Reconstruction of Missing Physiological Signals Using Artificial Neural Networks

AM Sullivan, H Xia, JC McBride, X Zhao

Department of Mechanical, Aerospace, and Biomedical Engineering
University of Tennessee, Knoxville, TN, USA

Abstract

Real-time monitoring of vital physiological signals is of significant clinical relevance. Disruptions in the signals are frequently encountered and make it difficult for precise diagnosis. Thus, the ability to accurately predict/recover the lost signals could greatly impact medical research and application. In response to the PhysioNet/CinC Challenge 2010: Mind the Gap, we develop an algorithm based on artificial neural networks to predict the missing signals in one channel using the measurements in other channels. An artificial neural network model is created for each record, which consists of 6, 7, or 8 signals acquired from bedside ICU patient monitors. We first train the network using data from the beginning 9.5 minutes of the record. Then, we reconstruct the missing data in the subsequent 30 seconds for a specific channel. A few techniques are utilized to improve the performance of the model. Principal component analysis is used to reduce complexity and computational cost. Noisy signals are smoothed using a wavelet-based de-noising algorithm before training and testing. We explored three different neural networks: focused time-delayed neural network, distributed time-delayed neural network, and nonlinear autoregressive network with exogenous inputs. The focused time-delayed neural network is more computationally efficient while the other two networks provide slightly more precise predictions. For highly correlated data sets, all three networks are able to produce accurate predictions; however, predictions of chaotic and highly noisy data sets are less satisfactory.

Figure 1. Clinical setting of an intensive care unit; Figure from [1].

Figure 2. Example showing sudden missing of physiological information. Here, the three channels are respiration, plethysmograms, and ECG. Figure obtained from [2].

1. Introduction

Real time monitoring of physiological signals, including ECG, blood pressure, respiration, plethysmograms, etc., is an essential tool in hospitals and institutes, especially in intensive care. Figure 1 shows the clinical setting of an intensive care unit (ICU).

There are a number of factors that interrupt data signals when patients are in clinical intensive care. Disruptions such as standing, walking, or eating meals could cause sensors to disconnect or become removed. Figure 2 shows one example, where one of the three channels is suddenly lost. Also, signals such as respiration can be consciously controlled while signals such as blood pressure can be quickly influenced by external triggers. Therefore, it is critical that an algorithm created to predict physiological signals address the fact that some of the features from the signals may not be representative of pertinent data needed to diagnose.

This work is in response to the 2010 Physionet Chal-

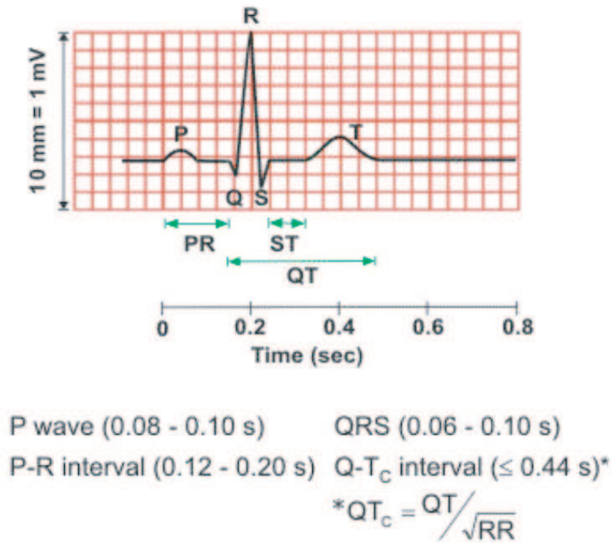

Figure 3. Example of typical ECG signal; Figure from [3].

lenge: Mind the Gap [2]. The purpose of the challenge is to construct models to predict the final 30 seconds of a channel from a 6-8 channel data set consisting of 10 minutes of data. The data is sampled at 125 Hz for the full ten minutes for each record. Each data set is representative of a different patient. The data is supplied from real patients receiving intensive care at hospitals. Each of the 6-8 channels include data from various physiological testing devices. Most of the signals consist of ECG I, II, and V data along with respiration and plethysmograph. Other signals that are presented less frequently because the sensors require invasive procedures include central venous pressure, arterial pressure and pulmonary arterial pressure. The ECG signals are all monitoring changes in the electrical activity from the heart using skin sensors placed in different positions on the body. Respiration is the measurement of patient breathing and the plethysmograph (pleth) is the measurement of blood flow through an optical sensor that is typically placed on the finger.

The normal electrocardiogram (ECG) is composed of a P wave, a QRS complex, and a T wave [3]; see Figure 3. The P wave and the components of QRS are respectively caused by the depolarization of the atria and the ventricles, while the T wave corresponds to ventricular repolarization. Since the ECG provides abundant information about the heart's activities, almost all severe abnormalities of the heart muscle can be detected through analyzing the ECG [4]. However abnormalities such as arrhythmias may appear intermittently. Thus real time monitoring is essential.

2. Methodology

Artificial neural networks are mathematical models that are designed to "think" like the human brain. A network model consists of interconnected group of artificial neurons that emulate the function of neuron cells. Neural network models can learn from data by certain training rules. Neural networks have now been widely used for modeling, prediction, classification, and control [5].

We attempt this challenge problem using several different neural networks, including focused time-delay neural network (FTD), distributed time-delay neural network (DTD), nonlinear autoregressive network with exogenous inputs (NARX) network, and feedforward network. In each of these models, we used all of the signals provided for each record to make a prediction of the missing data signal. The complete signals were considered training sets and the 30-second record of missing data were considered test sets. Because of the highly nonlinear and chaotic behavior of the signals, the training set was reduced from all 71,250 data points to various smaller intervals. Training on a smaller interval reduced the need for the network to train chaotic behaviors that may not have been present in the test set. While a smaller training set was used, each of the complete data signals from each record was used as an input in the network. The missing channel was used as an output and tested against the prediction for training purposes. In the training set, we used mean squared error (MSE) to determine the relative error for each test.

Neural network models have many parameters that must be formulated when creating a model. While some can be selected to match the model type (for example, the linear or tansig transfer function can be selected when a data set does not appear exponentially correlated), others must be selected carefully and fully tested. In our model, we tried to obtain optimality by testing training set range, delay-vectors, number of neurons, and computational efficiency.

Two scores are used by physionet to evaluate the prediction results. Denote the target signal by x_i and the predicted signal by y_i, where $i = 1, 2, \ldots, N$. Then score 1 is defined as follows

$$Q_1 = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2}{\frac{1}{N}\sum_{i=1}^{N}x_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N}x_i\right)^2}. \qquad (1)$$

Note that score 1 can be interpreted as $Q_1 = 1 - mse(x,y)/var(x)$, where mse stands for mean squared error and var stands for variance. The score 2 is defined as

$$Q_2 = \frac{\frac{1}{N}\sum_{i=1}^{N} x_i y_i - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} y_i\right)}{\sqrt{\left(\sum_{i=1}^{N}\frac{x_i^2}{N} - \left(\sum_{i=1}^{N}\frac{x_i}{N}\right)^2\right)\left(\sum_{i=1}^{N}\frac{y_i^2}{N} - \left(\sum_{i=1}^{N}\frac{y_i}{N}\right)^2\right)}}. \qquad (2)$$

Note that score 2 is essentially the correlation coefficient between the target and the predicted signals.

3. Results

We carry out extensive numerical studies using the neural network toolbox in Matlab [6]. Preliminary studies show that, using the same training data and training mechanism, the FTD network is faster and more accurate than the DTD and NARX network. In the following, we present only results using FTD networks. In all simulations, we use a network with 2 hidden layers, where the first hidden layer has 10 neurons and the second hidden layer has 1 neuron. We use *tansig* and *purelin* as the transfer functions [6].

Physionet provides 3 sets of data for the challenge problem. Set A is the only complete data set, where the missing signals of each record are provided, making it convenient to evaluate the results of prediction. The following results are based on records in set A. However, the final scores of all sets are given in Table 1.

The FTD network is robust and can provide very good predictions even for seemingly irregular time series. For example, results of a record (a02) from set A are shown in Figure 4. The scores of the prediction are $Q_1 = 0.92$ and $Q_2 = 0.96$.

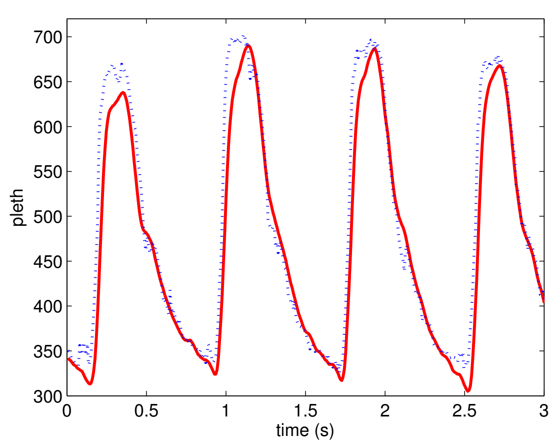

Figure 4. Comparison between predicted (dotted) and actual (solid) signals for a selected record (a02).

To explore the influences of delay and length of training data, we study different combinations of two delay vectors and three training length. We denote the two delay vectors by d_1 and d_2, respectively, and the three training lengths by l_1, l_2, and l_3, respectively. Both d_1 and d_2 contain 11 elements, where the components of d_1 continuously increase from 0 to 10 and d_2 is chosen to be $[0, 1, 3, 4, 8, 16, 32, 64, 128]$. Thus delay 1 can be regarded as a short-term delay whereas delay 2 is a mix of short-term and long-term delays. The three training lengths are chosen to be $l_1 = 10000$, $l_2 = 15000$, and $l_3 = 20000$, respectively. The training data are immediately prior the missing signal. We study the effects of the 6 different parameter sets: (d_1, l_1), (d_2, l_1), (d_1, l_2), (d_2, l_2), (d_1, l_3), (d_2, l_3). Figure 5 shows the boxplots of the 100 records of set A using the 6 different parameter sets. It is clear that, for the same training length, the mean score of d_2 is larger than that of d_1 whereas for the same delay vector, the mean score of l_3 is the largest.

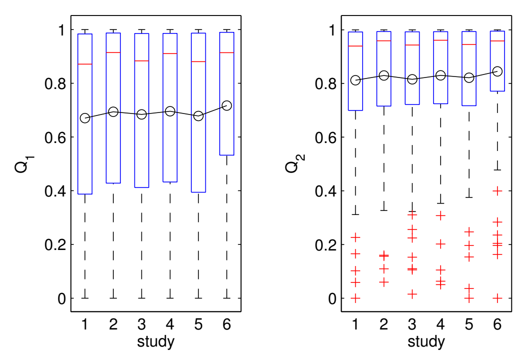

Figure 5. Box plots of the scores for 6 different studies; see text for details about the parameters. Black circles show the mean values of 100 records of set A.

We also study the effect of epochs using delay vector d_2 and training length l_1. Figure 6 shows the mean Q_1 and Q_2 scores under various values of epochs. It is interesting to notice that the dependence of the mean scores on the training epochs exhibits an oscillatory pattern. Moreover, the optimal scores are obtained at around 30 epochs.

Similarly, we investigate the influences of different training lengths when other parameters are kept constant; see Figure 7. Here, we choose the delay vector to be d_2 and the epochs to be 20. Then, the length of training data is varied from 2000 to 50000. It is noticed that best mean scores in Q_1 and Q_2 are obtained when the training length is around 25000 for the current delay vector and epochs.

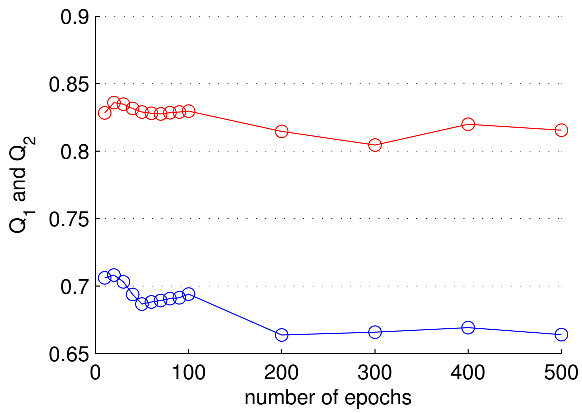

Figure 6. Mean scores using epochs: Q_1 (bottom) and Q_2 (top).

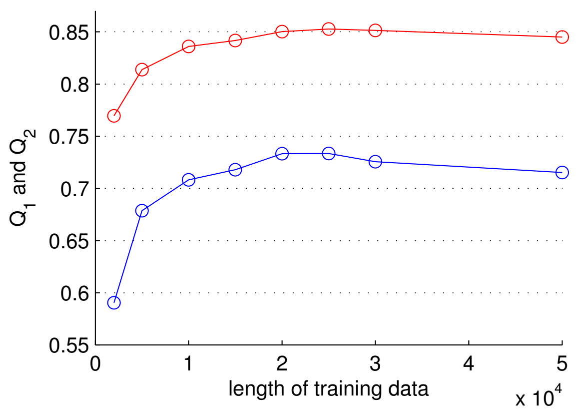

Figure 7. Mean scores using different training lengths: Q_1 (bottom) and Q_2 (top).

4. Discussion and Conclusions

We have explored the data reconstruction problem using different types of artificial neural networks. Empirical studies show that focused time delay networks are more robust, efficient, and accurate when compared to other networks. In order to achieve optimal prediction results, we have carried out extensive numerical simulations to explore the influences of various network parameters, including training length, delay vector, and training epochs. Due to limitation in time and computational resources, it is difficult to completely explore the feasible parameter space. An optimal searching approach such as genetic algorithm may be helpful. In the current studies, we have kept the layers and neurons of the network fixed. It will be interesting to investigate the influences of the different network structures.

We have observed a few useful techniques to improve the accuracy of prediction. First, better results can be obtained by a mix of short- and long-term delay outperforms short-term delays of the same length. This may be due to the "memory" effect of physiological systems. Second, we have designed a feedback algorithm that can improve the quality of prediction. Finally, we have found that averaging the predicted results from multiple iterations can further improve the accuracy.

As of the deadline on September 1st, 2010, the best scores of our tests for each data set are shown in Table 1. Both Q_1 and Q_2 values rank second among all participants of the challenge.

Table 1. Scores for different data sets.

	set A	set B	set C
aggregate Q_1	80.25	83.42	81.33
aggregate Q_2	89.15	89.78	89.67

Acknowledgements

We would like to thank the financial support of the NSF under grant No. CMMI-0845753. The authors are grateful to Dr. J. Wesley Hines for his insightful discussions. The authors would like to thank Mr. Gerald Ragghianti for technical support on cluster computing.

References

[1] http://en.wikipedia.org/wiki/Intensive-care-unit

[2] http://www.physionet.org

[3] Richard E. Klabunde, Cardiovascular Physiology Concepts, http://www.cvphysiology.com/Arrhythmias/A009.htm

[4] A.C. Guyton and J.E. Hall, *Textbook of medical physiology*, W.B. Saunders Company, ED-10, August 2000

[5] S. Haykin, *Neural Networks: A comprehensive Foundation*, New Jersey, Prentice Hall, 1999.

[6] http://www.mathworks.com/

Address for correspondence:

Xiaopeng Zhao
Department of Mechanical, Aerospace, and Biomedical Engineering
University of Tennessee
Knoxville, TN 37932
xzhao9@utk.edu