Finding events in noisy signals

L. S. Smith*, S. Shahid*, A. Vernier, N. Mtetwa***

**Department of Comp. Sc. and Maths*
University of Stirling
Stirling FK9 4LA, UK
{lss, ssh}@cs.stir.ac.uk

***Lab d'Informatique*
Université de Franche-Comté
Besançon
France

Abstract— **We produce a formal description of the problem of finding events in noisy signals. The particular problem of finding action potentials in noisy electrophysiological recordings is examined, and a method based on simulation of the underlying processes is used to assess different signal interpretation techniques. We discuss the broader application of this approach.**

Keywords **– data generation, spike detection.**

I INTRODUCTION

Finding events in noisy signals is an archetypical task throughout physics (and biosignal analysis is one corner of this area). The task can be straightforward if the signals associated with each event type are quite distinct, and the noise is small, but can be impossibly difficult if the signals for each event type are less distinct, and/or there is a large amount of noise.

Here we start by formalising the problem, and use this to introduce the different issues that arise particularly in biosignal analysis. We will then discuss one particular biosignal analysis problem (namely finding the spikes for different neurons in an electrophysiological recording) which is of importance to us as we are supplying signal processing services for the CARMEN neurophysiology data archive [1]). For this, we use a biophysically based signal and noise generator to compare processing techniques. We then discuss how this approach may be relevant in other areas.

II FORMALISING THE PROBLEM

The general formulation is that we have some type of experimental apparatus in which events occur, and we receive information about these events from instrumentation which provides signals. One example event from the biosignal domain could be an eye blink, measured (instrumented) using an infra-red camera. Normally, (i) there are many different types of events (and we are only interested in some of them), (ii) the relationship between the actual event and the signal resulting from that event is quite complex, and often only partly characterised, and (iii) the instrumentation also produces signals which are not related to the events of interest (noise). This formulation could apply to detecting or classifying sunspots as easily as to detecting or classifying events in biomedical systems.

We assume that there are a number of event types, $E_i : i = 1...N_E$, where N_E is the number of event types. For each event type E_i there is a sequence of events $e_i(j), j = 1...N_{E_i}$ where N_{E_i} is the number of events of type E_i. Each event is characterised by some form of change inside the experimental apparatus, which we write as $s(e_i(j))$. $s(e_i(j))$ is not directly measurable. We consider that the $s(e_i(j))$ are partially ordered in time. Each event may have a non-zero duration, so we fix on some specific point inside the event and call that the time of the event, $t(e_i(j))$. The measurements that we use have come from a set of detectors $D_k : k = 1...N_K$, each producing a signal $d_k(t)$. The general task is to find information about the $e_i(j)$ from the $d_k(t)$: this could be (for example) to find the event times ($t(e_i(j))$), or other characteristics of $s(e_i(j))$ such as location, duration, strength, etc. Exactly what we wish to know will depend on the nature of the events.

In order to tackle this problem, we need to understand the nature of the changes in the $d_k(t)$ caused by the $s(e_i(j))$. Further, we need to understand the nature of the noise in the system. Noise may arise from some of the event types in which we are not interested, from other activity inside the experimental apparatus, from the detectors themselves, to name three possibilities.

a) The Biosignal Context

In the biosignal context, the events could be (e.g.) heartbeats, neural action potentials, or even (for brain/computer interfacing, BCI) particular directed thoughts. The detectors might be (e.g.) skin-based electrodes, or an array of extracellular electrodes, or an array of pixels in a camera. In some cases, the nature of the transformation from $s(e_i(j))$ to $d_k(t)$ is known and characterisable: however in other cases (the BCI example above, particularly) this relationship is entirely unknown. Generally, we have some information about this transformation. The noise sources similarly may be characterisable (for example, noise from other neurons, or from other skin electrical potentials), or it may be more difficult to gauge. In general, the better we can characterise the signals from the events and the noise, the more we can prescribe particular processing techniques. If we have little information about the actual signal and noise it becomes more difficult to know how to proceed.

One major issue is determining the validity of the results. To achieve this, we need to have some independent (and accurate) measure providing the characteristics of the events in which we are interested. In some circumstances this may be possible, (for example, in many BCI paradigms, subjects are asked to direct their thoughts in response to a particular icon being displayed on a screen), but in others (for example measuring the different intervals in the ECG, or determining which neurons have fired) this may be almost impossible. Since knowing how accurately we are characterising events is vital to assessing different techniques, we need to find techniques for coping with the absence of such "ground truth" data. To investigate this further we concentrate on one specific area.

III: DETECTING NEURAL SPIKING ACTIVITY USING EXTRACELLULAR ELECTRODES

Detecting neural action potentials (spikes) using an extracellular electrode (that is, an electrode which is in the medium near the neurons) is a good example problem: the transformation from event to signal is not easily characterisable, and there is considerable interest in accurate timing of these events for later spike train analysis.

In this case, we can interpret the E_i to be either the event that some nearby neuron has generated an action potential (in which case $N_E = 1$), or we can have each E_i be that a specific nearby neuron has generated an action potential (in which case $N_E \geq 1$). The detector is a small high impedance metal electrode some (short) distance away from the neurons. The precise nature of the transfer of the signal from the inside of the neuron, through the membrane and intracellular fluid, to the electrode [2,3] may be complex and un-characterisable because the geometry of the electrode/neuron interface is unknown. The transduction of the signal at the electrode into a voltage may also be complex [4]. As a result we do not know the precise form that $d_k(t)$ takes for each $e_i(j)$. What we are interested in generally is the time of the spike $t(e_i(j))$, since this is what the downstream spike train analysis is based on.

However complex the transduction process may be, we do have a fairly clear idea of the nature of $s(e_i(j))$ from intracellular recordings. Further, from extracellular recordings in which we are confident that the noise level is low and from the experimental work in [2], and the theoretical work in [3], we have a clear idea of the types of signal that $s(e_i(j))$ is likely to result in. This knowledge can be used to help determine how to process $d_k(t)$.

Processing here takes a standard form. Firstly *d(t)* (we drop the subscript because we are dealing with one sensor) is processed to produce *d'(t)*, in such a way as to emphasize the type of signal expected to arise from $s(e_i(j))$. For the case where we seek to find out simply whether a nearby neuron has fired, *d'(t)* is thresholded to determine the $t(e_i(j))$. There are many different forms of processing that might be used, and a variety of ways in which the threshold might be set. Often this process ("spike detection") is used prior to attempting to assign the detected spike to a specific neuron ("spike sorting"): see [5] for a discussion of this. We are interested in finding ways to compare the many different possible techniques.

In earlier work we used real recordings, and compared the effects of a number of spike detection techniques [6], particularly investigating automated threshold setting. However, this is imprecise because we do not know exactly what the correct answers should be, so that making comparisons is difficult. More recently we have developed a technique for generating realistic extracellular signals, and used this to compare spike sorting techniques [3]. The signals are generated by first deciding on the $t(e_i(j))$ then determining the shape of the $s(e_i(j))$, and using a biophysically justified model of the transfer of the signal to determine the contribution of each event to *d(t)*. In [3], spike detection used a simple thresholding technique (i.e. *d'(t)* = *d(t)*). Here, we describe using the data generation technique to compare different spike detection techniques.

Another possibility would be to take some particularly good real data (for which the ground truth is clearly visible), and add noise to it so that we have test signals for which we know the ground truth, but in which this is not clearly visible. This has been quite a common approach, certainly in neuronal signal analysis. There are, however, still two issues: firstly, is the type of system for which clean data are available restrictive, and secondly, how should the noise be generated? The answer to the first question will be system dependent. For the second question, one might either take noisy parts of (real) signals, and add some scalar times that noise to the signal, or, indeed, one might use the biophysical model purely for noise generation.

a) The spike detection techniques

A number of spike detection techniques have been compared: each technique has at least one parameter which can be varied, and these are noted below:

plain: no processing (*d'(t) = d(t)*), thresholding both positively at the median plus a number (parameter) of standard deviations and negatively at the median minus a (different) number (parameter) of standard deviations

wav: Wavelet based preprocessing, as described in [7]. The following types of wavelet (parameter) were applied: biorthogonal spline wavelets (with of order 1 for reconstruction, and 3 and 5 for decomposition), Daubechies (order 2), Symlets (order 2), and Haar, all as supplied by the MATLAB toolbox. In addition, the parameter L (see [7]) was varied. The software used was provided by Z. Nenadic.

morph: A form of one-dimensional mathematical morphology was applied to the signal [8]. The signal was turned into two positive going signals, one being the positive-going part of *d(t) - median(d(t))*, and one being minus the negative-going part of *(d(t) - median(d(t))*. Both of these were then "opened" using a simple linear structuring element whose length is a parameter [8]. The opened signal is then subtracted from the positive-going signal, and the resultant signal thresholded at some value (parameter). A related technique is used in [9].

conv: This is a template based technique in which a section of the signal found by simple thresholding was used as a template for a spike, and this template was then convolved with the signal to produce *d'(t)*. The threshold for determining spiking events is the parameter. (see [6])

sum: This uses an averaging preprocessing technique which smooths the data from neighbouring points, based on [10] to produce *d'(t)*. The threshold for determining spiking events is the parameter. (see [6])

nced: The cumulative energy in the signal is computed, and then differentiated to produce *d'(t)*. The threshold for determining spiking events is the parameter. (see [6])

neo: This is also an energy based technique, and is an implementation of the technique described in [11]. The threshold for spike detection is the parameter.

b) Making the comparison

The formal description of the problem in section II only tells us where to look for assistance in determining how to process the $d_k(t)$: namely to the form of the $s(e_i(j))$, to how these are transformed into the $d_k(t)$, and to the nature of the noise in the $d_k(t)$. If we know these precisely, we can make the processing of the $d_k(t)$ optimal. However, we do not know these precisely: instead, we only know some information about each of them. Our knowledge is either based on experience of what the signals look like, or on some underlying biophysical model. In fact, we generally use our experience of the signals to develop and refine the models of the underlying biological system. It is therefore reasonable to use a model to generate synthetic data, particularly where real data which is annotated with the ground truth is very difficult to acquire.

To compare the different techniques, we (i) create a dataset of spike (event) times $t(e_i(j))$, (ii) transform these events into the signal that would be detected, *d(t)*, (iii) add variable amounts of realistic noise (generated from correlated and uncorrelated spiking events transformed into detected signals) to this dataset, (iv) apply each technique while varying the parameters for that technique. (Where spikes were detected less than 1 ms apart, the later spike was discarded.) Each application results in a value for the penalty, which is set to be the number of missed spikes plus the number of inserted spikes. (Clearly, different weights could be applied to missing and inserted spikes.) The penalty is calculated as a percentage of the total number of events. The parameter space is searched to find those which result in the minimum penalty. Each synthetic dataset is 5 seconds long, sampled at 24 Ksamples/second, and contains data from two neurons plus noise generated from 15 uncorrelated spiking neurons. The test is repeated 50 times for each technique to give a more reliable estimate of the penalty incurred. Different data and noise is used each time, always generated from the same distribution, and providing the same signal: noise level (measured peak:peak, as discussed in [7]). Since it may well be the case that the best values for these parameters will vary with the data and the noise level, we also recorded the parameters which gave rise to the best results for each test.

IV RESULTS

The overall results are shown in figure 1. Since we traverse the parameter space for all the techniques, one might have expected that they would all have the same best performance. This is clearly not the case. The plain technique is a version of the commonly used technique in which neurophysiologists adjust thresholds to provide the best results for each dataset. This performs well at high SNR, but is overtaken at lower SNR levels by the morph, wavelet and neo techniques. The neo technique does not perform as well at high SNR, but does perform better when the SNR is poor. The conv technique performs well at high SNR, but badly at low SNR, perhaps because it fails to find appropriate templates when the SNR is poor. The wavelet technique outperforms neo, but the morph technique outperforms all the other techniques both for high and low SNRs.

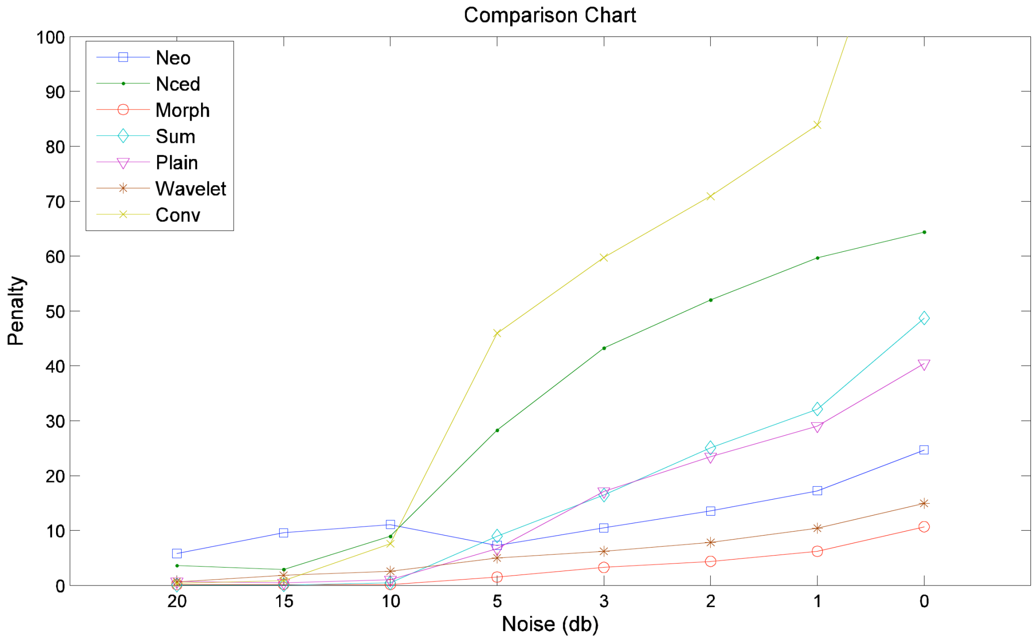

Figure 1: Comparison of the different techniques across a range of noise levels. Penalty (number of missing spikes + number inserted) shown is (total penalty/total number of spikes) * 100. Noise levels are peak to peak signal:noise ratios.

We note that for the wavelet technique, the most effective wavelet to use was the biorthogonal spline wavelet with order 1 for reconstruction and 5 for decomposition. For morph, the best structuring element length was 2.5ms.

The techniques of most interest are plain (because it is based on common usage), and wavelet and morph, since these perform best. We show the 50, 75 and 95% (percentile) penalty levels found over the 50 runs used.

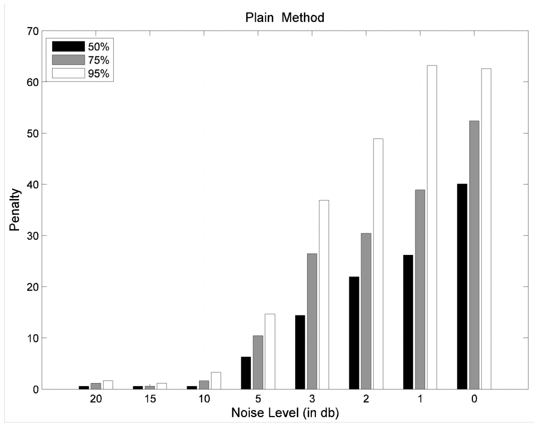

Figure 2: Error levels for plain method at different noise levels, calculated over 50 runs.

It is clear from figure 2-4 that the morph technique outperforms both other techniques at all SNRs tested. We note also that plain performs well when the SNR is good: it outperforms wavelet at 10 and 15dB SNRs. However, when the SNR worsens, wavelet outperforms plain.

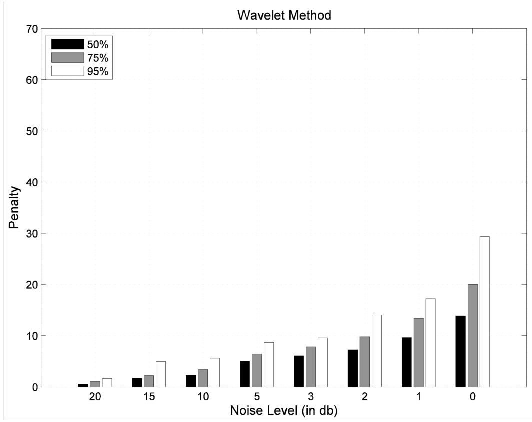

Figure 3: Error levels for wavelet method at different noise levels, calculated over 50 runs.

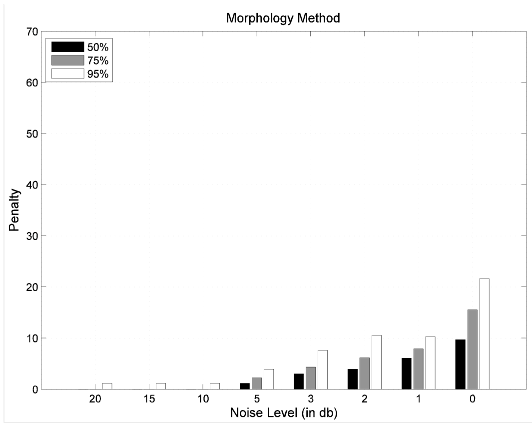

Figure 4: Error levels for morph method at different noise levels, calculated over 50 runs.

V CONCLUSIONS: HOW VALID IS THIS APPROACH?

From the experiments reported here, we have been able to make a justified comparison between a number of commonly used spike detection techniques, and we conclude that morph is the most effective technique for spike detection. We note that experiments (both with new detection techniques and other sources of data) are continuing.

How valid is this type of experiment for assessing spike detection techniques? This depends on the validity of the synthetic data: the more realistic it is (both in the sense of providing appropriate signals and noise whose statistics correspond to the actual noise in $d(t)$) the more valid is this technique. In this case, we generated the data from a realistic biophysical model, and looked at real data to attempt to determine the appropriate parameters for the biophysical model. Of course, each actual experiment will be different, depending on the precise type of electrode, and on the types of neurons used. Nonetheless, by determining the best results for each technique by searching the parameter space, we have, we believe contributed to the debate on what is the best spike detection technique.

Real neural systems do suffer from other problems not included in the biophysical model. For example, in vivo experiments often have problems with the electrode moving relative to the neurons over time, thus altering the nature of the signal received.

Considering the problems set out in section I, what can we take from the specific problem examined here? It is likely that any model of a system used for data generation will only model some of the variables in the real situation. Thus care needs to be taken that those aspects of the problem which are modelled are those that matter the most for interpretation techniques. Particularly where no ground truth data is available, we believe that biophysically based realistic data (and noise) generation provides a good test system for comparing different event detection techniques.

ACKNOWLEDGEMENTS

We acknowledge the support of the UK EPSRC (grant numbers GR/R65602/01 and EP/E002331/1 (CARMEN)). We thank Damien Coyle for useful comments.

REFERENCES

[1] The CARMEN consortium, Code analysis, repository and modelling for e-neuroscience, http://www.carmen.org.uk.

[2] G. R. Holt and C. Koch, Electrical interactions via the extracellular potential near cell bodies, Journal of Computational Neuroscience, 6(2): 169—184, 1999.

[3] L. S. Smith and N. Mtetwa, A tool for synthesizing spike trains with realistic interference, Journal of Neuroscience Methods, 159(1): 170—180, 2007.

[4] G. T. A. Kovacs, Introduction to the theory, design, and modeling of thin-film microelectrodes for neural interfaces, In D.A. Stenger and T.M. McKenna, editors, Enabling technologies for cultured neural networks, Academic Press, 1994.

[5] M. S. Lewicki, A review of methods for spike sorting: the detection and classification of neural action potentials, Neural Computation, 9:R53—R78, 1998.

[6] N. Mtetwa and L. S. Smith, Smoothing and thresholding in neuronal spike detection, Neurocomputing, 69(10-12): 1366—1370, 2006.

[7] Z. Nenadic and J. W. Burdick, Spike detection using the continuous wavelet transform, IEEE Transactions on Biomedical Engineering, 52(1): 74—87, 2005.

[8] J. Serra, Image analysis and mathematical morphology, Academic Press, 1982.

[9] E. E. Zelniker, A. P. Bradley, J. S. Castnet, H. J. Chenery, D. A Copland, and P. A. Silburn, Estimation of neuronal firing rates with the three-state biological point process model, submitted to Journal of Neuroscience Methods.

[10] M. Mariscotti, A method for automatic identification of peaks in the presence of background and its application to spectrum analysis, Nuclear Instruments and Methods, 50:309—320, 1967.

[11] K. Kim and S. Kim. Neural spike sorting under nearly 0-db signal-to-noise ratio using nonlinear energy operator and artificial neural network classifier, IEEE Transactions on Biomedical Engineering, 47(10): 1406—1411, 2000.