PREDICTION OF TSUNAMIGENIC POTENTIAL FROM SEISMOGRAM SIGNAL

[1]A S Sarika, [2]J L Lakshmi, [3]Allan. J.Wilson

[1]Assistant Professor, [2]Assistant professor, [3]Assistant Professor
Department of Electronics and Communication Engineering,
Amrita College of Engineering and Technology, Nagercoil, India

***Abstract*: Tsunami and earthquake are most dangerous natural disaster. Hence an early and accurate warning is necessary in such fields. Earthquakes are hard to predict but its resulting tsunami's can be predicted from seismogram. The available tsunami warning systems are not much effectively used in practical situations and they are time consuming since the time taken for data processing and modeling is high. In this proposed work, the characteristics of seismograms are used to distinguish the tsunamigenic earthquake from nontsunamigenic earthquakes. Feature extraction is done by applying wavelet decomposition and Shannon entropy method. Wavelet decomposition is chosen for determining the energy while Shannon method retrieves the entropy value. A classifier is defined and its performance level is tested. Accuracy up to 93% is achieved and time consumption is low when compared to other tsunami warning system. This result may contribute major part in the assessment of tsunami early warning system.**

***Keywords*: Earthquake, Tsunami, Seismogram, Signal processing.**

I. INTRODUCTION

All earthquakes in sea with high magnitude (above 7mw) will not cause tsunami. Tsunami is a very long wave caused by an underwater earthquake or underwater volcanic eruption at the epicenter point. This disturbance has an unbelievable amount of force and the energy dissipated will be high. Tsunami includes four stages. Those are generation, propagation, run up and inundation. In December 26 2004 an earthquake occurred in Indian Ocean of magnitude 9.0mw hit India, Thailand, Malaysia, Indonesia and many other Islands and killed around 275,000 people. This revealed the importance of early and accurate Tsunami warning system. Seismograms are the ground movement recorded by a seismometer and obtained as a time Vs amplitude graph. The generation and propagation of tsunami depend on the type of fault, fault rupture, fault rupture mechanism. Hence current Tsunami warning system use seismic data to determine the epicenter and magnitude level but it's impossible to calculate whether the earthquake lead to tsunami. As a result false alarms are issued.

Earthquake seismograms are chosen to analyze the characteristic of the signal since they reach seismic station at a speed of about 4km/s while tsunami travels at a rate of 500-700km/h. A widely used approach in early warning system for tsunamigenic earthquake is by receiving the information from DART (Deep ocean Assessment and Reporting of Tsunamis) buoys placed in the deep ocean. But such system involves time delay and its maintenance leads to high cost. Complementary methods have been developed to predict Tsunamigenesis of earthquake from seismic signals.

II. RELATED WORK

Past studies are done by spectral analysis for discriminating tsunamigenic behavior [1-4]. Fourier transform were used to analyze the spectral behavior of seismogram [5]. Kanamori and Kikuchi proposed slow propagation of rupture along the fault plane in W phase signal [6]. Lockwood and Kanamori proposed that W phase of tsunamigenic earthquake signal has high amplitude [7]. But W phase is found at far stations only.

Stein and okal proposed that the Andaman and Sumatra subduction zone is one of the active plate tectonic margins in the world [8]. The idea of using seismic signals to determine the Tsunamigenesis was first proposed by Ewing et al [9]. Following this many researchers used the frequency content and duration of T waves to determine the effect of earthquake. But till date no accurate predictions has been made.

III. DATA USED:

The study of tsunamigenisis of earthquake using seismic signals from different seismic station was started from late 1990's. After the devastating effect of tsunami on December 26 2004, more seismic station has been established. For understanding the physical characteristics and to develop new methodologies the high quality of earthquake seismogram signals from different stations incorporated with IRIS were used.

Global tsunamigenic and nontsunamigenic events are also used for testing and validating the methodology. The locations and magnitude of each event are also listed in the table. The entire earthquake events tabulated are recorded at 20 Hz frequency at different stations. The events used in the following study are given in the below table.

Table 1: Some of events used as data

Category	Date	Region	Location			Station Used	Mag
			Lat	Lon	Depth		
Ts	2012-04-11	Sumatra	2.33	93.06	22.9	PALK-00	8.6
	2010-10-25	Sumatra	-3.52	100.4	20.6	BTDF	7.7
	2007-09-12	Sumatra	-4.46	101.4	34.0	KAPI-00	8.5
	2005-03-28	Sumatra	2.10	97.11	30.0	PALK-00	8.7
	2004-12-26	Sumatra	3.3	95.98	30.0	DGAR-00 DGAR-10	9.3
NTs	2009-09-30	Sumatra	-0.71	99.97	81.0	CHTO-00	7.9
	2009-08-10	Andaman	14.05	92.87	33.1	CHTO PALK	7.5
	2002-11-02	Sumatra	2.98	96.11	33.0	COCO-00	7.3

Subscripts 00 and 01 are for indicating different sensors used at same site

IV.METHODOLOGY:

The main aim of this project is to find out the tsunamigenic potential in the earthquake seismogram signal. The methodology includes feature extraction technique, which is used to extract the information from the signals. SVM classifier is used for testing and training in which the data are aligned into single row as a vector. The SVM is a supervised learning machine for classification problems that transforms the attributes space into multidimensional feature space. Kernel function is used to separate dataset instances by an optimal hyper plane [10].

One of the important metrics in information theory is Shannon entropy. Entropy can be defined as the uncertainty associated with a random variable.

The Shannon entropy is used to quantify the uncertainty of the prediction of the outcome of a probabilistic event [11], being a minimum if such a prediction is exact, and, consequently, zero for deterministic events. The Shannon entropy finds wider applications in geophysics [12]. The obtained value from Shannon entropy will be positive for discrete distributions. If the outcome of an event can be predicted then the corresponding Shannon entropy value will be zero. For continuous distribution the Shannon value may be either positive or negative. Shannon entropy power has to be calculated in order to eliminate the problem arising due to the negative information measure. To differentiate Shannon entropy and Shannon power entropy different notations were used.

The Shannon entropy is calculated using the formula

$$H(x) = \sum_{i=1}^{n} p(xi) log p(xi) \quad (1)$$

p(xi) is considered as the probability density of the signal x.

The higher order statistical attributes of a time series are considered as skewness and kurtosis. Skewness can be simply defined as a measure of the symmetry of the data around the sample mean. For a normal distribution the skewness value may be either positive or negative or zero.

The skewness of a distribution is defined as

$$s = \frac{E(x-\mu)^3}{\sigma 3} \quad (2)$$

Where μ is considered mean value of the input and σ is considered as the standard deviation of the input signals. The skewness value is positive if the time series has large values in right tail. The skewness value is negative if the time series has large values in left tail.

Energy is determined using wavelet decomposition technique. Thus both the approximation and detailed energy values are obtained for the earthquake seismic signals. In high frequency band, the corresponding total energy is calculated by

$$\mathbf{Ea} = \sum_{a} |\mathbf{w}^2| \quad (3)$$

The total energy Ea at different times for high frequency is calculated for characterizing the tsunamigenisis of earthquake.

For classification SVM classifier is used. SVM classifier is an supervised learning method. Two class SVM is used. Training and testing is performed for classification. Binary classification is done. The expected output is either tsunamigenic signal or nontsunamigenic signal.

V. RESULT

The figure 1 shows the seismogram of the Sumatra earthquake signal which is indicated as an event in the table 1. The p wave train of the nontsunamigenic signal are shown in fig 2. The p wave train is considered since it holds the maximum information that can be used to characterize the seismogram signal.The Shannon entropy defines the degree of uncertainty involved in predicting output of a probabilistic event. If the entropy values increases it implies that disorder in the system also increases. The wavelet coefficients are calculated for the first few minutes of the seismograms with high frequency content are used for identifying the Tsunamigenesis.

The skewness value may be either positive or negative. Kurtosis is the measure of the taildness of the probability distribution of a real valued random variable. Kurtosis is the fourth standardized moment. From the result obtained the kurtosis value is high for nontsunamigenic signal when compared to tsunamigenic signal. The calculated Shannon entropy values and other statistical features for the input signal are shown in table 2.

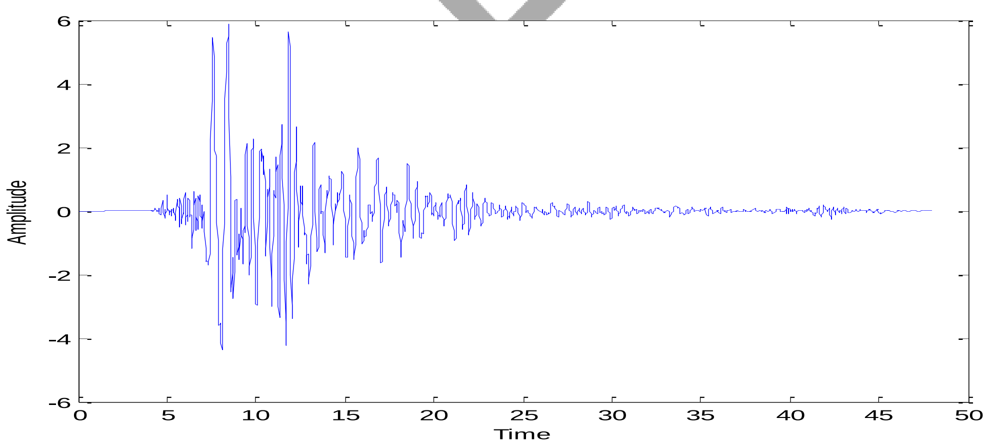

Figure 1: seismogram of Sumatra earthquake event

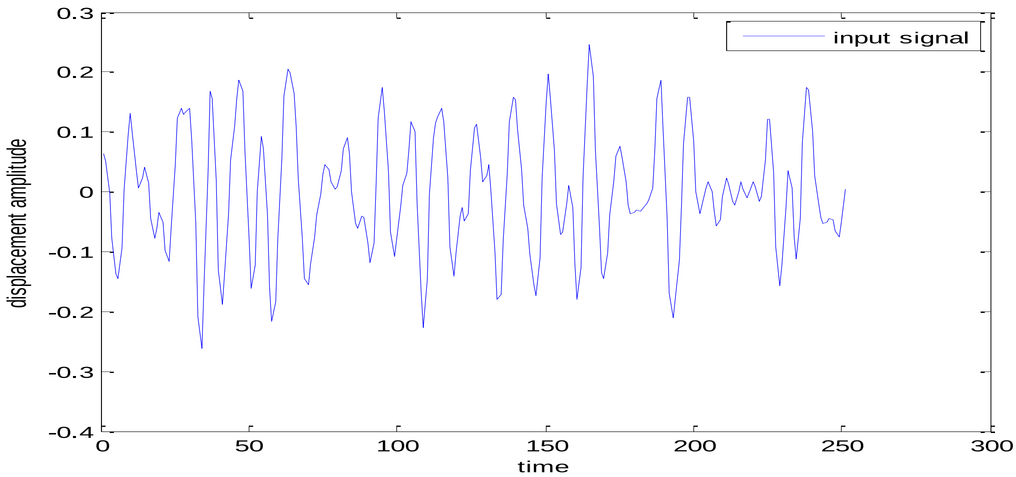

Figure 2: P wave of seismogram signal

The calculated Shannon entropy values and other statistical features for the input signal are shown in table 2

Ea	Ed				Kurt	Entropy	Skewness
99.9176	0	0	0.003	0.0821	8.8742	0	8.874244
93.0485	0	0.03	0.9069	6.0061	7.7561	0	22.2795
37.375	0.13	1.26	12.35	48.87	19.1405	0	137.258
4.9667	3.6692	81.326	9.6121	0.426	2.5072	0.191	2.634
0.1946	3.7645	84.0856	11.8356	0.1156	1.4356	0.278	2.541
1.8588	3.5075	83.2941	11.1881	0.1515	0.9368	0.424	2.729

Table 2 : Extrated feature values

Based on the methodology explained above the Shannon entropy is calculated along with its skewness and energy. The classification is done by SVM classifier using linear kernel. The two groups specified in classification are tsunamigenic as class 1 and nontsunamigenic as class 2. Training and Testing is done on the input signals. In training process out of five tsunamigenic signal two signals are used and out of thirty one non tsunamigenic signals five signals are used. The obtained result shows an evident discrimination between Tsunamigenesis and non Tsunamigenesis signal.

VI. CONCLUSION

To predict the Tsunamigenic potential from earthquake seismogram signals, the characteristics of earthquake seismograms are analyzed. The arrival of the seismic waves to the seismic stations is much faster than tsunami and this information is used for warning. The statistical features of seismograms of tsunamigenic and nontsunamigenic earthquake were investigated. The Shannon entropy alone did not show any evident pattern while the other features along with it provided an better descrimnation between tsunamigenic and nontsunamigenic event. The result shows that the above method could develop more new finding related to seismogram Time consumed for producing early warning is reduced and discrimination is effectively achieved. From the above result the accuracy level calculated is about 93%.

REFERENCES

[1] N.M. Shapiro, S.K. Singh, J. Pacheco, A fast and simple diagnostic for identifying tsunamigenic earthquakes, Geophysical Research Letters 25 (20)(1998) 3911–3914.

[2] O.G. Lockwood, H. Kanamori, Wavelet analysis of the seismograms of the 2004 Sumatra–Andaman earthquake and its application to tsunami early warning, Geochemistry Geophysics Geosystems 7 (9) (2006).

[3] A. Chamoli, V. Swaroopa Rani, K. Srivastava, D. Srinagesh, V.P. Dimri, Wavelet analysis of the seismograms for tsunami warning, Nonlinear Processes in Geophysics 17 (5) (2010) 569-574.

[4] S.K. Singh, J.F. Pacheco, M. Ordaz, R.S. Dattatrayam, G. Suresh, P.R. Baidya, Estimating tsunami potential of earthquakes in the Sumatra Andaman region based on broadband seismograms in India, Natural Hazards 64 (2) (2012) 1491–1510.

[5] A. Chamoli, A.R. Bansal, V.P. Dimri, Wavelet and rescaled range approach for the Hurst coefficient for short and long time series, Computer & Geosciences 33 (2007) 83–93

[6] Kanamori, H., Kikuchi, M., 1993. The 1992 Nicaragua earthquake: a slow earthquake associated with subducted sediments. Nature 361, 714–716.

[7] Lockwood, O.G., Kanamori H., 2006. Wavelet analysis of the seismograms of 2004 Sumatra-Andaman earthquake and its application to tsunami early warning. Geochemistry, Geophysics, Geosystems (G3) 7, Q09013.

[8] Stein, S. and Okal, E. A.: Speed and size of the Sumatra earthquake, Nature, 434, 581–582, 2005.

[9] Ewing, W.M., Woollard, G.P., Vine, A.C., Worzel, J.L., 1946. Recent results in submarine geophysics. Geological Society of America Bulletin 57, 909–934.

[10] V. Joevivek, N. Chandrasekar and Y. Srinivas Improving Seismic system for small to intermediate Earthquake Detection, International Journal of Computer Science and Security (IJCSS), Volume (4): Issue (3)

[11] C.E. Shannon, A mathematical theory of communication, Bell System Technical Journal 27 (1948) 379–423. 623–656.

[12] V.P. Dimri, Deconvolution and Inverse Theory, Elsevier Science Publishers, Amsterdam, London, New York, Tokyo, 1992.