УДК 519.4

G. Koziel
Lublin University of Technology, Poland
Poland, 20-618 Lublin, Nadbystrzycka 36b

Neural Nets Usage for Problems of Hiding Data in Audio Signal

Г. Козел
Люблинский технический университет, Польша
Польща, 20-618 Люблин, Надбыстрзицка 36 б

Использование нейросетей для задач сокрытия данных в аудиосигнале

Г. Козел
Люблінський технічний університет, Польща
Польща, 20-618 Люблін, Надбистрзіцька 36

Використання нейромереж для задач приховування даних у аудиосигналі

The proposal of the new steganographical method of hiding data in audio signal is shown in this article. The solution idea is based on the mechanism of linear prediction of audio signals samples valued made by neural nets. Usage of two neural nets allows for obtaining two slightly different results who are ascribed to two binary values to hide. This method allows for successful data hide in sound signal without introducing regularity inside the signal.
Key words: steganography, neural net, hiding data

В статье предложен новый метод стеганографического укрытия данных в аудиосигнале. Этот метод базируется на механизме линейного предсказания значения образца аудиосигналов с помощью нейронных сетей. Использование двух сетей позволяет получить немного разные результаты, которые приписываются двум разным двоичным значениям для укрытия. Этот метод позволяет укрыть данные в цифровом звуке без введения регулярности внутри сигнала.
Ключевые слова: стеганография, нейронные сети, сокрытие данных

В цій статті запропоновано новий метод стеганографічного укриття даних в аудіосигналі. Цей метод базується на механізмі лінійного прогнозування значення зразків аудіосигналів з використанням нейронних мереж. Використання двох мереж дозволяє отримати дещо інші результати, які приписуються до двох різних двійкових значень для приховання. Цей метод дозволяє приховати дані цифрового звуку без введення регулярності всередині сигналу.
Ключові слова: стеганографія, нейронні мережі, приховування даних

Introduction

In era of fast technology development we come across the need of often and quick communication. Repeatedly data transferred is considered to be confident, and as such it shouldn't be exposed to unauthorized persons. This means that a safe way of communication Tis needed. It can be a private communication channel. But only numerous companies own one or – because of various reasons – don't want to have one. In this situation public ways of data transmission are being chosen – in most cases, it means communicating via Internet. But every method of communication via World Wide Web is laden with the risk of

eavesdropping, or even interception and change of transmitted information. In this cases additional protection is being used – most often data is encrypted and digitally signed and only then is send to the addressee. It allows not only to protect information's confidentiality, but also allows sides to define its authenticity. In most cases this solution is sufficient enough. But the development of quantum computers calls further usage of traditional cryptography into question. Alternative methods of data protection need to be found. One of the possible solutions to this problem is to use a steganographic methods [1].

Steganography is a science dealing with inventing methods of concealing information in different carriers in such a way, that encoded content cannot be detected but a third person party [2]. This is acquired by including information into another, meaningless, or by generating new carrier, which is adapted to information concealed. The result of such an operation is a piece of data remarkably similar to the original one, which can be transmitted via public channel without gaining any attention. Additionally, to better protect information included it can be previously encrypted, which raises the security level and it allows changing data into sequence more similar to quasi-random [3-5].

Steganographic attitude to data protection problem has one more advantage – it allows parties to conceal the fact of communication itself. It can be acquired by placing prepared data in a public place, where it does not raise any suspicions. Cases of steganographically prepared data placed in public services are known. Everybody is allowed to see it, but only those, who have knowledge of hidden broadcast and have key allowing reading it are able to gain access to this data [5].

Steganography is a field of steganology – a science interested in hiding information in different carriers. An opposition to it is steganoanalisys – which interests in detection and reading of hidden messages. These two sciences complete one another, allowing to create new methods, checking their reliability and detection vulnerability. It is not possible to analyze values of gained results without checking the flaws of considered solution. It comes out from the definition of steganographic system, which says, that every system of data protection has to be considered from two points of view – of a party, which protects the data transmission channel and the party trying to read information transmitted. Schema of steganographic system is shown in Fig. 1.

This system is based on a 'game' between three parties. Two of them communicate with each other via public channel by transmitting information hidden in other carrier. Key, needed to create and read hidden content is known only to those two parties. The third is trying to detect, which part of transmitted message contains hidden data. If he succeeds, he begins to separate message form a carrier [4], [6], [7].

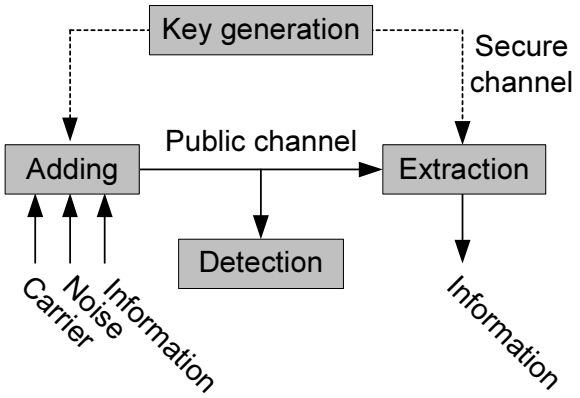

Figure 1 – Schema of steganographic system

Every kind of digital data, which can be modified without losing its basis properties, can act as a carrier. Multimedia data is especially useful because of two reasons: little sensibility to modification and huge changeability and unpredictability of signal recorded. The flaw of this solution is, that hidden data can be possibly destroyed during frequent modifications of carrier, such as compression or other form of conversion [5], [8].

1 Sound properties allowing it to be used as a hidden information carrier

Audio record is one of multimedia data type. Due to unpredictability of transient signal and its frequent changes there is a possibility of changing its record. Though, it has to be considered, that those changes cannot affect sound quality. It applies not only to frequencies audible by human being, but also to sound parameters, which can be detected as record modification. It is also essential not to allow included information to be removed while transmitting the carrier. It can occur while transmission through filter or while compression.

Every carrier has many attributes, which can be modified. Many of them are common to all digital data. Each carrier has also many specific properties. They are based on peculiarity of each data format, but also on properties of human perception.

Digital audio record is based on discrete imaging analog acoustic signal. During sound conversion into digital form inaccuracies of signal mapping appear. This phenomenon causes appearance of noise, which is saved in samples' least significant bytes. It makes hiding information by modifying this part of record, which carries only noise possible. But during this operation higher frequencies, which are removed by filters appear. Besides, the characteristics of noise saved in the record also changes, which can point out concealment of information. It also should be noticed, that this kind of record is quite undurable. It will be destroyed during any kind of conversion or format change. Including additional noise, carrying hidden information to the record is also frequently used. It is possible, mainly because it is difficult to state the original ratio of noise in the original record. The flaw of this solution is a loss of sound quality and easiness of filtering additional data while reducing noise ratio [3], [8], [9].

Usage of frequency band creates a variety of possibilities. Including additional data into record by saving it in inaudible frequency range, which is further added to container is used quite often. Such a recording is not very durable and is easy to detect. Inaudible frequency bands are often cut out from the record during process of compression. Similar operation can be performed while filtering. To avoid those inconveniences a frequency band can be filtered out of original, modified and afterwards, insert it again into the record. It will improve immunity of hidden action for removal during compression and filtering. Thus it won't improve immunity either for detection or damaging information included by steganoanalitic. To improve those the method of scattered spectrum is often used. It is based on inserting hidden information into much larger spectrum than is needed to send it. It allows to transmit high power signal, because it will be scattered on many frequencies. It allows also to gain a small distance between signal and noise in every frequency range, which will hamper either detection or damaging hidden information [3], [8], [10].

Furthermore, there is a possibility of using individual properties of audio signal and imperfection of human perception. Adding an echo to a record can be used as example. It is a natural phenomenon originating while sound reflects from an obstacle. It does not affect sound characteristics significantly. Human audition is able to separate signal echo only if it appears at least 0,02 s after the sound is emitted. Echo, which appears quicker is inaudible

for a human. It creates another possibility for hiding information, which can be encoded by using variable distances between echo and signal in some range of distance. This is considered as a good method, because separation of echo is very difficult and time consuming. Furthermore, this method shows immunity to compression and record modification.

2 Linear prediction

Linear prediction is a method that allows for calculating the signal sample estimate on a base of previous samples values. Estimate is an approximation some unknown value or parameter. This method can be also used to calculate sound signal samples values. Prediction algoithm have iterative character. It means that they work in a program loop. In each step one sound signal sample value is calculated.

The prediction rank defines how many previous samples will be a base to calculate the analyzed sound signal sample value. The bigger predictions rank the smaller the error in the prediction result. Unfortunately increasing prediction rank causes the calculative complexity rise. It demands bigger calculative power to predict signal sample value or increases the operation time.

After estimating the signal sample value the real sample value can be replaced. Usually this operation is used to replace the broken samples containing interference or some cracks. It allows for improving the signal quality. In some cases, when played signal is send via network the latencies can occur. In such case it is also possible to estimate following samples values to play without any pause. Of course it demands estimating on the base of previously estimated samples. Such procedure causes quick error rise and can not work for a long time. In steganographical tasks the author of the article proposes to use prediction operation to hide additional data inside the signal.

3 Neural nets usage in linear prediction

Artificial neural net is a structure containing a group of neurons connected together with weights. Connections between neurons can have various structures. Connections scheme defines the neural net type and its possible applies. The most often used structure in sound signal repairing is linear neural net. In this structure neurons are organized in layers. Each layer can contain various numbers of neurons. Signals in such organized neural net are send from one layer to the following layer. It means that signals flow only in one direction. There are no loops to send signal to the previous layer neither connections between neurons in the same layer [11]. The scheme of the linear neural net having three neurons in each of three layers is presented in the fig. 2.

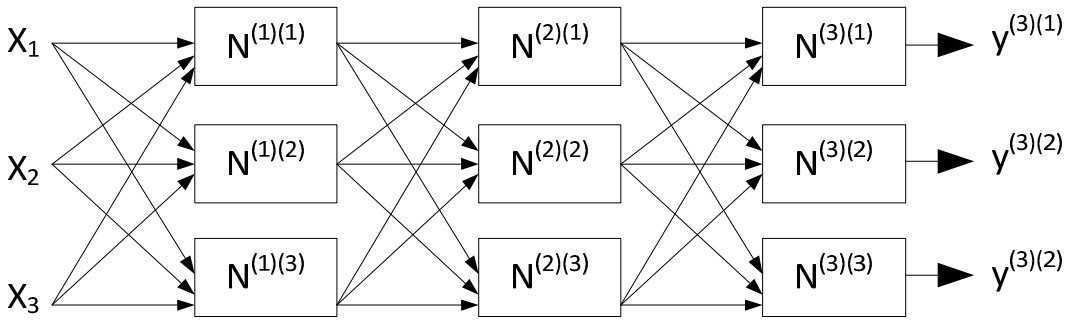

Figure 2 – Linear neural net structure scheme

Connections in linear neural net are organized in the following way: each neuron from the n-th layer is connected with all neurons from the n+1 layer. The value from the n-th layer neuron output is send to the inputs of all neurons in the n+1 layer of the neural net. Of course even one neuron can function as a simple neural net but it limits possibility of this neural net to classify objects of two linearly separable collections. Realizing prediction demands resolving much more complicated problems. To ensure such a possibility it is necessary to create network having more than one layer of neurons. The best possibility offers the natural neural nets, where a great number of neurons exists. Also connections among neurons create a very complicated structure. It is impossible to create such a big structure because it demands huge calculative power. Moreover the learning process length rises proportionally to the network size growth [12].

In real solutions a linear neural net is used. The number of neuron layers is usually reduced up to three. Three layers of network allows for resolving very complicated problems. To find the optimal neural net structure to predict the value of the signal sample on the basis of previous samples the research was done. A various neural net sizes were examined. To grade the network usability in the prediction task various parameters were examined. First of them was the time necessary to train the neural net to predict the sample value [3], [4], [8], [13]. Second graded parameter was time necessary to predict all samples in the mono signal having 450 000 samples (it is about 11 seconds length CD quality recording). Third parameter to grade was the precision of prediction possible to achieve – greater precision ensures lower level of interference and better signal quality. Precision was graded as a mean difference between original samples values and values predicted by neural net for the processed samples. The experiment was conducted on about 11 seconds length wav file, probed with 44 100 Hz. File contained pop music.

To prepare the research the neural net structure was defined first. Next the learning set was prepared. It was done by building structures containing defined number of the following sound signal samples values as an input vector and the original value of the sample which will be predicted by the neural net. These sets were used to train the neural net. After training the neural net was tested by predicting the values of the sound signal samples in the part of the recording that was not included in the training set. This was defined as verifying set. Difference between original values of these samples and the values obtained with the neural net usage is shown as the neural net imprecise.

Neural net and calculations were realized in Matlab. This software was installed on the computer supported with 4GB RAM memory and Intel i5 M560 processor working with 2,67MHz frequency.

Results of the experiment are presented in the table 1. In the first column the neural net structure was presented. Following numbers mean the number of neurons in the subsequent layers of the neural net. The first layer contains only the input neurons. Their number is equal to the number of sound samples values given to the input of the neural net. The second column defines how many elements were in the learning ser. Each element contains the target sample value and the preceding samples values. The third column presents the time necessary to train the neural net. In the fourth column the time necessary to predict all sound samples values from the verifying set was placed. In the last column the mean error was placed. The mean error was calculated on the base of all elements of the verifying set according to the formula 1:

$$ME = (\sum_{t=1}^{N} (A - T)) / N, \quad (1)$$

where: A is the original sample value, T is predicted sample value, N is the number of elements in the verifying set.

The result presented in the table 1 analysis shows that two layer neural net is enough to predict sound signal samples values. Three layer neural net usage is also possible, but training of that type neural net is more difficult and not always gives satisfying results. The number of samples given to the neural net input is also important. Obtained results analysis shows that it is not worthy to predict signal samples on the base a very short signal samples values vector. The very long vector usage is not efficient too. The optimal input vector length ranges from 40 to 100 sound signal values. Not less important is to choose the appropriate number of learning vectors. If this number is too small the mean error is high. The very big number of learning vectors usage decreases a mean error a bit but significantly increases the training time. Optimal number of learning vectors ranges from 500 to a few thousands.

Table 1 – Neural nets results in the sound signal sample prediction

Neural net structure	Number of learning vectors	Training time [s]	Prediction time [s]	Mean error (ME)
10:1	1000	1,1	4,6	0,013
40:1	1000	1,37	5,3	0,008
60:1	1000	1,37	5,7	0,007
100:1	1000	1,31	7,8	0,01
1000:1	1000	5,4	60,4	0,31
60:3:1	1000	2,65	6,9	0,008
60:6:1	1000	3,7	6,9	0,016
60:16:1	1000	18,6	7,3	0,024
60:1	100	0,57	38,9	0,1
60:1	500	0,83	9,4	0,0075
60:1	2000	1,68	4,5	0,0074
60:1	10000	16,07	3,3	0,0062

Good results of predicting the sound signal sample value allow to use the neural net to calculate the signal sample value. The graph of obtained values of samples and error calculated as a difference between calculated value and the original one is presented in a fig. 3.

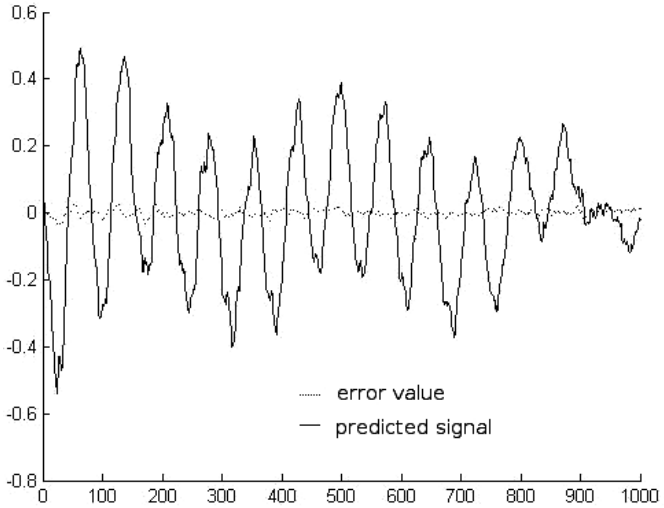

Figure 3 – Sound signal samples values predicted by a neural net and the error value

4 Neural nets as a tool in steganography

As it was presented in the previous chapter neural net can be successfully used in linear prediction. Although this tool is really effective, introduces a certain level of interference. This phenomena can be exploited to the steganography purposes. If we use two independent neural nets to predict the same signal sample, it should result in two various results. Of course this difference will not be significant. It should be really small, but each difference allows us to use it to hide additional secret data. To check the difference between results obtained with two independent neural nets the following experiment was conducted. Two neural nets having the same structure were implemented. Each of them was initialized with random weight values. Next the training process was conducted with the same training set usage. Both neural nets where then used to predict samples values in the same signal. Difference in their results was calculated and analyzed. Results obtained are presented in figure 4.

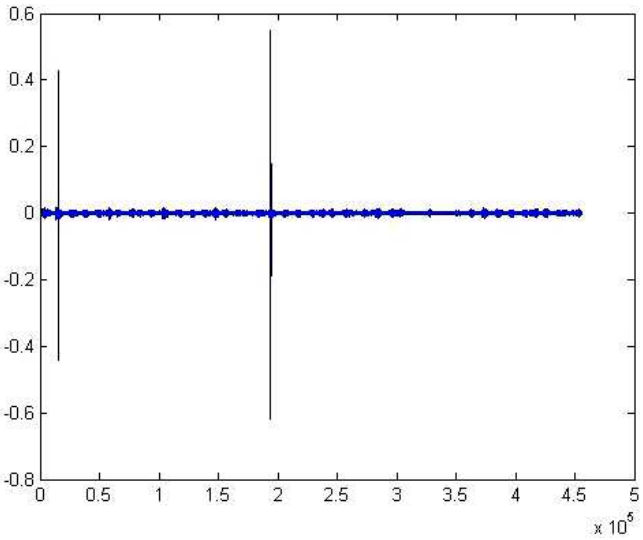

Figure 4 – Difference between values predicted by two neural nets.

Analysis of the results presented in the figure 4 shows that there are small differences between results obtained with two independent neural nets. Of course big errors occurs too. We can easily notice two peaks on the graph. In these samples error was really high. To precisely define the difference between values given by these two neural nets the obtained results were classified to several ranges. To define the desired ranges it is necessary to calculate the precision of its quantization. The signal sample value ranges from -1 to +1. If we use 16 bit to keep one sample value it allows us to write 65536 various values. If we divide the maximum possible range of sample value which is equal to two by 65536 we obtain the value difference between subsequent values possible to write. This difference is equal to 1.2095e-004. Further in the article that value will be marked as diff.

The most important thing is to define how many predicted values have difference smaller than diff. So the first range size is diff. The following ranges are a multiplicity of the diff. The results obtained are presented in the table 2.

Analysis of the results presented in the table 2 shows that over 54% of results given by two independent neural nets have very close values. Their difference is smaller than diff. It means that most of them will be quantized as the same value. From the steganography point of view it is disadvantage. It would be better if most of results would have difference bigger than diff. Of course it is possible to obtain such property. It is enough to increase the acceptable error level during training. It will result in worse neural net training result and significantly

varies results given by independent neural nets. But such procedure also results in less precise prediction what introduces unnecessary interference. To avoid these disadvantages and use projected method the additional test was conducted. It was calculated how many results returned by independent neural nets have the same value. It was found that there are no identical values. The probability of identical results occurrence is close to zero. It gives us a possibility of similar results given by neural nets usage. If we know results given by two neural nets and their difference is smaller than diff, it is possible to round the bigger of them up and round the smaller down. It will make it possible to do not loose that difference during write process.

Table 2 – Differences between values predicted by two independent neural nets

Range r	r< diff	diff <r r<diff*10	10*diff <r r<diff*50	50*diff <r r<diff*100	100*diff<r r<diff*500	500*diff <r
Number of results	249176	81320	120265	4199	979	62
Percentage of results	54,64%	17,83%	26,37%	0,92%	0,21%	0,00%

If we have a possibility to write results given by two neural nets and it is possible to determine which result was given by defined neural net we can build steganographic method to effectively hide secret data. The method is based on an ascription binary values zero and one to two independent neural nets. The neural net which is ascribed to the binary zero will be named “zero net”, the other one the “one net”.

In proposed method the one bit of data can be hidden in one sound signal sample. First it is necessary to determine the sample to hide data. After that the value of this sample is predicted by two neural nets. If the binary zero value is hidden, the value given by “zero net” is inserted apart of the original one. If the binary one value is hidden, the value given by “one net” is inserted apart of the original one. In this way the binary value is coded inside the signal.

Samples to hide bits of secret data are chosen by an algorithm. In the conducted experiment it was a regular distance – once per defined number of samples. Of course it is possible to use any function to determine samples to modify. It is very important to predict samples on the base of signal with all previously modified samples values inserted into the signal because if distance between modified samples is smaller than number of samples used to predict the value of chosen sample the inserted modification influences the value predicted by neural net. This phenomenon can interfere the hidden data extraction process. So it is necessary to insert one change into the signal before the next sample value prediction.

The strength of the method depends on the steganographic key. The key contains such elements as:

- Two neural nets structure and all data necessary to run them,
- Number of samples used to predict sound signal sample value,
- Assignment each of the neural nets to binary value,
- Distance between following samples used to hide secret data,

To proper detect hidden signal we do not have to have original signal. It is enough to analyze the modified one. Also necessary is the knowledge about hiding algorithm and used steganographic key.

Conclusion

This work shows the reader the concept of steganography and shows possibilities of usage of this science for secret communication protection purposes. It is just a rapidly

developing branch of science. It is caused mainly by the need for new ways of protecting information. Besides, steganography, opposite to cryptography allows parties to stay anonymous, which becomes very desired feature.

There are many steganographical methods based on audio record. But most of them focus on modifying parameters analyzed during the filtration of the signal. It creates the necessity of committing huge changes, which cannot be completely removed from the carrier. Scattering the record technique is also used, as it decreases the risk of damaging the record. But it results also in decreasing the steganographic capacity. Author considers that using the parameters of sound, which haven't been used so far, will benefit in the field of steganography.

Neural nets usage allows for obtaining the steganographical capacity of one bit per one modified sample. The total capacity of the method depends on the number of used samples. In the case of usage one per ten signal samples it is possible to obtain the steganographic capacity at the level of 4410 bits per one second in each channel of multichannel record.

Literature

1. Chun-Shien Lu. Multimedia Security / Chun-Shien Lu. – Idea Group, 2005
2. Fiok J. Proovably secure steganography / Fiok J. – 2003.
3. Garbarczuk W. Information protecting basis / W. Garbarczuk, A.Świć. – Lublin, 2005.
4. Sencar H. T. Data Hiding Fundamentals and Applications / H.T. Sencar, M. Ramkumar, A.N. Akansu. – Elsevier, 2004
5. Korbicz J. Artificial neural nets. Basis and applications / J. Korbicz, A. Obuchowicz, D. Uciński. – Akademicka Oficyna Wydawnicza PLJ, Warsaw, 1994.
6. Cachin C. An Information-Theoretic Model of Steganography / Cachin C. – 1998.
7. Zhang K. Information Security / K. Zhang, Y. Zheng // 7th International Conference, (ISC 2004, Palo Alto, CA, USA, September 27 – 29). – Springer-Verlag Berlin 2004
8. Wayner P. Disappearing cryptography / P. Wayner // Elsevier Science. – San Francisco, 2002.
9. Johnson N.F. Information Hiding Steganography and Watermarking-Attacks and Countermeasures / N. F. Johnson, Z. Duric, S. Jajodia. – Kluwer Academic Publisher, 2001
10. Katzenbeisser S. Information Hiding Techniques for Steganography and Digital Watermarking / S. Katzenbeisser, Fabien, A.P. Petitcolas. – Artech House, 2000.
11. Żurada J. Artificial neural nets / J. Żurada, M. Barski, W. Jędruch. – PWN, Warsaw, 1996.
12. Czyżewski A. Some methods for detection and interpolation of impulsive distortions in old audio recordings / Czyżewski A. – New York, 1995.
13. Czyżewski A. Digital Sound / Czyżewski A. – EXIT Warsaw, 2001.

RESUME

G. Koziel

Neural Nets Usage in Secret Communication

In the given article the proposal of the new steganographical method of hiding data in audio signal is shown.

The solution idea is based on the mechanism of linear prediction of audio signals samples valued made by neural nets. Usage of two neural nets allows for obtaining two slightly different results who are ascribed to two binary values to hide.

This method allows for successful data hide in sound signal without introducing regularity inside the signal.

The paper is received by the edition 08.04.2013.