Linear Support Vector Machines for Error Correction in Optical Data Transmission

Alex Metaxas[1], Ray Frank[1], Alexei Redyuk[2], Yi Sun[1], Alex Shafarenko[1], Neil Davey[1], Rod Adams[1]

[1]Biological and Neural Computation Research Group, School of Computer Science
University of Hertfordshire, Hatfield, Herts. AL10 9AB UK
http://homepages.feis.herts.ac.uk/~nngroup/
(a.metaxas, y.2.sun, a.shafarenko, n.davey, r.g.adams)@herts.ac.uk
[2] Novosibirsk State University, Novosibirsk, 2 Pirogova street, 630090, Russia
alexey.redyuk@gmail.com

Abstract. Reduction of bit error rates in optical transmission systems is an important task that is difficult to achieve. As speeds increase, the difficulty in reducing bit error rates also increases. Channels have differing characteristics, which may change over time, and any error correction employed must be capable of operating at extremely high speeds. In this paper, a linear support vector machine is used to classify large-scale data sets of simulated optical transmission data in order to demonstrate their effectiveness at reducing bit error rates and their adaptability to the specifics of each channel. For the classification, LIBLINEAR is used, which is related to the popular LIBSVM classifier. It is found that is possible to reduce the error rate on a very noisy channel to about 3 bits in a thousand. This is done by a linear separator that can be built in hardware and can operate at the high speed required of an operationally useful decoder.
Keywords: Error correction, classification, optical communication, adaptive signal processing

1 Introduction

Fibre optic communication links are extensively used for high-speed and long-distance data transmission. For example, the internet backbone primarily consists of fibre optics trunk lines, bundles of fibre optic cables combined together [11]to provide increased capacity [11] (e.g. Trans-Atlantic links). Furthermore, Nielsen's Law of Internet bandwidth [9] states that "a high end user's connection speed grows by 50% per year", an exponential growth of bandwidth year on year. Dutton highlights the problem in [4]: The faster the link the lower we need the error rate to be! But the harder that low error rate becomes to deliver. Therefore, improving the performance (lowering the *Bit Error Rate*, *BER*) of fibre optic links is not only an important task but is one that is also difficult to achieve. Fibre optic link performance is affected by a variety of phenomena, including attenuation, chromatic dispersion and non linear effects, which combine to cause signal degradation. In addition, each particular link has its own characteristic signature of transmission impairments [10] [6]. As stated by Hunt et al in [7]: There is great value in a signal post-processing system that can undo some of these signal distortions, or that can separate line-specific distortions from non-recoverable errors. Signal post-processing in optical data communication can offer new margins in system performance in addition to other enabling techniques.

In this paper we build on our earlier work by using a much bigger and noisier data set than we have previously analysed. In order to work with such a data set we have used an optimised linear SVM which improves upon our previous use of a neural network approach using a pereptron based method.

2 Background

Communication of digital signals along physical media typically requires that the bits are encoded into a time-varying signal at the transmitter, transmitted along the medium, and then decoded back into a digital signal at the receiver. The basic operation of an optical communication system is as follows (see Figure 1)[4]: A serial bit stream in electrical form is presented to a modulator, which encodes the data appropriately for fibre transmission. A light source (laser or Light Emitting Diode - *LED*) is driven by the modulator and the light focused into the fibre. The light travels down the fibre (during which time it may experience dispersion and loss of strength). At the receiver end the light is fed to a detector and converted to electrical form. The signal is then amplified and fed to another detector, which isolates the individual state changes and their timing. It then decodes the sequence of state changes and reconstructs the original bit stream. The timed bit stream so received may then be fed to a using device. The process is described in Figure 1.

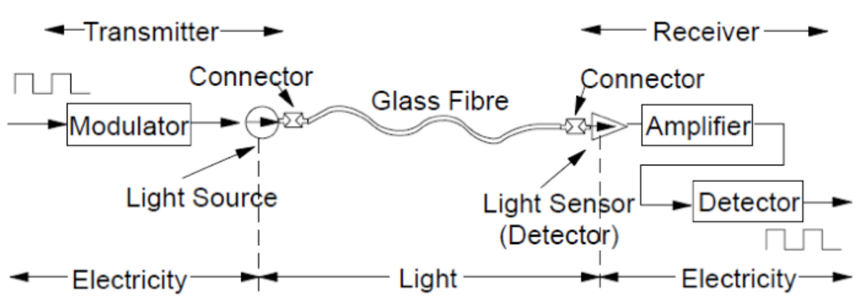

Fig. 1. The process of optical transmission

There are three broad categories of signal degradation in optical systems [1]: Attenuation – decay of signal strength, or loss of light power, as the signal propagates through the fiber. Chromatic dispersion – spreading of light pulses as they travel down the fiber. Nonlinear effects – cumulative effects from the interaction of light with the material through which it travels, resulting in changes in the light wave and interactions between the light waves.

In this paper we attempt to use a trainable classifier to help reduce the number of bits that are incorrectly decoded due to degradation.

One important feature of this problem domain is that if the resulting trained classifier is going to be useful it must be extremely fast. Optical channels can operate at speeds of over 50GHz. Clearly a classifier will only be useful if it is built in hardware. To this end we have used a simple linear separator, which can easily be built in hardware. In a related application this speed requirement is discussed and an SVM is instantiated on a *FPGA* (Field-Programmable Gate Array) board and classification is done at over 10GHz (see [8]). In our earlier work we found the linear separator using perceptron learning in a neraul network abased approach. However training a perceptron to find a good separator, particularly on a large data set is known to be difficult [3]. So in this work we use a linear SVM to find an effective linear separator.

3 The Data

The data is of the form described in [7], consisting of a large number of bits encoded as the electrical signal produced following the conversion of the optical signal into an electrical current. Each bit is therefore encoded as a waveform. The waveform is represented by using 32 evenly spaced samples of the intensity level within a bit time slot, producing a 32-ary vector of real numbers. The sum of these 32 values is the *energy* of the signal. Figure 2 shows an example of a stream of five bits. The original bit stream is also recorded so that each wave has an associated binary label.

The data was produced by a simulation of single transmission channel, which was deliberately made to have a high level of noise, in order to produce misclassifications. So the data set we use is large, consisting of a sequence of 611,430 bits of which 105,890 or 17.32% are misclassified by an optimal energy thresholder. We divided this data set into 4/5 training and 1/5 testing, by using the last 122,286 bits as the test set. For the whole data set we searched for the energy threshold value that gave the best decoding of the data stream, that is it gave the best reconstruction of the original binary data stream. As this is a very noisy channel the error rate even with the best threshold is high. We denote those bits in the data stream that are correctly decoded by the threshold as *easy*, and those that are incorrectly decoded as *hard*. Table 1 gives a breakdown of the data set.

The hard bits usually come from either the sequence "101" or "010" where the central bit is often distorted by the energy of the bits surrounding it. Figure 3 shows some examples of the "010" subsequence. It can be seen that the red misclassifications do not have sufficient energy to be classified as 1's.

In order to represent this data for a trainable classifier we simple took the 32-ary wave vector for each bit and catenated the representation of the bits to its left and right, giving a 96-ary real vector. The motivation for this was that

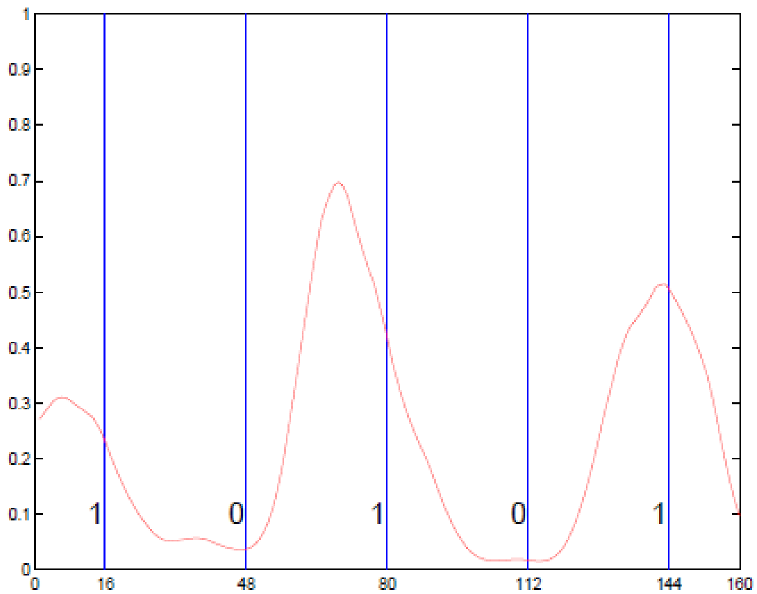

Fig. 2. An Example bit stream

Table 1. The breakdown of the data set

Data Set	*Training*	*Testing*
Easy	404,432	101,108
Hard	84,712	21,178
Total	489,144	122,286
Easy(%)	82.68	
Hard(%)	17.32	

the surrounding bits have a clear influence on the wave of the bit between them and this information could be of use to the classifier. In summary our data set consists of 611,430 96-ary labelled real valued vectors.

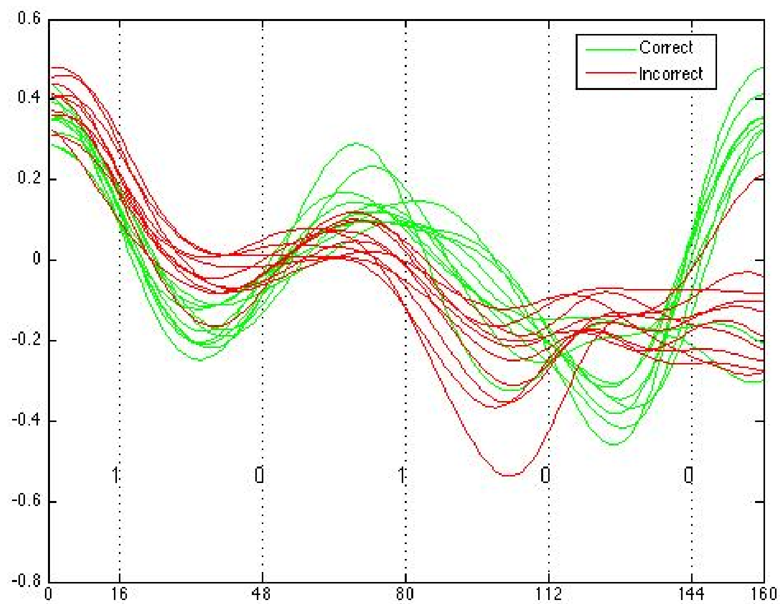

Fig. 3. Waves where the central bit is correctly or incorrectly decoded by the threshold

4 The Classifier

As we have said earlier we use a linear SVM to find a good separator of our data.

4.1 Software Used

The actual tool we used is LIBLINEAR [5] which is a linear classifier produced by the authors of the well known LIBSVM [2]. It supports the same data formats as LIBSVM but is more suited to classification of large data sets with [5]: "millions of instances and features".

4.2 Training

The only hyper parameter in a linear SVM is the regularising cost parameter C. To find a good value for C we simply undertook an empirical search using 5 fold cross validation in the training set.

5 Results

The first thing to note is that LIBLINEAR handled this huge data set without difficulty. This is quite impressive as the training set alone contained 489,144 96-ary vectors, or 46,957,824 real numbers. The search for a good value for C took about an hour on an Intel QX6700 Core 2 Extreme processor. Table 2 gives the final classification rates on the test set.data in order to demonstrate their effectiveness at reducing bit error

Table 2. Final Results

Classifier	*Accuracy (%)*	*Error Rate (%)*	*Error Split*	
Threshold	82.68	17.32	*easy set (%)*	*hard set (%)*
LIBLINEAR	99.62	0.38	62.7	37.3

We can see that the SVM has corrected many of the original errors. In numerical terms the 21,178 original errors have been reduced to just 437. We cannot make a direct comparison with our earlier work, using perceptrons, as we have never before used such a large data set. However on a subset of this data, about one fifth, we were previously able to get a best error rate of 1.15% [6], as against 0.38% here.

Figure 4 shows a wave which the SVM was able to correct and one that it could not correct. For example it is able to correct the red wave which has poor alignment but in context is recognisably a “one”. However the magenta wave has both poor alignment and poor shape and the classifier is unable to correct it.

Table 3 shows the number of incorrectly identified 3 bit sequences (in which the middle bit is incorrectly identified). Notice that, as is usually the case it is the “101” and ’010” sequences that present the biggest problem for both thresholder and the SVM. Nevertheless the classifier is able to correct many of the errors of the thresholder. For example the 6,411 thresholder decoding errors for the “010” sequence are reduced to just 128 by the SVM. The thresholder makes no misclassifications of the “000” bit sequences. It does make a small number misclassification of “011” and “100”.

6 Discussion

6.1 Analysis of Results

The results of the paper show, quite definitively, that error correction of optical signals using linear support vector machines can approach the target BER (as

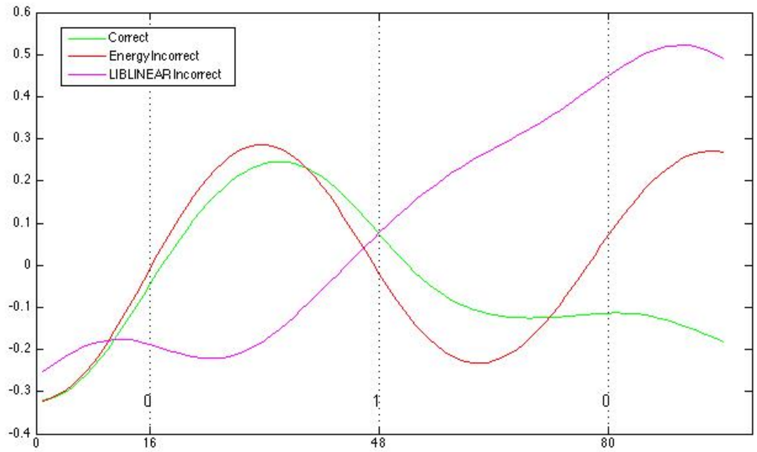

Fig. 4. The Classification of various waves

Sequence	Threshold Errors	SVM Errors
000	0	1
001	4,102	35
010	6,411	128
011	19	34
100	16	26
101	6,439	179
110	3,987	20
111	138	14

Table 3. Number of Errors Made

stated in [10]) of 0.1%, or less than one erroneous bit in a 1000. This is true even in the case of a very noisy channel with high thresholded BER, as demonstrated by the massive reduction in error produced by the SVM. In the experiments the linear kernel SVM (LIBLINEAR) achieved significant gains over the previously achieved results using neural networks and other trainable classifiers. Importantly it is possible to build a hardware based classifier that can work at speeds of over 10GHz, and by parallelising the classification in an appropriate way speeds of over 100GHz should be possible. Moreover a FPGA board based classifier can be reprogrammed should the characteristics of the data channel, being decoded, change.

6.2 Linear Kernel SVM

The performance of LIBLINEAR was notable due to both its improvement over the previous results in [6], a 70% reduction of the BER, and its high operating speed. It by far outstripped the training and prediction speeds of other classifiers we have used, making it possible to analyse the very large data set presented here. This indicates that it may be more easy to implement in hardware and certainly that, due to its reduced dimensionality, its computational cost is low. While it didn't achieve the target BER of 0.1% it is certainly worth further investigation. In other work [10] we have used different representations of the wave, for example adding the energy of the waves to the input and it would be interesting to try this out with LIBLINEAR. Also, a more extensive search of the C space may locate more optimal settings as a number of local minima were observed.

References

1. Greg Bernstein, Bala Rajagopalan, and Debanjan Saha. *Optical Network Control: Architecture, Protocols, and Standards*. Addison Wesley, August 2003.
2. Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a library for support vector machines. 2001.
3. Ronan Collobert and Samy Bengio. Links between perceptrons, mlps and svms. In Carla E. Brodley, editor, *ICML*, volume 69 of *ACM International Conference Proceeding Series*. ACM, 2004.
4. Harry J. R. Dutton. *Understanding optical communications*. Prentice Hall PTR, 1998.
5. R. E Fan, K. W Chang, C. J Hsieh, X. R Wang, and C. J Lin. LIBLINEAR: a library for large linear classification. *The Journal of Machine Learning Research*, 9:1871–1874, 2008.
6. Stephen Hunt, Yi Sun, Alex Shafarenko, Rod Adams, Neil Davey, Brendan Slater, Ranjeet Bhamber, Sonia Boscolo, and Sergei Turitsyn. Correcting errors in optical data transmission using neural networks. In Konstantinos Diamantaras, Wlodek Duch, and Lazaros Iliadis, editors, *Artificial Neural Networks – ICANN 2010*, volume 6353 of *Lecture Notes in Computer Science*, pages 448–457. Springer Berlin / Heidelberg, 2010.

7. Stephen Hunt, Yi Sun, Alex Shafarenko, Rod Adams, Neil Davey, Brendan Slater, Ranjeet Bhamber, Sonia Boscolo, and Sergei K. Turitsyn. Adaptive electrical signal post-processing with varying representations in optical communication systems. In *Engineering Applications of Neural Networks*, pages 235–245. 2009.
8. Dzmitry Maliuk, Haralampos-G. Stratigopoulos, and Yiorgos Makris. An analog vlsi multilayer perceptron and its application towards built-in self-test in analog circuits. In *Proceedings of the 2010 IEEE 16th International On-Line Testing Symposium*, IOLTS '10, pages 71–76, Washington, DC, USA, 2010. IEEE Computer Society.
9. J. Nielsen. Nielsen's law of internet bandwidth. *Online at http://www. useit. com/alertbox/980405. html*, 1998.
10. Yi Sun, A. Shafarenko, R. Adams, N. Davey, B. Slater, R. Bhamber, S. Boscolo, and S.K. Turitsyn. Adaptive electrical signal Post-Processing in optical communication systems., 2008.
11. Jeff Tyson. How internet infrastructure works, April 2001.