Identification of Individual Sleep Apnea Events from the ECG Using Neural Networks and a Dynamic Markovian State Model

W Bystricky[1], A Safer[2]

[1]TSD, Künzelsau, Germany
[2]Abbott, Ludwigshafen, Germany

Abstract

A two step method is introduced that uses heart beat related information from the ECG to identify individual sleep apnea events. In the first step a set of features based on the RR interval, QRS dynamic, T-wave morphology and high frequency noise is used as input of a neural network to assign each heart beat to one of four consecutive apnea states. In the second step the output of the neural network is proceeded by a dynamic Markovian state model which determines the most probable state trajectory. The 35 records of the learning set from challenge 2000 are used for model adjustment. For that purpose the authors examined about each 10th minute of the learning records and manually annotated the location of apnea states.

Although only a part of the information is used we find 89% agreement with the original annotations within the learning set and 84% agreement if applied to the 35 records of the test set from challenge 2000.

1. Introduction

Detection of sleep apnea from the ECG has been studied by a number of groups during the Computers in Cardiology challenge 2000 [2]. One objective of this challenge was to quantify the apnea stress by the number of minutes spent with disordered breathing. The surprisingly good results and methods are summarized in [5]. Here we want to get rid of the 1 minute restriction and identify onset and length of each individual apnea. We do this by training a neural network to recognize the apnea state for each heart beat. As this information is not given in the challenge data we had to identify the exact position of a larger set of apnea by ourselves. We furthermore differentiate between three states of an apnea to account for its progression in time.

We use a set of heart beat related features as input for the neural network which seem to be correlated with the dynamic behavior of an apnea. The primary feature surely is the RR interval, but there are other characteristics which show a clear relation to apnea events. In the challenge 2000 the heart beat related parameters R, T-wave and S amplitude as well as a so-called pulse energy and R duration have been investigated in various methods [5]. Here we use a set of parameters which contain all this information plus another feature which we call jitter. It is high frequency noise which often can be observed during the respiration activity that terminates the cessation of breath.

The output of the neural network are probabilities of the apnea states. We apply a dynamic Markovian state model to that time series of probabilities and identify the most probable state trajectory. This resulting state trajectory is compared with the original annotations of the challenge. We further compare the contribution of the individual features to the prediction quality.

2. Methods

2.1. ECG pre-analysis

Heart beats are detected in all 70 ECGs using a standard QRS detector. Non-normal events are manually labeled and excluded from further calculations. For each normal heart beat the following set of event parameters is calculated:

- RR: Time interval to preceeding event
- QRS_{dyn}: Amplitude of the smoothed norm of the first derivative of a 5-dimensional phase space representation of the QRS complex. QRS_{dyn} is furthermore scaled in a way that the median of all QRS_{dyn} values within a single ECG has size 1.
- T_{area}: Area under the T-wave, calculated as integral over the smoothed signal in the time interval [100ms;350ms] after the event trigger.
- Jitter: High frequency noise outside of the QRS complex, measured as spectral power of the frequency band 35-45 Hz, using the time interval [50ms;400ms] after the event trigger. This value is furthermore power-transformed by $f(x) = x^{0.25}$ to obtain a more symmetric distribution.

2.2. Neural network

A 2-layer perceptron neural network is used to classify each heart beat as belonging to one of four apnea states {N: no apnea); O: onset; M: maximum; R: recovery} (see Figure 1).
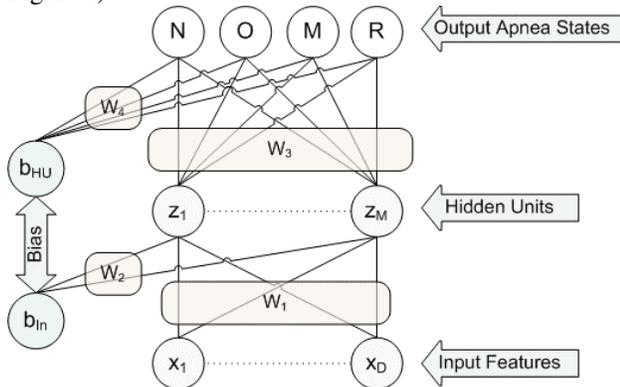

Figure 1. Neural network configuration with two layers of weights, organized in four regularizer groups W_1 to W_4.

Input of this neural network are six features which are measures of the event parameters described above, calculated from all normal events in one or more time intervals related to the current event:

- Mean(RR): Mean value of RR in the time interval [–4s; 4s]
- Slope(RR): Slope of RR calculated in the time intervals [-10s;-2s], [-4s;4s] and [2s;10s]
- Mean(QRS_{dyn}): Mean value of QRS_{dyn}, calculated in the time intervals [-10s;-2s], [-4s;4s] and [2s;10s]
- Stdev(QRS_{dyn}): Standard deviation of QRS_{dyn}, calculated in the time intervals [-10s;-2s], [-4s;4s] and [2s;10s]
- Mean(Jitter): Mean value of Jitter, calculated in the time intervals [-10s;-2s], [-4s;4s] and [2s;10s]
- Delta(T_{area}): Difference of the medians of T_{area}, calculated between the two time intervals [-10s;-2s] and [-4s;4s] and between the two intervals [-4s;4s] and [2s;10s]

These features measured in their different time intervals produce 15 input parameters for the neural network which are standardized for numerical reason. They represent dynamic information of the heart rate, the amplitude and energy within the QRS complex, modulations of the T-wave morphology and high frequency noise which probably represents muscular stress.

Output of the model are four real numbers, each between 0 and 1, which may be interpreted as probability that the current heart beat belongs to the appropriate apnea state.

To train the neural network each 10th minute of all 35
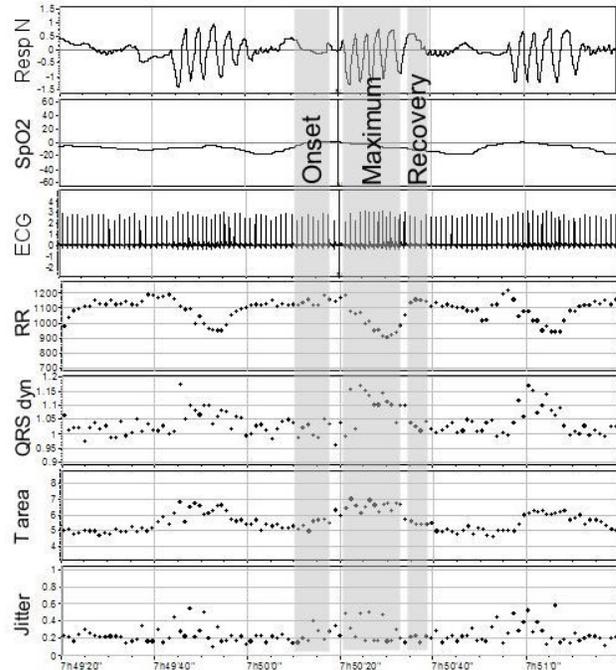
Figure 2. Example for annotation of apnea states used for training the neural network. Time axis length is 2 minutes. The upper two signals (Resp N: oronasal airflow; SpO2: oxygen saturation) are given for only 8 records. The lower four graphs show the event parameters described in section 2.1.

ECGs from the learning set were examined for location of the apnea if present. Onset, maximum and recovery of each apnea was marked as one thinks best (see Figure 2) (the autors tried to learn from those recordings where respiration and oxygen saturation signals are present). If no apnea happened within the examined minute, a time interval of about 30 seconds length was annotated as 'absence of apnea'. All heart beats lying in those marked regions were used for training (65707 out of about 1106000 normal events in the learning set).

Various model sizes are investigated and compared by their cross validation error. All neural networks use hyperbolic tangent activation functions and the model weights are restricted by four regularizers according to the four weight groups W_1 to W_4. These regularizers are hyper model parameters and determined by applying the Bayes-motivated evidence framework [1][3]. The importance of the individual features for prediction of sleep apnea is determined by leaving out one feature at a time and comparing the remaining prediction quality with the full model.

2.3. Dynamic Markovian state model

The output of the neural network is weighted by three adaptable weights and used by a Markovian state model
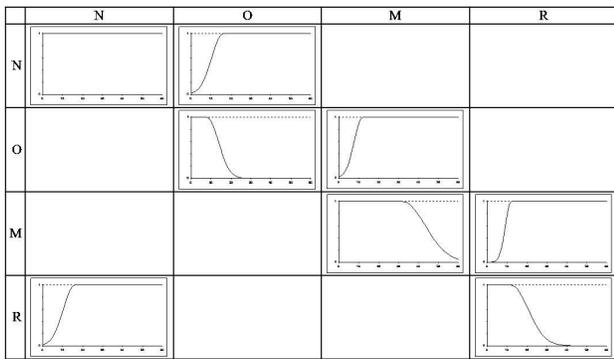
Figure 3. Matrix of transition probability functions between apnea states. N: no apnea; O: onset; M: maximum; R: recovery. Time axis length of each graph is 1 minute. Empty cells indicate forbidden transitions.

as probability of being in a specific state. Transitions between states are controlled by a matrix of probability functions (Figure 3) which depend on the time how long the current state already lasted (hence the term 'dynamic'). Only transitions in the sequence N→O→M→R→N are permitted. The time which can be spent in the states A, M and R is restricted, indicated by the diagonal elements of the matrix. The non-diagonal elements of the matrix enforce that a given state should only last for a certain time.

The probability of a state trajectory is calculated as the product of its state probabilities times the transition probabilities. The final trajectory is the one with maximal probability. As the number of all possible trajectories is intractable to handle, the optimization algorithm makes some simplifications: The trajectory hierarchy tree starts at the first time point (heart beat) with four possible states. Each time the trajectory tree is expanded by a new step, all trajectories which end in the same state are compared by their trajectory probabilities. Only the trajectory with maximum probability is kept while the others are removed. This prevents the number of trajectories to explode. Once the hierarchy root is unique, this root is cut off and used as solution. The probabilites of the remaining trajectories are rescaled to prevent the trajectory probabilities from becoming less than the numerical precision boundary. This algorithm is fast and can be applied to any number of heart beats.

Each transition probability function is specified by a position parameter which determines the time where maximum is reached or ends and by a scale parameter which adjusts the descending or increasing rate of the function. The seven non-trivial transition probability functions thus deliver 14 free parameters. Together with the three weights for the neural network output 17 parameters have to be estimated. This is done by maximizing the agreement whith the expert annotations in
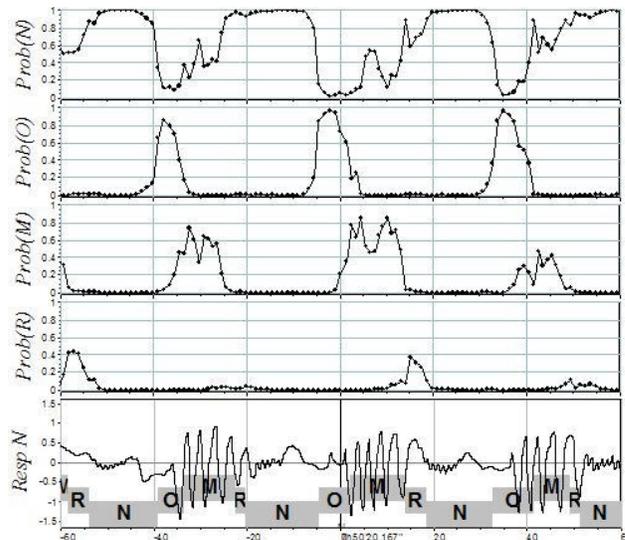
Figure 4. Probabilities of apnea states N, O, M and R as output of the neral network (same time interval as in Figure). Lower graph: respiration signal plus the apnea states generated by the state model.

the learning set from [2]. Agreement means that in a 1-minute segment with apnea present the model identifies at least one apnea state of type M, or if this minute does not contain apnea the model does not identify any apnea state of type M. The optimizing procedure is based on the wellknown simplex algorithm of Nelder and Mead [4].

3. Results

Figure 4 is showing the output of the neural network as probabilities of the four apnea states together with the solution generated by the state model. The time interval is identical to Figure 2. It can be seen that the apnea states O, M and R describe the later period of breath cessation and the following respiration activity. Cessation start is typically coded as state N. Distribution of the length of all apnea states is given in Table 1.

Best prediction is found using a neural network with 30 hidden units and transition matrix shown in Figure 3. Cross validation error is 0.5201 and the comparison with the expert annotations give 89.0% agreement within the learning set and 84.1% within the test set. Using only 20 hidden units worsens prediction quality only marginally (see Table 2, row: ⊕ all features).

Table 2 shows the change of prediction quality when

Table 1. Distribution of the length of apnea states.

Apnea state	Mean	Stdev	Min	Max
O	6.65	3.67	0.57	54.54
M	12.43	3.95	0.50	55.45
R	5.30	3.56	0.47	27.08

Table 2. Prediction quality depending on input features.

Neural network input features	Cross validation error	Agreement with challenge set	
		Learning	Test
⊕ All features	0.5390	88.1%	83.7%
⊕ RR	0.6997	76.0%	76.7%
∅ RR	0.6828	82.9%	79.0%
∅ QRS_{dyn}	0.5979	83.5%	82.6%
∅ T_{area}	0.5637	87.1%	82.8%
∅ Jitter	0.5664	85.1%	83.9%

⊕: feature used. ∅: feature excluded

All models use 20 hidden units and the transition matrix shown in Figure 3.

individual features are taken out. While removing the QRS_{dyn}, T_{area} or Jitter information individually has only minor influence of the prediction quality, removing the complete RR information reduces agreement of about 5%. However, the QRS_{dyn}, T_{area} and Jitter features are valuable information as can be seen when these features are completely ignored (row: ⊕ RR).

4. Discussion and conclusions

The manual annotations of the apnea states by the authors are definitely not perfect and so they are adding some uncertainty to the training data. Considering further that only about 6% of the heart beats from the learning records are used for training the neural network the observed prediction power of about 84% with respect to the test set is comparable to the methods presented in challenge 2000.

The features QRS_{dyn}, T_{area} and Jitter investigated here turned out to be valuable information for identification of apnea events. They enhance prediction quality about 7% compared to a model that uses only RR information. If they are used without RR information they are even more predictive than a model which is solely based on RR.

The largest neural network uses 604 adaptable weights. Applying the evidence framework in determining these weights together with their regularizer hyper parameters reliably prevents from overfitting and produces good generalization properties. The dynamic Markovian state model proved to be a usefull descision rule to identify onset and offset of individual apnea events, even if the output of the neural network is quite ambiguous.

The method introduced here was originally developed to identify sleep stages of animals with a high time resolution using the EEG. It could be shown that it is also applicable for determination of individual events of disordered breathing delivering a direct estimation of the apnea index in terms of events per hour.

References

[1] Bishop CM. Neural Networks for Pattern Recognition. Oxford University Press, 1995.
[2] Computers in Cardiology Challenge 2000. Detecting Sleep Apnea from the ECG. Computers in Cardiology 2000; 27:739-760.
[3] MacKay DJC. Bayesian Methods for Adaptive Models. Ph.D. Thesis, California Institute of Technology, 1992
[4] Numerical Recipes in C: The Art of Scientific Computing. Cambridge Univ Press 1988-1992.
[5] Penzel T, McNames J, de Chazal P, Raymond B, Murray A, Moody G. Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. Medical & Biological Engineering & Computing 2002; 40:402-407.

Address for correspondence

Werner Bystricky, TSD
Pfarrweg 5, D-74653 Künzelsau, Germany
e-mail: tsd@werner-bystricky.de