# A Genetic Algorithm for the Classification of Earthquake Damages in Buildings

**Petros-Fotios Alvanitopoulos, Ioannis Andreadis and Anaxagoras Elenas**

Faculty of Engineering.

Democritus University of Thrace, Xanthi, Greece.

{palvanit, iandread}@ee.duth.gr, elenas@civil.duth.gr

**Abstract** In this paper an efficient classification system in the area of earthquake engineering is reported. The proposed method uses a set of artificial accelerograms to examine several types of damages in specific structures. With the use of seismic accelerograms, a set of twenty seismic parameters have been extracted to describe earthquakes. Previous studies based on artificial neural networks and neuro-fuzzy classification systems present satisfactory classification results in different types of earthquake damages. In this approach a genetic algorithm (GA) was used to find the optimal feature subset of the seismic parameters that minimizes the computational cost and maximizes the classification performance. Experimental results indicate that the use of the GA was able to classify the structural damages with classification rates up to 92%.

## 1 Introduction

Earthquake engineering can be defined as the branch of engineering devoted to mitigate disasters caused by earthquakes. This research area involves designing, constructing and managing earthquake-resistant structures. The main aim of the proposed approach is the automatic approach of the post-seismic status of buildings.

In high potential seismic areas there is an obvious need to have direct knowledge of the damage suffered especially in constructions of special interest (such as schools, bridges, hospitals etc.). This paper examines the structural damages in buildings. It is well known that various damage indicators through nonlinear dynamic equations can represent the damages after severe earthquakes [1, 2].

For the present classification system the proposed algorithm consists of three processing stages. First, a set of artificial accelerograms have been used to describe the earthquake ground motion. Then a set of twenty seismic parameters have been extracted from them to express the damage potential of earthquakes. In addition, a satisfactory number of damage indices have been used to estimate the earthquake damages in structures. Previous works prove that there is a correlation between the damage indices and the aforementioned intensity parameters [3, 4].

At the second stage of processing a GA has been used to reduce the number of the seismic parameters and find the subset that maximizes the classification rates. The GA starts the feature extraction process using an initial population of individuals (combination of seismic parameters) and after a specific number of generations produce an optimal single solution.

To select the optimum representation of seismic signals different kinds of classifiers have been used. Previous studies proposed artificial neural networks and artificial neuro-fuzzy inference systems for the classification of earthquake damages [5]. The classification accuracy of these systems has been used to evaluate the fitness value of the individuals.

The last part of the research was the investigation of the classification performance. The classifiers have been trained and simulated using the optimal subset of the intensity parameters. Classification results prove the effectiveness of this method.

## 2 Genetic Algorithms

A Genetic algorithm is an adaptive search and an optimization model which have been inspired from the principles of natural evolution [6]. GAs were first introduced in the early 1970s by John Holland. They are able to exploit the information from the acceptable solutions and select the optimal one.

The implementation of a GA starts with the generation of the initial population of the candidate solutions. Usually, the selection of the first population is random. GA is an iterative process which modifies the current population by selecting individuals to be parents and uses them to produce the children for the next generation.

GA moves from generation to generation and terminates until a converging criterion is met. The maximum number of generations or a threshold in fitness value, may be used to find the optimal solution.

## 3 Proposed Method

### *3.1 Seismic parameters*

Accelerograms are records of the acceleration versus time measured during an earthquake ground motion. The seismic accelerograms are a useful tool in earthquake engineering since they are able to provide an explicit description of the seismic excitation. However, due to the random sizes and shapes it is very difficult to exploit their similarities. Therefore, a set of twenty seismic parameters have been used to represent the seismic signals, whose connection to the structural damages is studied through correlation studies in the literature. In our method the GA attempts to produce better classification results.

### *3.2 Structure of the proposed method*

A GA was used to find the optimal feature set to produce the best classification accuracy of the proposed classifiers. First several subsets of seismic parameters have been examined. The classifiers have been trained according to these features. The fitness function of these subsets has been evaluated and the optimal set of seismic parameters have been extracted.

Let $L=20$ (twenty seismic parameters) be the number of feature descriptors. Assume a population of $N$ individuals. In this research a population size of $N=20$ individuals has been used. A chromosome of $L$ genes is an individual which represents the subset of seismic parameters. In the initial population $p = \{x_1, . . . ,x_N\}$ the first sample $x_1$ has all the genes equal to 1. The genes were allowed to take either values 0 or 1. A value of 1 implied that the corresponding parameter would be included in the feature subset. The seismic parameter would be excluded from the feature subset if its gene value was set to 0. In our method the negative classification accuracy of the classifiers is equal to the fitness function of the subset. We use the negative classification accuracy because the algorithm selects as elitist individuals the subsets which have the lowest fitness value. The GA was allowed to run for a maximum of 100 generations.

The GA creates three types of children to the next generation. The first type of children is the elite children. These are the best individuals in the previous generation which are guaranteed to survive to the next generation. In this approach the elite children parameter was set to 2. Besides elite children, the algorithm creates the crossover and mutation children. The crossover operation recombines genes from different individuals to produce a superior child. After the crossover the mutation step was used to search through a larger search area to find the best solution. In each generation 80% of the individuals in the population excluding the elite children were created through the crossover operation and the remaining 20% were generated through mutation. Using these parameters it is clear that for a population equal to 20 there are 2 elite children from the previous generation, 14 crossover and 4 mutation children.

### *3.3 ANN as classifier*

In the present study an ANN was used for the classification of seismic signals. This network consists of one input layer, a hidden layer of 17 neurons and one output layer. The proposed classifier is a supervised feed-forward ANN with hyperbolic tangent sigmoid activation function. The first layer presents the inputs on the network. The number of the inputs to the first layer is not fixed. All the individuals from the GA are passed through the ANN to estimate the classification accuracy of them and their fitness function. Each time the inputs are equal to the number of genes which their value is set to 1. The number of output units is fixed to four, since four are the categories of possible damages. During the training of the ANN a set of representative vector samples have been used. Then the ANN was simulated using the entire set of seismic signals to evaluate the classification performance. Each time a seismic signal was represented in ANN with the set of seismic parameters according to the individuals of the GA. During the supervised training process whenever a training vector appears, the output of the neuron, which represents the class, where the input belongs, is set to 1 and all the rest outputs are set to 0. The training algorithm for the network is the Levenberg/Marquardt (LM) and is described in [7] in more detail.

### *3.4 The neuro-fuzzy classifier*

The last method for the classification of damages in structures is a neuro-fuzzy approach. This system combines the fuzzy set theory and the ANNs. The neuro-fuzzy system has six layers. The first layer is the input layer where the inputs corespond to the subset of seismic parameters that represent the individual of the GA. The second layer implements the fuzzification process. Each neuron in the second layer represents a membership function. The key point of this step is the fuzzyfication of the inputs using four membership functions for each intensity parameter.

The next layer of the classifier consists of fuzzy rules. Each neuron corresponds to a fuzzy rule. The number of rules is related with the volume of the training samples (training accelerograms). The last three layers comprise an embedded ANN. The inputs of the embedded ANN are the firing strength of rules according to the input seismic sample and the output declares the winning class.

## 4 Results

After the nonlinear dynamic analysis of the structure, for the entire set of artificial accelerograms, three damage indices (DI), namely, the DI of Park/Ang, of Di-Pasquale/Cakmak and the maximum inter-storey drift ratio (MISDR) have been computed. According to the damage indices, the damages caused by seismic signals, were classified in four to classes. In this experiment a total set of 450 artificial accelerograms have been used. The representation of the artificial accelerograms has been studied using different subsets (individuals) of the twenty intensity parameters. Each individual is a 1x20 bit matrix.

Due to the bit string type of individuals the total number of the possible candidate solution is $2^{20}$. A GA with a population of 20 individuals was employed and executed for a maximum number of 100 generations. This means that the GA searches for the optimal feature selection and tests up to 2000 possible solutions. Using only the selection process in the GA without the crossover and mutation step it will create a negative effect on the convergence. On the other hand using mutation alone is similar to a random search. The GA has been used once time for each of the damage indices. Two types of classifiers have been used to estimate the fitness function of the GA. Tables 1 and 2 show the classification rates for the three damage indicators. Fig. 1 presents the total best individuals for the representation of seismic signals using the MISDR and DI of DiPasquale/Cakmak.

Table 1. Classification results using GA and ANN.

| | MISDR | DI of DiPasquale/Cakmak | DI of Park/Ang |
|---|---|---|---|
| Number of unknown samples | 450 | 450 | 450 |
| Number of intensity parameters | 13 | 13 | 13 |
| Number of well recognized samples | 417 | 410 | 408 |
| Total % of the recognized vectors | 92,60% | 91,10% | 90,66% |

Table 2. Classification results using Neuro-Fuzzy System.

| | MISDR | DI of DiPasquale/Cakmak | DI of Park/Ang |
|---|---|---|---|
| Number of unknown samples | 450 | 450 | 450 |
| Number of intensity parameters | 13 | 13 | 13 |
| Number of well recognized samples | 415 | 404 | 410 |
| Total % of the recognized vectors | 90,21% | 89,70% | 91,10% |

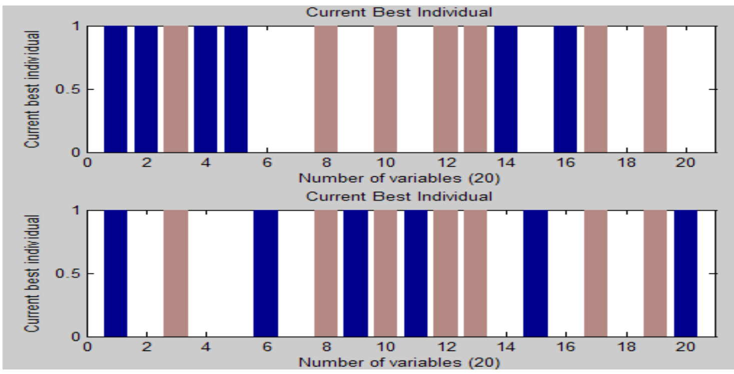


Fig. 1. Seven common Intensity Parameters.

## 5 Conclusions

GAs are a popular tool in Artificial Intelligence applications. It is well known that GAs can be used for feature extraction. With the use of GAs this approach examined the structural seismic damages in buildings. A training set of 450 artificial accelerograms with known damage effects was used to derive the parameters which are able to describe the seismic intensity. The proposed algorithm was based on a set of seismic parameters. The advantage of this approach is that the proposed algorithm was able to produce high level classification accuracy using a subset of seismic features. The number of intensity parameters was reduced from 20 to 13. The experimental results show that the classification rates are better from previous studies [8, 9]. It was demonstrated, that the algorithm developed herein, presents classification rates up to 92%. The results prove the effectiveness of the proposed algorithm. Until today, survey is performed with on-site examination by expert engineers. With the proposed technique engineers will have an additional tool which can guide them to a faster and more confident estimation of the structural adequacy of constructions.

## References

1. E. DiPasquale and A.S. Cakmak, On the relation between local and global damage indices, Technical Report NCEER-89-0034, State University of New York at Buffalo, 1989.
2. Y.J. Park and A.H.S. Ang, Mechanistic seismic damage model for reinforced concrete, Journal of Structural Engineering 111, 1985, 722-739.
3. Elenas and K. Meskouris, Correlation study Between Seismic Acceleration Parameters and Damage Indices of Structures, Engineering Structures 23, 2001, 698-704.
4. Elenas, Correlation between Seismic Acceleration Parameters and Overall Structural Damage Indices of Buildings, Soil Dynamics and Earthquake Engineering, 20, 2000, 93-100.
5. P. Alvanitopoulos, I. Andreadis and A. Elenas, A New Algorithm for the Classification of Earthquake damages in Structures, IASTED Int. Conf. on Signal Processing, Pattern Recognition and Applications, Innsbruck, Austria, February 2008, CD ROM Proceedings, Paper No. 599-062.
6. S.N. Sivanandam and S.N. Deepa, Introduction to Genetic Algorithms, Springer Verlag, Germany, January 2008.
7. Hagan, M.T., and M. Menhaj, Training feed-forward networks with the Marquardt algorithm, IEEE Transactions on Neural Networks, Vol. 5, No. 6, 1994, 989-993.
8. Tsiftzis, I. Andreadis and A. Elenas, A Fuzzy System for Seismic Signal classification, IEE Proc. Vision, Image & Signal Processing, 153, 2006, 109-114.
9. Andreadis, Y. Tsiftzis and A. Elenas, Intelligent Seismic Acceleration Signal Processing for Structural Damage Classification, IEEE Transactions on Instrumentation and Measurement, 56, 2007, 1555-1564.