GENERATION OF SYNTHETIC EARTHQUAKE RECORDS BY ARTIFICIAL INTELLIGENCE TECHNIQUES

M. Fadavi Amiri*,1, S. A. Soleimani Eyvari2, H. Hasanpoor3 and M. Shamekhi Amiri4
1,3School of Computer & Information Technology, Shahrood University of Technology, Shahrood, Iran
2School of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran
4School of Civil Engineering, Shahrood University of Technology, Shahrood, Iran

ABSTRACT

For seismic resistant design of critical structures, a dynamic analysis, based on either response spectrum or time history is frequently required. Due to the lack of recorded data and randomness of earthquake ground motion that might be experienced by the structure under probable future earthquakes, it is usually difficult to obtain recorded data which fit the necessary parameters (e.g. soil type, source mechanism, focal depth, etc.) well. In this paper, a new method for generating artificial earthquake accelerograms from the target earthquake spectrum is suggested based on the use of wavelet analysis and artificial neural networks. This procedure applies the learning capabilities of neural network to expand the knowledge of inverse mapping from the response spectrum to the earthquake accelerogram. At the first step, wavelet analysis is utilized to decompose earthquake accelerogram into several levels, which each of them covers a special range of frequencies. Then for every level, a neural network is trained to learn the relationship between the response spectrum and wavelet coefficients. Finally, the generated accelerogram using inverse discrete wavelet transform is obtained. In order to make earthquake signals compact in the proposed method, the multiplication sample of LPC (Linear predictor coefficients) is used. Some examples are presented to demonstrate the effectiveness of the proposed method.

Keywords: artificial spectrum; time series analysis; wavelet analysis; artificial neural network; particle swarm optimization

Received: 14 march 2017; Accepted: 20 June 2017
1. INTRODUCTION

The main reason of destruction in many buildings during earthquakes is their response to the earth’s movements or, in other words, the way buildings respond to the power released from earthquakes. The most important aspect of designing earthquake resistant structures maybe the way of calculating the power transferred from an earthquake to the structure. Nowadays, one of the most common ways of analysis and design of earthquake resistant buildings is spectral analysis, which, itself, is the basis of calculation of the force induced to the building from an earthquake, especially average buildings. A major problem associated with this method is its inability in providing time information about structural behavior and response, which is resulted from the basis of the plan’s response spectrum [1].

Due to the lack of appropriate accelerograms in many areas (either because of quantitative insufficiency or the unreliability of the registered records) and on the other hand, the benefits of employing time history analysis in many cases, it is necessary to produce some artificial accelerograms in the area. These accelerograms must, firstly, have the area’s geotechnical features and, secondly, be synchronized with a definite design spectrum [2].

In this respect, some efforts have been made to introduce methods for producing artificial accelerograms of earthquakes which are closest to the reality. There are two points that must be taken into account in different methods of producing artificial accelerograms of earthquakes:

1. Behavior of the time history of the produced accelerogram must be similar to and indiscernible from the associated real earthquake. In other words, the appearance of the produced accelerograms should not have a significant difference from recorded earthquakes.

2. The response spectrum of the produced accelerograms must, to a great extent, be in line with the design response spectrum such that it contains the special characteristics of the area.

Although engineers and seismologists are aware of the stochastic nature of earthquake, different techniques are suggested so far to estimate the movements. Engineers rely deeply on experiential methods in which frequency content and the time of movement with high amount of information are created [3].

The subject of producing artificial accelerograms has been taken into consideration by many researchers in different areas, and several papers have been published on this matter. Along with this, many efforts have been made for the development of Time-Frequency methods, considering a temporary display from the behavior of a permissible signal, one of the oldest samples of which is the Stationary-White Noise sample, which has been used to model earthquake movements [4]. The Short-Time Fourier Transform (STFT) is also one of the most prevalent methods used. The most important challenge in this method is dividing signals into small pieces (webs) with similar width in order for using Fourier analysis in all parts to gain new frequencies in all pieces [5]. In 1995, Arfken and Weber [6] proposed a new form of STFT based on Gabor Transform, which was the best way in Invert STFT. Among these models, Kanai-Tajimi is also one of the most efficient models used by researchers and engineers [7].

According to Kanai, considering the frequency content of different earthquakes records, Tajimi proposed a relationship for the function of spectral density of strong motions with
important frequency, separately. This function can be described according to stimulation of
an ideal White Noise caused on the surface of bedrock and filtered among different layers of
residual soil. After stationary methods, new methods such as Generalized Non-stationary
Kanai-Tajimi Model, have also been introduced to produce artificial accelerograms of
earthquakes based on non-stationary time dependent patterns [8]. In this model, the non-
stationary process is considered stationary by a Moving Time-Window Technique in proper
and changeable sizes [9].

The Generalized Non-stationary Kanai-Tajimi Model is used to express and simulate the
earth’s time history in two parts of time and non-stationary frequency. In this method, a
moving time-window technique is used to calculate the variable time parameters of the
model using a real earthquake record [10 and 11]. Raoufi et al. [12] used this method to
study the production of artificial accelerograms for the earthquakes in Naqan, Tabass, and
Manjil. Recent studies, however, have led to the advancement and introduction of models
that record the non-stationary spectrum of the frequency content of anaccelerogram. One of
the most famous models introduced for the production of necessary accelerograms is the
Auto Regressive Moving Average (ARMA), which is a general line model to analyze
discrete time series [13]. Auto Regressive models (AR) can build different stochastic
processes, and Moving Average (MA) models can depend the current levels of a process to
previous models of a series of white noises. Combining the terms of AR and MA, the mixed
model of ARMA is obtained which is a line model to analyze time series. Aghababaei
Mobarake et al. [14] have studied the above method in Iran earthquakes (Tabass, Manjil, and
Naqan) in two heavy and hard soil types, and in Mexico City earthquake in soft soil. Time
variable model (ARMA) can receive the non-stationary nature as well as the frequency of an
earthquake. In this method, in order to receive a continuous description of time variable
parameters, Moving Time-Window Techniques is used [15-17].

From the studies about producing artificial accelerograms compatible with response
spectrum using artificial neural network, Qabousi and Lin’s studies can be pointed out [18].
one of the other ways of producing artificial accelerograms of earthquakes is the analysis of
accelerogram wavelets for getting the features of input accelerogram synchronized with the
design spectrum. In this method, the Multi-resolution characteristic of wavelet has been used
for the betterment of instructing the system [19 and 20]. Considering the importance of the
subject in recent years, different studies have been carried out all of which were innovative
methods about using the theory of information to get artificial accelerograms synchronized
with response spectrum and other design parameters [21-24].

In references [1, 2 and 25] new methods have been introduced by the authors. The idea
lying behind these methods is the use of wavelet multiplications. In wavelet analysis, one
signal can be broken up into two signals. Each of this minor signals covers a special
frequency. If the signal is broken up into many levels, the majority of frequency contents are
covered in some special levels. Therefore, if other multiplications of wavelet are
stochastically selected, the reversed conversion of wavelet can be used to gain a signal that
is consisted of content similar to the primary frequency signal. Studying different models of
wavelet together with selecting the main wavelet and employing the above methods, a
proper accelerogram can be accessed for the considered area. In these methods the
difference factors was using different models for compacting earthquake signals and
different algorithms for improving the training mechanism of Neural Networks.
There are generally four methods of producing artificial accelerograms:
1. Stochastic Methods
2. Ray-Theory Methods
3. Hybrid Method
4. New Biologically Soft Computing Methods

Most researchers have used artificial neural networks in order to produce accelerograms of earthquakes [10]. Two important drawbacks of neural networks include getting stuck in local minima and slow instruction speed. Using the algorithm of optimization of mass ingredients, one of the most successful methods in optimizing issues, a new method is developed in this article so that the instruction problems of neural networks can be solved and the produced accelerograms can have more adaptability with the design spectrum.

In this paper firstly, some applied contexts are explained and secondly, an effective method for generation of artificial accelerogram compatible with response spectrum will be presented. In this method, artificial neural networks, wavelet analysis, an evolutionary computing algorithm and LPC coefficients are utilized. Finally, by controlling the results, the accuracy of the method is evaluated.

2. ARTIFICIAL NEURAL NETWORK (ANN)

Artificial Neural Networks which have shown to have numerous uses nowadays, have been created based on the biological model of human and animal brains and the way of acting and reacting in humans and animals. In the past, engineers made use of different tools in order to solve converse mapping problems such as statistics, regression, probabilities and optimization, systems based on knowledge. One of the methods of Artificial Intelligence to solve inverse problems is Artificial Neural Network. Methods based on Artificial Networks are approximate and not exact compared to mathematical methods, but considering the advanced mechanism of learning as well as their capability for generalization, they are becoming increasingly common.

As mentioned above, Artificial Neural Network has been formed based on the idea from humans’ neural system. An average Artificial Neural System is consisted of some layers each of which includes some small data analysis parts called Neurons, cells, units, or knots. The structure of a network includes different layers of related neurons. The first layer of each network is called the input layer, the last one is called the exit layer, and the middle layers are idiomatically called Hidden layers. Normally, the neurons of each layer are related to the side neuron layers through an oriented relationship. A sample structure of an artificial neural network has been shown in Fig. 1. The intra-neural data are connected through these connections. Each of these connections contains a special characteristic (weight) which is multiplied in a neuron to neuron transferred data. Each of these neurons operates a non-linear activation function for its final calculation. Along with the entrance node, an extra node called bias with the unit value is related to the amount of the unit or the whole neurons of the next layer. The existence of this neuron and its calculated weight is considered as a fixed amount for the input data and causes the movement of the graph in the entrance space. This amount can be the same as a fixed number in a multi-sentence theorem. Therefore, in order for this number to be eliminated, the movement function must consist of a threshold.
opposite to zero which is similar to a shift in each neuron’s entrance [26].

Figure 1. The general structure of a neural network [27]

3. WAVELET ANALYSIS

Wavelet Transform is one that extracts frequency specification of a signal from a short period, and explain show the frequency parts change as the time passes. In this conversion, a series of basic vectors are made in which the display of a signal on basis of the vectors is equivalent to frequency element of the signal. Since for each frequency resolution these basic vectors are changed, frequency elements are obtained with different resolutions. Wavelet theory gives us the opportunity of getting information about the frequency elements in non-stationary processes once in a second. Additionally, wavelet analysis shows a local behavior in contrary to Fourier analysis, which shows a global behavior. Wavelets are developed independently in different areas such as mathematics, quantum physics, electronic engineering [9].

When a discrete wavelet conversion is used, two filters are applied to the signal. One filter conquers the type of signal approximation, and the other one masters the deviations or signal approximation details. The former refers to scale filter or scale function, and the latter refers to wavelet or the wavelet filter. Fig. 2 shows the process.

Figure 2. Wavelet analysis tree [28]
In fact, applying the scale function or wavelet is the same as a moving window in the extension of the signal. Capturing local behavior is carried out by the scale function and wavelet in time. Every window is determined by the displacement of the scale function and wavelet to the right side by multiplying the main signal by the displaced scale function and wavelet [28].

The above-mentioned process is repeated so that the signal’s approximate series are obtained. Then, a new separation between details and approximate signals from signal’s approximation is made in each level. The method has been discussed in Fig. 2, which is called wavelet’s resolution tree [9].

A detail part in Wavelet analysis in level \( j \) defined as below:

\[
D_j(t) = \sum_{k \in Z} c_{D,j,k} \psi_{j,k}(t) \tag{1}
\]

In which \( Z \) is a set of positive integers, and \( c_{D,j,k} \) are Wavelet factors in level \( J \) which are defined as below:

\[
c_{D,j,k} = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt \tag{2}
\]

An approximate part in level \( j \) defined in this manner:

\[
A_j(t) = \sum_{k \in Z} c_{A,j,k} \phi_{j,k}(t). \tag{3}
\]

In which \( c_{A,j,k} \) are scale factors in level \( j \) which are defined in the as below:

\[
c_{A,j,k} = \int_{-\infty}^{\infty} f(t) \phi_{j,k}(t) dt \tag{4}
\]

Finally, the signal \( f(t) \) shown in the following manner:

\[
f(t) = A_j + \sum_{j \in J} D_j \tag{5}
\]

Often, signals can be represented well as a sum of sinusoids. However, consider a non-continuous signal with an abrupt discontinuity; this signal can still be represented as a sum of sinusoids, but requires an infinite number. Wavelets are more useful for describing these signals with discontinuities because of their time-localized behavior (both Fourier and wavelet transforms are frequency-localized, but wavelets have an additional time-localization property). Because of this, many types of signals in practice may be non-sparse in the Fourier domain, but very sparse in the wavelet domain. This is particularly useful in signal reconstruction, especially in the recently popular field of compressed sensing. (Note that the short-time Fourier transform (STFT) is also localized in time and frequency, but there are often problems with the frequency-time resolution trade-off. Wavelets are better signal representations because of multiresolution.
analysis. This motivates why wavelet transforms are now being adopted for a vast number of applications, often replacing the conventional Fourier transform [29].

4. PARTIAL SWARM OPTIMIZATION ALGORITHM (PSO)

Kennedy and Eberhart were first to propose PSO in 1995 as a non-crucial searching method for functional optimization [30]. PSO algorithm is a social searching algorithm modeled based on the social behavior of birds’ swarms. At first, this algorithm was implemented in order to discover the dominant patterns on simultaneous flight of birds and sudden drift in their direction and the optimal transformation of the swarm. In PSO, the particles are flowed in searching space. Particle transform in the search space is influenced by the experience and knowledge of both themselves and their neighbors. Therefore, the situation of other particle swarms affects the quality of searching a particle. The result of modeling this social behavior is a searching process in which the particles incline towards successful areas. The particles learn from each other and travel towards their best neighbors according to the obtained knowledge. As it is observed in Fig. 3, the foundation of PSO is based on the principle that in each moment, each particle regulates its place in the search space regarding the best place in which it has been placed so far and the best place where exists in the hole neighborhood [30]. Heuristic Particle swarm optimization and its Derivatives is known as an advanced algorithm in the field of engineering design problems [31].

- for each particle
- Initialize particle
- End For
- do
- for each particle
- calculate fitness value of the particle fp
  /*upating particle’s best fitness value so far*/
  - If fp is better than pBest
  - set current value as the new pBest
  - end For
  /*upating population’s best fitness value so far*/
  - Set gBest to the best fitness value of all particles
  - for each particle
  - calculate particle velocity
  - update particle position
- End for while maximum iterations or minimum error criteria is not attained

**Figure 3. Pseudo code of the Particle swarm optimization algorithm [30]**

5. PROPOSED METHOD

In this section, the proposed pattern for generation of artificial accelerograms compatible with special earthquake spectrum is illustrated. In this method, Wavelet transform, LPC contraction method, artificial neural networks and PSO algorithm will be used. Using biological methods such as artificial neural networks, evolutionary algorithms and other
nature-inspired algorithms is an efficient method for generation of artificial earthquake accelerograms. Simultaneous implementation of neural networks and Particle swarm optimization algorithm provides the possibility for simultaneous use of the advantages of all aforementioned biological methods. Nature inspired algorithms such as particle swarm optimization can play a significant role in optimization and external exploration of the domain of input data due to their use of extensive searching spaces. On one hand, artificial neural networks can perform internal exploration of input data with their exclusive capability in training and implementing similar behaviors. Simultaneous use of these methods can be implemented for the effective generation of artificial accelerograms compatible with the design spectrum.

In the proposed method, it is aimed to generate a relationship between wavelet-based coefficients and response spectrum of accelerograms until ultimately obtaining the most compatible accelerogram by presenting the earthquake response spectrum of the design. One of the advantages of using neural networks in this area is that since the networks benefit from high internal-searching capability, depending on the input, they will follow the taught outputs at the stage of output production. This decreases its random nature to some extent. In the proposed MLP (Multi-Layer Perceptron) artificial neural network, the input will be formed by the response spectrum and the output of accelerogram corresponding to that spectrum. After teaching the network, by exercising any shape of the target spectrum; a corresponding record can be obtained.

In order to train the networks, 186 earthquake records recorded in Iran are used herein which are divided into two soil and stone ground types according to the shear wave velocity of the recording station. The basis of this division is Iranian Code of Practice for Seismic Resistant Design of Buildings, Standard No. 2800 [32] which considers the border between soil and stone on a velocity of 375 m/s (the soil types I and II are considered as stone and the soil types III and IV are considered as soil).

The process of extracting essential characteristics and compressing the signals are addressed in the section of “signals processing”. In addition, if the goal is the compression of information (data) in a limited space, LPC\(^1\) can be used. In the processes of signal processing, Linear Predictor Coefficients Model is a proper tool for signal analysis. This model generates the coefficients by the aid of a system called IIR\(^2\) and can be better displayed with fewer points with regard to the main signal (Fig. 4). Generally, a problem in LPC can be explained as finding coefficients of the following model in which the average aggregate of squared errors of signal sample S is minimized [33].

\[
H(z) = -a(0)z^{-1} - a(1)z^{-2} - \ldots - a(m + 1)z^{-1}
\]

\[
\sum_{n=-\infty}^{\infty} x(n) + \varepsilon(n)
\]

Figure 4. Linear predictor coefficients model [33]

\(^1\)Linear predictor coefficients model
\(^2\)Infinite impulse response
In this model, finding the coefficients of a linear predictor is conducted utilizing the real X time series current values according to the previous states.

\[
\hat{x}(n) = -a(2)x(n-1) - a(3)x(n-2) - \cdots - a(p + a)x(n-p)
\]  

This method performs some signal compressions so that it would be able to have an appropriate expression from the original signal, and the recovered signals will be close to the original ones, as much as possible. Regarding the wavelets, there are different methods for data compression, since, as mentioned before, the wavelet will ultimately present an approximation and a detail coefficients of the signal. All compressional methods are able to eliminate the additional characteristics of a record and present its obvious features.

At this stage, by training several networks, the input of all of which is the response spectrum and the output is the wavelet transform, better training of the network is conducted by the composed compression method with a lower number of characteristics to the artificial neural network. In fact, the calculation of different levels of a signal is accomplished in this way. It is obvious that the sum output of the above-mentioned networks will present the desired record. The wavelet implemented in this project is a db10 wavelet, which has the simplest form of the scale function. As mentioned before, this simplicity will not affect the accuracy of the method since it is intended to achieve a relatively static pattern in output of each network here, and whatever the form of this pattern shall be, is not significantly important. In Fig. 5, the flowchart of the proposed method is shown.

![Flowchart of the proposed method](image)

To train the desired networks, 80 records with distinguished characteristics in terms of magnitude, duration, and maximum acceleration, are selected from different regions of Iran. The above-mentioned records are all separated in time intervals of 0.02 seconds, and then
they were normalized. In training each network, 60 records are selected as training series (TR) and 20 records are selected as testing series (TS). In selecting the testing data, no general selection criterion is used, and it is done randomly. It is only finally checked out that at least one pattern is followed for each soil type.

In this research, in the stage of training the networks, Particle Swarm Optimization algorithm is used regarding the required precision. After optimizing the networks using Particle Swarm Optimization algorithm, gradient reduction algorithm is used to train the network in order to achieve the weight matrix and the exact bias. In training the proposed artificial neural networks, by the use of PSO algorithm, each network is regarded as a particle. In this implementation, the training of artificial neural networks is performed simultaneously. For each input from the data set, all weights will change. In this method, each particle is a structure and an arrangement of these structures is a set of all points. The value of each particle is emerged repeatedly according to equations of this method to reach a stable state. In this way, not only are the networks involved in local extremes, but training time of the networks will also decrease severely, such that the required time for training each network will take up to 7 minutes regarding the expected precision from the algorithm. In Figs. 6 and 7, the examples of results tested by the proposed method are shown.

Figure 6. Artificial (Left-Top) and Original accelerograms (Left-Button) of Tabas earthquake and related response spectra (Right)
In this paper, a method for generation and simulation of the earthquake accelerogram has been presented. By comparing the results, it can be concluded that generated accelerogram contain most of the frequency content of the real earthquake accelerogram, and the response spectrum of the artificial accelerogram is properly adjusted to response spectrum of the real earthquake accelerogram.

By comparing the figure of artificial and real accelerograms, it can be concluded that the mentioned pattern has the ability of simulating the shape of records. Moreover, by comparing the response spectra, it is observed that the average of generated record’s response spectrum is properly adjusted to the average of generated record’s response spectrum. Therefore, it can be resulted that the presented pattern can simulate the earthquake with respect to the domain and frequency content.

In the presented method, the records are decomposed using Wavelet analysis and LPC, and then the relationship between response spectrum, LPC factors and Wavelet is approximated using MLFF neural networks. The desired accelerogram is finally obtained using reverse Wavelet transform. As a result of the simulations, it is demonstrated that the
implementation of wavelet analysis, LPC comparison method and particle swarm optimization algorithm in training artificial neural networks not only can increase the accuracy of the networks response, but also can decrease the training time.

From advantages of the above mentioned method it can be referred to:

- The non-random nature of outputs, so that the accelerograms can be considered the results of input patterns.
- The use of Wavelet analysis on extraction and recognition of record frequency features.
- The reduction of neural network size using LPC analysis.
- The use of PSO algorithm on training artificial neural networks for preventing the network from getting stuck in local minima.
- Desired flexibility, which means that with a few number of patterns, a better response for a particular input is obtained.

**REFERENCES**

3. Namiranian P. Generation of Artificial Earthquake Record using Wavelet Analysis, Artificial Neural Network and Ant Colony Algorithm, MSc Thesis, Iran University of Science and Technology, School of Civil Engineering, 2011.
17. Ebeling RM. Introduction to the Computation of Response Spectrum for Earthquake Loading (No. WES-TR/ITL-92-4), Army Engineer Waterways Experiment Station Vicksburg MS Information Technology Lab, 1992.
