

# Early Warning System and Man-Made Noise

Konstantin Kislov<sup>1</sup>, Fedor Vinberg<sup>2</sup>, Aleksandr Kushnir<sup>3</sup>, Valentin Gravirov<sup>4</sup>

Institute of Earthquake Prediction Theory and Mathematical Geophysics, Moscow, Russia

<sup>1</sup>kvkislov@yandex.ru, <sup>2</sup>winberg@mitp.ru, <sup>3</sup>afkushnir@gmail.com, <sup>4</sup>gravirov@rambler.ru

**Abstract**— We present the results of development of methods and algorithms for automatic real time identification of waveforms arrival from local earthquakes in increased level of man-induced noises for the purposes of earthquake early warning.

**Keywords**— *earthquake; early warning system; earthquake detection in noise; wavelet transform; artificial neural network; earthquake disaster mitigation*

## I. INTRODUCTION

Earthquake early warning systems (EWS) rely on the capability of advanced electronic systems to process and transmit information faster than seismic waves can propagate. The relevant information arrives a few seconds to a few tens of seconds before strong ground motion begins. EWS have proven the need to reduce losses in the earthquake [1].

Most current systems rely on high density seismic networks. For example in Japan instruments are spaced every 25 km across the entire country. However if there is no dense seismic network, for example, if the purpose of early warning is only protection of a large city or a nuclear power plant, a hybrid system with "single sensors" being incorporated in the overall early warning network [2]. As the seismic network collects more data on an earthquake, the predictions will improve, but the time until shaking will decrease. The single-station approach is the fastest way to give warnings near the epicenter.

Seismic stations are generally located in remote areas, as far as possible away from any human activity. Nevertheless, road and railway traffic, heavy industry, mining and quarry activities, extensively exploited agricultural areas, and many other sources of manmade seismic noise around the seismic stations, along with natural sources can be strong noise sources. If a seismic signal is noisy and the loss of information content even at one station only, it can decrease the effectiveness of the system.

Thus, at the present stage there were following requirements to an EWS based on the "single sensor" principle:

- Identification of first earthquake waves in a noise and determination of earthquake parameters. Information must be made accessible to external systems.
- The system must be capable of operation after the waves due to the main event have arrived. The option of self-contained operation.
- The use of fast algorithms. Simple maintenance and deployment, low costs.

## II. BACKGROUND

The sufficiently complete information on an earthquake is obtainable in real time from the analysis of its first arrival based on recordings of a single seismometer. The automatic determination of the characteristics in question should be quite effective and will take about 3-5 seconds [3-7].

The noise seriously impedes the processing of seismograms. The standard detection methods of seismic waves are based on the assumption that the noise is stationary on a long enough segment of record. When a seismic record made in a megacity is to be analyzed, one has to deal with noise types having very diverse origin and characteristics. Also, the noise level is comparable with the amplitude of the signal to be detected. Thus it is impossible to make use of traditional detectors STA/LTA (Short Time Average to Long Time Average) and on others that model the signal/noise ratio.

There is generally a database of recorded events for each specific region. The database includes sample seismograms to characterize both the earthquakes and the natural noise background occurring in the region since the start of observation. All existing classification methods use this database as the training set [8].

The tests of the algorithm were carried out using seismic data from the real earthquakes, as well as the most recent large earthquake in Japan, March 11, 2011. The efficiency of system operation was verified with the help of test samples of signals belonging to certain classes, i.e., both noise and earthquake ones, and these samples were not used in training of the neural network.

The volume of the series was 2688 processes. The one part consisting of 2500 processes was used to debug the algorithm, while the testing was based on samples from the second part (188 samples).

We shall restrict the notion of a "useful signal" to the first 4 seconds in the P wave of a near large earthquake. This is a serious restriction, because of all possible signals we choose to deal with low-energy waves of earthquakes with certain parameters. First, we select epicentral distances in the range of 20 to 600 km. The damaged zone of a large earthquake is usually within this value, while at distances under 20 km the system is of little use. Secondly, the earthquake magnitude must exceed the value M6.5. An earthquake of magnitude M5 may also pose some threat. However, the damaged zone will correspond to the lower bound of the range of epicentral distance chosen. Thirdly, the focus depth is within 80 km. The rest is treated as noise.

This work is conducted with the financial support of the European Community by the International Science and Technology Center (ISTC).

### III. METHODS OF SIGNAL DETECTION

Estimates show that the time difference between P and surface waves will be about 30 seconds when the epicenter is about 200 km distant from the monitoring site. Some users choose rather more warning time and may tolerate more errors of prediction. For example, schools may prefer to get the warning sooner so children can take cover. In the case of very short pre-warning times of few seconds, it is still possible to slow down trains, to switch traffic lights to red, to close valves in gas and oil pipelines, to release a SCRAM in nuclear power plants, etc.

Our warning system is implemented to include three detection processes in parallel. Further, applying a decision rule we carry out the final earthquake detection and estimate its reliability.

#### A. An Adaptive Algorithm for Detecting the Onset Times of Low Amplitude Seismic Phases

The wavelet analysis has advantages in that it is possible to investigate not only stationary signals but also irregular series. In contrast to the standard filtering technique using the Fourier transform, the wavelet transform provides better representations of seismic wave onsets [9] and, secondly, requires a computation time for signal processing smaller by a factor of a few times.

We shall illustrate the processing taking the example of a signal shown in Fig. 1. The first 10 seconds (see Fig. 2) involve data acquisition. It was assumed for the moment that the initial segment reflects the behavior of the process before the measurements began as regards its statistical characteristics. The next portion is shifted by one second and is processed in a similar manner. These portions are little different, except for the lateral edges. Since second 74, the process begins to change. Accordingly, so does the wavelet transform (Fig. 3). Two neighboring images are compared each second. When two consecutive images are significantly different, that marks a change from one state to another. That instant of time can thus be identified as the time of onset for low-amplitude seismic phases that precede an earthquake.

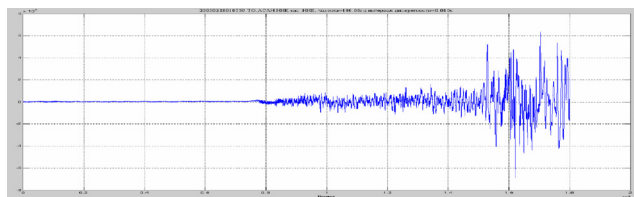


Figure 1. An example of a seismic signal.

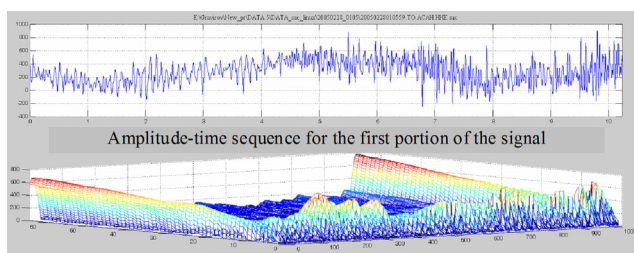


Figure 2. Coefficients of the wavelet transform (scale) as functions of time.

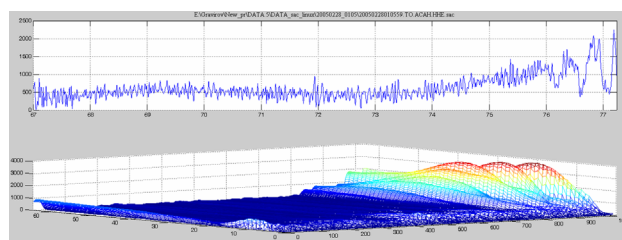


Figure 3. Variation of the parameters in the process being discussed (the arrival of first onsets of the earthquake signal).

The sequence of calculations consists of the following points:

1. Transforming a vector  $X$  with size 1:1024 elements (a segment of the original one-dimensional process lasting 10.24 s) by the wavelet transform into a 1024x64  $Y$  matrix. The matrix  $Y$  is treated as a color image.
2. The image is divided into two halves whose textures are compared. If the criterion tells us that the two images are different, then the change point is found within the later half.
3. A new  $X$  vector is formed by a shift and by adding new 100 measurements at the end; the procedure is repeated.

We use the co-occurrence matrix for comparing images. The color image  $Y$  is transformed to a gray-level image  $I$ . The co-occurrence matrix  $C$  is a 8x8 matrix and is determined by the number of gray-level values [10]. Several methods of texture characterization were proposed in the literature such as that of Haralik [11]. We choose among these parameters the most pertinent ones. We found that four parameters (Entropy, Energy, Contrast and Homogeneity) are sufficient for the most simulations that we have carried out.

The co-occurrence matrix and the above features are calculated once every second for each half of the image. For the problem in hand, we chose to use the Euclidean distance as measuring the closeness of two images. In Fig. 4, the sequence of values of the Euclidean distance is shown as a broken green line. If that sequence is treated as a sequence of random numbers in time, then one can calculate estimates of current statistical characteristics for it ( $m_{Dr}$  and  $d_{Dr}$ ), which are the mean and the mean square deviation, respectively. The mean  $m_{Dr}$  is shown by a black line. The two red lines mark the confidence interval  $m_{Dr} \pm d_{Dr}$ . The statistical characteristics are calculated in real time using only preceding values in a random time sequence  $Dr$ . Consequently, the resulting estimates are biased. The criterion to distinguish between two adjacent images is the occurrence of the current value of  $Dr$  outside the confidence interval.

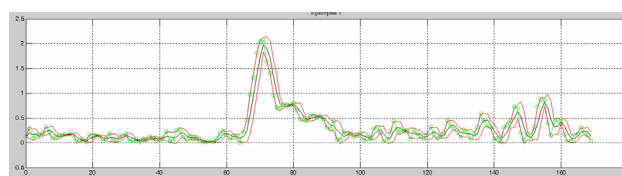


Figure 4. Coefficients of the wavelet transform (scale) as functions of time.

B. Identifying a Change Point in a Random Process and Signal Detection in a Moving Time Window

There are the following important factors that impede the construction of optimal statistical processing procedures: (a) the statistical characteristics of earthquakes signals are not known beforehand and strongly vary from event to event and (b) seismic noise is nonstationary over long intervals of time, that is, its probability characteristics vary over time (in particular, because of changes in manmade activities in the area of seismic observations).

The change of a random process  $x_t, t \in Z$  is a sudden change in its statistical properties occurring at time  $t_0$ . For example, a change-point of a stationary time series (s.t.s.)  $x_t$  would be a change of its matrix power spectral density  $F(\lambda)$ :

$$x_t = \begin{cases} \xi_t, t \leq t_0 - \text{s.t.s. with spectral density } F_1(\lambda); \\ \eta_t, t > t_0 - \text{s.t.s. with spectral density } F_2(\lambda). \end{cases}$$

Such a change arises, e.g., when the physical properties of the source that "generates" the time series  $x_t$  changed at time  $t_0$  or else another time series  $s_t$  was superposed on a time series  $\xi_t$ :  $x_t = \xi_t + s_t, t > t_0$ , the superposed series being practically independent of the original. For Gaussian stationary time series  $\xi_t$  and  $s_t$  with zero means, both these causes are equivalent (and so admit of the same mathematical treatment) [12].

The concept of the "fastest detection" is a traditional one in change-point problems [13] in which it is required to minimize the time required for correct detection of the change-point with a fixed probability of false alarms. The various options for optimality criteria of decision rules in the change-point problem as following from this concept, as well as the structure and characteristics of these rules can be found in [14].

However, in a number of cases (e.g., in signal detection) the time  $N$  allotted to decision making (whether or not a change has occurred) is strictly bounded from above, while a more rapid decision (for less than  $N$ ) is simply meaningless. In that case, a change-point is better detected by the hypothesis testing method, whether or not a signal is available based on observations in a moving window  $x_N(t) = (x_{t-N+1}, \dots, x_t)$ , and optimal tests should be used to test these hypotheses, which maximize the probability of correct detection with a limited probability of false alarms.

It sometimes is possible to construct an effective recurrence procedure to compute successive values  $g(x_N(t)), t \in N$  of the optimal test statistic; in that case, the hypothesis testing method in a moving window turns out more advantageous in terms of computer resources. This is generally the case in the absence of exact prior information on the characteristics of the process  $x_t$  after a change-point, when sequential algorithms become especially complex [15]. The procedure of detecting a change-point in a process can be reduced to determining the values of the autocovariance matrices  $R^x_t$  of  $x_t$  in a moving time window and to subsequent computation, at each detection step, of a linear or a quadratic form consisting of all the elements of these matrices [16].

For the sake of brevity, we call the detection algorithm  $\chi^2$  detector: it responds to a change (compared with the power spectral density (PSD) of the noise) in the average power and average PSD of the observations within the moving window. The noise PSD is reflected in the parameters of autoregressive approximation of the noise and correlations, which are involved in the algorithm. We note that a change in the PSD of observations is detected by optimal manner according to the asymptotic performance criteria. The statistically optimal algorithms we have developed provide sufficiently accurate solutions to the problem of decision making when we are required to recognize the signal of a local earthquake. Such an investigation of the asymptotic properties can be carried out using the methods developed in [17-19].

C. A Method for Identifying Earthquake Onsets in Man-Made Noise Using Artificial Neural Networks

One efficient tool for event detection in the presence of noise is also the artificial neural network (NN) [20, 21]. However, these powerful tools have so far been of limited use as classifiers in the identification of seismic signals in noise.

To deal with classification problems (earthquakes v. noise) we shall use a fully-connected NN with a feedforward structure, a multilayer perceptron (MLP) with a single hidden layer (see Fig. 5). The use of this architecture is frequently justified for dealing with many problems. Different neural networks can be used, since their areas of application intersect. We investigated whether it is possible to use artificial NN for the identification of seismic signal type. It has been made clear by how much neural networks are efficient as a classifier compared with the other available intellectual methods. The Neural Networks package in the MatLab system was chosen to simulate the structure of NN.

MLP neurons are organized into several layers [22]. A neuron calculates a weighted sum of its inputs and transforms this in a nonlinear manner into an output signal. The activation function  $f$  (a nonlinear transformation) must be differentiable. This requirement is met by the sigmoid function we chose to use. Since the classification is to be into two classes (earthquakes and noise), MLP will contain a single output unit that describes the classification result.

The structural diagram of the NN system for earthquake detection can be represented as follows (see Fig. 6).

We have used the simplest back propagation training method. This is to be recommended when the data set is large and there are redundant data.

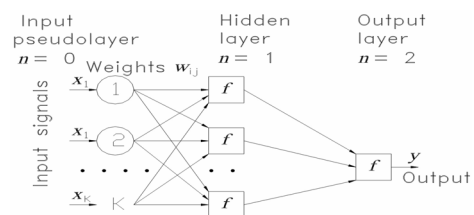


Figure 5. Structure of a multilayer perceptron with a single hidden layer and a single neuron in the output layer.

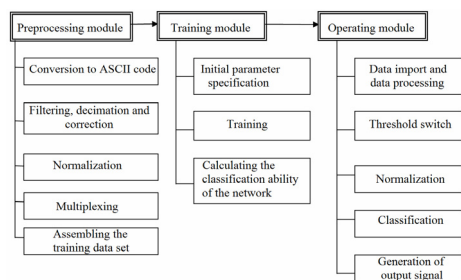


Figure 6. Structural diagram of the neural networks system for earthquake detection.

The Neural Networks package envisages calculating the performance of a trained algorithm. The process can be visualized by a training curve, which shows the decreasing error versus the number of training epochs. The criterion is the error level obtained on the training data set, while the derivative of the training curve based on the test set must remain negative. Unfortunately, this calculation gives an estimate of how the network architecture fits the training data set and is not a measure to provide an unambiguous statement of the subsequent performance of a trained network.

#### IV. CONCLUSION

A new method for automatic P onset picking for very high noise level signals has been presented. Selection criteria of methods were the following two conditions: the algorithm should give a stable result, and a minimal time for signal processing should be taken (time interval not in excess of four seconds). We used three methods in parallel. They are equally effective, but are mistaken on different examples. Our results confirm that the first earthquakes arrival can be identified for more than 97%.

Our algorithm has been tested on synthetic signals contaminated with real noise of different time-frequency characteristics and amplitudes, providing good results for all the analyzed signals, independent of the kind of noise. Each particular test was followed by constructing a plot of the process marking the detected events where there were errors of the first (failure to detect an event) and second (false alarm) kinds, as well as recording the fact of event detection during a longer (5 seconds instead of 4) interval. These data were used to find the frequencies of the errors, which are estimates of the respective probabilities. The probability of the first kind errors is 0.02, the second kind is 0.06.

If the interval during which an event is to be detected is increased to reach 5 seconds instead of the assigned 4, this improves the results very little, but significantly inflates the processing time.

We have tested the methodology against standard STA/LTA phase picks and found that this approach performs better, especially for low signal-to-noise ratios.

We already can speak about the possibility of the next step, i.e., development and production of devices for early warning of an earthquake. Both individuals and organizations can be final consumers.

#### REFERENCES

- [1] R. M. Allen, "The current status of earthquake early warning around the world", *EOS Trans. AGU*, 90(52), Fall Meet. Suppl., S21C-01, 2009.
- [2] K. V. Kislov, V. V. Gravirov, "Developing a block diagram for the earthquake warning device", *Book of Abstract CTBT S&T*, p.75, 2011.
- [3] Y.-M. Wu, H. Kanamori, "Experiment on an onsite early warning method for the Taiwan early warning system", *BSSA*, Vol. 95, N 1, pp. 347-353, February 2005.
- [4] R. M. Allen, P. Gasparini, O. Kamigaichi and M. Bose, "The status of earthquake Early Warning Around the World: An Introductory Overview", *Seism. Res. Lett.*, Vol. 80, N 5; pp. 682-693, September/October 2009.
- [5] A. B. Lockman and R. M. Allen, "Single station earthquake characterization for early warning". *BSSA*, Vol. 95, N 6, pp. 2029-2039, December 2005.
- [6] T. Odaka, K. Ashiya, S. Tsukada, S. Sato, K. Ohtake, and D. Nozaka, "A new method of quickly estimating epicentral distance and magnitude from a single seismic record." *BSSA*, Vol. 93, N 1, pp. 526-532, February 2003.
- [7] Y. M. Wu & L. Zhao, "Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning", *Geophys. Res. Lett.*, 33, L16312, 2006.
- [8] K. Mogi, "Earthquake Prediction", Academic Press, Tokyo, 1985, 166 p.
- [9] A. N. Yakovlev, "Applications of the wavelet transform for the processing of hydroacoustic signals", In *Proc. Sixth Intern. Sci. Tech. Conf. Urgent Problems in Electronic Design-2002*, Novosibirsk; Vol. 4, pp. 47-52, 2002.
- [10] O. A. Mishulina, V. Tkhei, "Correlation-Type Features in a Recognition System for Textual Images", in *MIFI-2007 scientific session*, vol. 2: Technologies of Software Systems Design. Information Technologies, Moscow: MIFI, 2007.
- [11] R. M. Haralik, "Statistical and structural approaches to textures", *Proceedings of the IEEE*, Vol. 67, N 5, pp. 786-804, 1979.
- [12] A. F. Kushnir, "Algorithms for identification of linear systems with correlated noise at input and output", *Problemy Peredachi Informatsii*, Vol. 23, N2, pp. 61-74, 1986.
- [13] A. N. Shiryaev, *Probability*. Moscow: Nauka; 574 p, 1980.
- [14] I.V. Nikiforov, "Sequential Detection of Change-Points in Time Series", Moscow: Nauka; 199 p, 1983.
- [15] A. F. Kushnir, I. V. Nikiforov, I. V. Savin, "Statistical adaptive algorithms for automatic detection of seismic signals, I. The one-dimensional case, in *Earthquake Prediction and Studies*", In the Earth's Structure (Comp. Seism. 15), Moscow: Nauka, pp. 134-143, 1983.
- [16] A. F. Kushnir, "Statistical and Computational Methods in Seismic Monitoring", *Comp. Seism.*, in press.
- [17] A. F. Kushnir, "Estimating the vector of apparent slowness in a plane wave based on data from a three-component seismic array: A statistical problem with nuisance parameters", in *Theoretical Problems in Geodynamics (Comp. Seism. 29)*, Moscow: Nauka, 1997, pp. 197-214
- [18] A. F. Kushnir, "Asymptotically effective estimation of parameters of linear systems. II. Simplified asymptotically efficient estimators" In *Numerical Modeling and Analysis of Geophysical Processes (Comp. Seism. 20)*, Moscow: Nauka, p. 148-166, 1987.
- [19] A. F. Kushnir, A. I. Pinski, "Asymptotically effective parameter estimation for linear systems. I. Local asymptotic normality", In *Mathematical Methods in Seismology and Geodynamics (Comp. Seism. 19)*, Moscow: Nauka, pp. 101-118, 1986.
- [20] S. Gentili, A. Michelini, "Automatic picking of P and S phases using a neural tree" *J Seismol.*, Vol. 10, N. 1, pp. 39-63, 2006.
- [21] K. V. Kislov, V. V. Gravirov, "Earthquake Arrival Identification in a Record with Technogenic Noise", *Seismic Instruments*, Vol. 47, N 1, pp. 66-79, 2011.
- [22] V. V. Gravirov, K. V. Kislov, T. V. Ovchinnikova, "Neural Network Method for Identification of Earthquake Phases in Increased Noise Level Conditions", *Geophysical Research Abstracts*, Vol. 12, EGU2010-2434-1, EGU General Assembly 2010, Vienna, 2010.