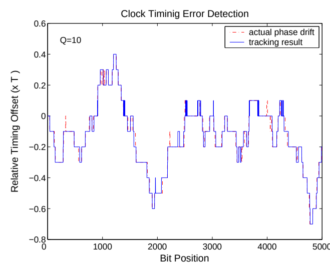
Therefore, there are three possibilities for the number of samples in each symbol interval: no sample, one sample, or two samples.

Define the $k-th$ symbol interval as $((k-1)T, kT]$ on the time axis. Our states for timing uncertainty at each symbol interval are defined so as to represent all sampling possibilities during the interval $((k-1)T, kT]$. Generally, if we quantize each symbol interval to $Q$ levels, we form a timing uncertainty trellis with $Q+2$ states.

We assume we have a pulse shape $s(t)$ which occupies $3$ symbol intervals. If all samples are taken at multiples of $T$, the ISI length is $1$, and the partial response polynomial has the form $h(D) = 1 + D$. However, if there is a timing offset, the ISI length is 2.

The final trellis is obtained as a cross product of the timing uncertainty trellis and the ISI trellis and thus has $4(Q+2)$ states, which jointly characterize the timing uncertainty and ISI. One interesting and important example is the special case when we have only $Q = 1$ quantization level in each symbol interval, and the pulse shape $s(t)$ is a perfect rectangular pulse that occupies only one symbol interval. In this case, the timing error trellis will contain only 3 states in each column: the 0-sample (deletion) state, the 1-sample state (synchronous), and the 2-samples (insertion) state. Since $s(t)$ is perfect rectangular shape, the ISI trellis will have only two states. So the final trellis has only 6 states.

The algorithm we derive resembles the well-known BCJR algorithm [2], but is appropriately modified to handle synchronization errors. An important difference is that the BCJR algorithm requires a memory size proportional to $n$, while the optimum version of the proposed algorithm requires a memory size proportional to $n^2$.



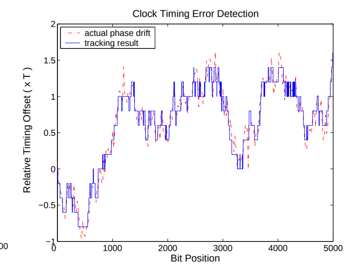




Fig.1 and Fig.2 demonstrate the timing offset tracking results of the soft output algorithm when the timing error is quantized to 10 and 5 levels per sampling period, respectively. In Fig.1 the random phase change is generated by a discrete Markov chain in (2) with $\delta = 0.01$ and $Q = 10$ quantization levels. That is, the sampling phase in the next period will drift to the adjacent phase level with probability $0.01$, and we build the trellis with the same model. Fig.2 shows the results of a more realistic case that models the timing offset increment at each sample to be an independent and identically distributed Gaussian random variable with mean $0$ and variance $\frac{\sigma}{T} = 3\%$. We still track the timing uncertainty by our discrete quantized timing offset model. The model we use has $Q = 5$ levels, and we choose $\delta = 0.01$ so that we have the same variance for the independent timing offset increment. For each case, we transmit $5000$ bits and the signal-to-noise ratio (SNR) is $8$dB. All the information bits are decoded without errors, although we notice that in Fig.2, we've skipped $5$ input symbol and inserted $3$ symbols at the receiver side over the course of $5000$ symbol intervals.

## ACKNOWLEDGMENTS

We thank Michael Mitzenmacher for helpful discussions.

## REFERENCES

[1] M. C. Davey and D. J. C. Mackay, "Reliable communication over channels with insertions, deletions and substitutions," *IEEE Trans. Information Theory*, vol. 47, pp. 687–698, Feb 2001.

[2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Information Theory*, vol. 20, pp. 284–287, March 1974.

[1]This work was supported by Seagate Technology.