Computational Models of Stress in Reading Using Physiological and Physical Sensor Data

Nandita Sharma and Tom Gedeon

Research School of Computer Science
Australian National University, Canberra, Australia
{Nandita.Sharma,Tom.Gedeon}@anu.edu.au

Abstract. Stress is a major problem facing our world today and it is important to develop an objective understanding of how average individuals respond to stress in a typical activity like reading. The aim for this paper is to determine whether stress patterns can be recognized using individual-independent computational models from sensor based stress response signals induced by reading text with stressful content. The response signals were obtained by sensors that sourced various physiological and physical signals, from which hundreds of features were derived. The paper proposes feature selection methods to deal with redundant and irrelevant features and improve the performance of classifications obtained from models based on artificial neural networks (ANNs) and support vector machines (SVMs). A genetic algorithm (GA) and a novel method based on *pseudo-independence* of features are proposed as feature selection methods for the classifiers. Classification performances for the proposed classifiers are compared. The performance of the individual-independent classifiers improved when the feature selection methods were used. The GA-SVM hybrid produced the best results with a stress recognition rate of 98%.

Keywords: stress classification, artificial neural networks, genetic algorithms, support vector machines, physiological signals, physical signals, reading.

1 Introduction

Stress is part of everyday life and it has been widely accepted that stress which leads to less favorable states (such as anxiety, fear or anger) is a growing concern for people and society. The term, *stress*, was coined by Hans Selye. He defined it as "the non-specific response of the body to any demand for change" [1]. Stress is the body's reaction or response to the imbalance caused between demands and resources available to a person. Stress is seen as a natural alarm, resistance and exhaustion [2] system for the body to prepare for a fight or flight response to protect the body from threats and changes. When experienced for longer periods of time without being managed, stress has been widely recognized as a major growing concern. It has the potential to cause chronic illnesses (e.g. cardiovascular diseases, diabetes and some forms of cancer) and increase economic costs in societies, especially in developed countries [3, 4]. Benefits of stress research range from improving day-to-day activities, through

J. Pei et al. (Eds.): PAKDD 2013, Part I, LNAI 7818, pp. 111–122, 2013.

increasing work productivity to benefitting the wider society - motivating interest, making it a beneficial area of research and posing technical challenges in Computer Science. Various computational methods have been used to objectively define and classify stress to differentiate conditions causing stress from other conditions. The methods developed have used simplistic models formed from techniques like Bayesian networks [5], decision trees [6] and support vector machines [7]. These models have been built from a relatively smaller set of stress features than the sets used in the models in this paper.

The human body's response signals obtained from non-invasive methods that reflect reactions of individuals and their bodies to stressful situations have been used to interpret stress levels. These measures have provided a basis for defining stress objectively. Stress response signals used in this paper fall into two categories – physiological and physical signals. Physiological signals include the galvanic skin response (GSR), electrocardiogram (ECG) and blood pressure (BP). Unlike these signals, we define physical signals as signals where changes by the human body can be seen by humans without the need for equipment and tools that need to be attached to individuals to detect general fluctuations. However, sophisticated equipment and sensors using vision technologies are still needed to obtain physical signals at sampling rates sufficient for data analysis and modeling like the ones used in this paper. Physical signals include eye gaze and pupil dilation signals. GSR, ECG, BP, eye gaze and pupil dilation signals have been used to detect stress in literature [5, 8, 9] but this combination has not been reported in literature so far. We use this combination of sensor signals in this paper and refer to them as primary signals for stress.

Hundreds of stress features can be derived from primary signals for stress to classify stress classes. However, this set of features may include redundant and irrelevant features which may outweigh the more effective features showing stress patterns. This could cause a stress classification model to produce lower quality classifications. Since this paper is dealing with sensor data, some features may suffer from corruption as well. In order to achieve a good classification model that is robust to such potential features that may reduce the performance of classifications, appropriate feature selection methods must be developed and adopted by classifiers. A feature selection method that selects features in order to reduce redundancy using correlation based analysis could be used [10]. In addition, a genetic algorithm (GA) could also be used to select subsets of features for optimizing stress classifications. A GA is a global search algorithm and has been commonly used to solve optimization problems [11]. The search algorithm is based on the concept of natural evolution. It evolves a population of candidate solutions using crossover, mutation and selection methods in search for a population of a better quality. GAs have been successfully used to select features derived from physiological signals [12, 13].

This paper describes the method for collecting and developing computational models for recognizing stress patterns in response signals observed from individuals while reading *stressed* and *non-stressed* labeled text validated by participants. It details an experiment conducted to collect sensor and participant-reported data where experiment participants read text with stressful and non-stressful content during which multiple response signals were recorded. Several approaches for stress recognition in reading are developed, compared and discussed including methods for selecting features from hundreds of features derived from the response signals. The paper concludes with a summary of the findings and suggests directions for future work.

2 Data Collection from Reading Experiment

Thirty-five undergraduate students were recruited as experiment participants. The participant cohort was made up of 25 males and 10 females over the age of 18 years old. Each participant had to understand the requirements of the experiment from a written set of experiment instructions with the guidance of the experiment instructor before they provided their consent to take part in the experiment. Afterwards, physiological stress sensors were attached to the participant and physical stress sensors were calibrated. The instructor notified the participant to start reading, which triggered a sequence of text paragraphs. After finishing the reading, participants had to do an assessment based on the reading. An outline of the process of the experiment for an experiment participant is shown in Fig. 1.

Each participant had physiological and physical measurements taken over the 12 minutes reading time period. During the reading period, a participant read *stressed* and *non-stressed* types of text validated by participants. Stressed text had stressful content in the direction towards distress, fear and tension. Each participant read three stressed and three non-stressed text. Each text had approximately 360 words and was displayed on a computer monitor for participants to read. For consistency, each text was displayed on a 1050 x 1680 pixel Dell monitor, displayed for 60 seconds and positioned at the same location of the computer screen for each participant. Each line of the paragraph had 70 characters including spaces.

Results from the experiment survey validated the text classes. This is a common method used in literature to validate stress classes for tasks [14]. Participants found the paragraphs that were labeled stressed stressful and text labeled non-stressed as not stressful with a statistical significance of $p < 0.001$ according to the Wilcoxon test.

The physiological and physical sensor signals (which we refer to as primary stress signals) captured during the experiment were GSR, ECG, BP, eye gaze and pupil diameter signals. Biopac ECG100C, Biopac GSR100C and Finapres Finger Cuff systems were used to take ECG, GSR and blood pressure recordings at a sampling rate of 1000 Hz. Eye gaze and pupil dilation signals were obtained using Seeing Machines FaceLAB system with a pair of infrared cameras at 60 Hz. There were other signals that were derived from the primary stress signals to form other stress response signals. These signals included the heart rate variability signal, which was calculated from consecutive ECG peaks and another popular signal used for stress detection [15, 16].

Features were derived from the primary stress signals. Statistics (e.g. mean and standard deviation) were calculated for the signal measurements for each 5 second interval during the stressed and non-stressed reading. Measures such as the number of peaks for periodic signals, the distance an eye covered, the number of forward and backward tracking fixations, and the proportion of the time the eye fixated on different regions of the computer screen over 5 second intervals were also obtained. The statistic and measure values formed the stress feature set. There were 215 features altogether.

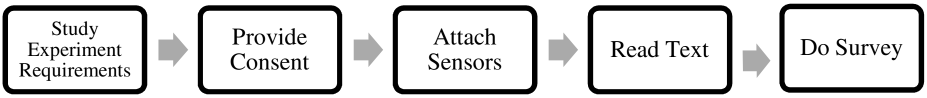


Fig. 1. Process followed by participants during the reading experiment

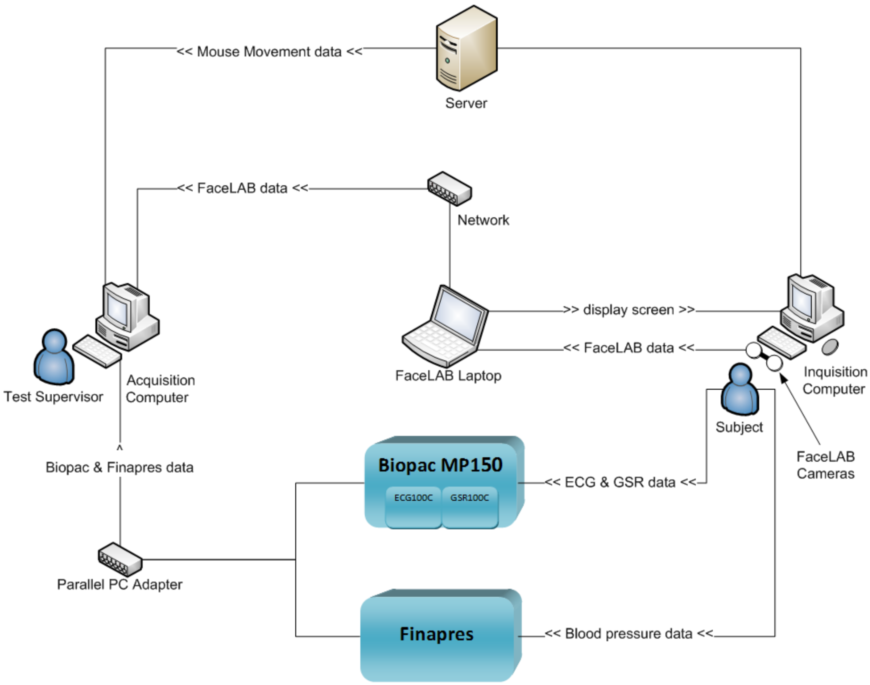


Fig. 2. Equipment setup for the reading experiment

3 Feature Selection

Features used for developing models for classification had an effect on the performance of classifiers. Selecting effective features by reducing redundant and irrelevant features have been known to improve the quality of pattern recognition [17] because it generalizes the patterns in the data better and helps develop a generalized model that captures necessary data patterns. In turn, this improves the quality for classifications. In this paper, a feature selection method based on correlation of features and a genetic algorithm (GA) approach were developed and used as feature selection methods for stress classification.

The features derived from the stress primary signals may have had redundant data so a feature selection approach based on *correlation coefficients* for features was developed. Correlation analysis using correlation coefficients has been reported to detect some redundancies in data [10]. It also took into account the time-varying nature of features and enabled comparison of features on this basis.

A correlation coefficient is a measure for the strength of the linear relationship between features. Consider two features, X and Y, with x_t and y_t values at time-step t in X and Y respectively, $\overline{X}$ and $\overline{Y}$ the means and σ_X and σ_Y are the standard deviations for X and Y, then the correlation coefficient r_{XY} is defined by

$$r_{XY} = \frac{\sum_{t=0}^{T}(x_t - \bar{X})(y_t - \bar{Y})}{(T+1)\sigma_X \sigma_Y} \tag{1}$$

The values for r_{XY} fulfill the following equation:

$$|r_{XY}| \leq 1 \tag{2}$$

If the value for $r_{XY} = 0$, then features X and Y are independent otherwise the features are correlated. However, stress features may have noise originating from data collection and the human body so as a result. In addition, the definition for independence of features for stress classification may be too strict. This motivated the use of different degrees for feature independence. In order to distinguish from independence defined by the strict criteria, we coin the term *pseudo-independence* to mean independence at a certain degree. Suppose the degree of independence is set at $r_{XY} \leq 0.1$ and r_{XY} is found to be 0.05, then features X and Y are pseudo-independent. On the other hand if r_{XY} is found to be 0.25, then features X and Y are not pseudo-independent.

The pseudo-independent based feature selection method, *pseudo-independent feature selection algorithm* (PISA), was used to select stress features that a classifier was provided to detect stress patterns. Given a set of features, each feature was compared with each of the other features to determine whether they were not pseudo-independent features. Features and their not pseudo-independent features were found and used to generate a set of pseudo-independent features for classification. The process by which a set of pseudo-independent features were obtained was by PISA. PISA is defined in Fig. 3.

Given a set of features and a set of features that are not pseudo-independent to each feature, PISA firstly finds features that are pseudo-independent to every other feature. Then it selects one feature from a cluster of not pseudo-independent features. The feature selected from a cluster is based on its not pseudo-independent features – the number of features and whether the features have been already selected. Features with a higher number of not pseudo-independent features have a greater chance of being selected. This characteristic of PISA minimizes the impact of highly correlated noisy features on classification.

Algorithm: Pseudo-Independent Feature Selection Algorithm (PISA). Find features that are pseudo-independent to other features given a set of features and their pseudo-independent features

Inputs:

- A set of features, *feature_set*
- Collection of features that are correlated to some degree to the other features in *feature_set* with a mapping to the features that they are not pseudo-independent with in *feature_set, corr_feat_mapping*

Output:

- Collection of pseudo-independent features

Method:

1. *notcorr_feat_collection* ← get features that are in *feature_set* and not in *corr_feat_mapping*
2. for each feature *feat* selected in descending order of the number of features that they are correlated with in *corr_feat_mapping* {
3. *corr_feats* ← get the features correlated with *feat* from *corr_feat_mapping*
4. if none of the features in *corr_feats* are in *notcorr_feat_collection* {
5. then add *feat* in *notcorr_feat_collection* }}

Fig. 3. An algorithm to generate a set of pseudo-independent features (PISA) for stress classification

To illustrate PISA, consider an example input set comprising features and the features that are not pseudo-independent to the features presented in Fig. 4. Firstly, PISA determines features that do not have any not pseudo-independent features. A set {f4} is the result at this point. Now, the rest of the pseudo-independent features need to be appended to the set. There are multiple possibilities and they are {f1, f3, f4} and {f2, f4}. After applying the algorithm the result will be {f2, f4}, which is the smaller feature set of the multiple feature sets with pseudo-independent features. This approach reduces the risk of selecting features that were derived from corrupted sensor data and negatively affect the performance for stress classification. Suppose the feature set {f1, f2, f3} were derived from corrupted sensor data and if {f1, f3, f4} was selected as a set of features for classification, then intuitively the possibility of a classification model to capture better stress patterns would be lower than if {f2, f4} was chosen instead.

Inputs:
feature_set: {f1, f2, f3, f4}
corr_feat_mapping:

Feature		*Feature(s) not pseudo-independent to associated feature*
f1	→	{f2}
f2	→	{f1, f3}
f3	→	{f2}

Output: {f2, f4}

Fig. 4. An example to illustrate the inputs and corresponding output for the PI algorithm presented in Fig. 3 using a simple set of features

Another feature selection method used for stress classification in reading was based on a GA. A GA is a global search algorithm that was used to select features to improve the quality of stress classifications. The GA search evolved a population of subsets of features using crossover, mutation and selection methods in search for a population of subset of features that produced a better quality stress classification. A subset of features is referred to as an *individual* or *chromosome*. The quality for each chromosome in the population was defined by the quality of classifications produced when a classifier was provided with the features in the chromosome.

The initial population for the GAs was set up to have all the features. The number of features in the chromosomes varied but the chromosome length was fixed. The length of a chromosome was equal to the number of features in the feature space. A chromosome was a binary string where the index for a bit represented a feature and the bit value indicated whether the feature was used in the classification.

The parameters for the GAs implemented were set as provided in Table 1.

Table 1. Parameter settings for GAs used for feature selection

GA Parameter	Value/Setting
population size	100
number of generations	2000
crossover rate	0.8
mutation rate	0.01
crossover type	MATLAB's Scattered Crossover
mutation type	MATLAB's Uniform Mutation
selection type	MATLAB's Stochastic Uniform Selection

4 Computational Stress Classifiers

Classification models developed for stress pattern recognition from primary stress signals were based on an artificial neural network (ANN) and a support vector machine (SVM). The models were extended to incorporate feature selection phases either using the PISA or GA approaches. Each classification model was defined to capture individual-independent stress patterns. The accuracy and F-Score were calculated for each approach to determine the quality of the classification.

The stress reading data set was divided up into 3 subsets – training, validation and test sets – where 50% of the data samples were used for training a classification model and the rest of the data set was divided up equally for validating and testing the model. MATLAB was used to implement and test the models.

4.1 Artificial Neural Network Based Stress Classifiers

ANNs, inspired by biological neural networks, have capabilities for learning patterns to recognize characteristics in input tuples by classes. An ANN is made up of interconnected processors, known as *artificial neurons*, which are connected by weighted links that pass signals between neurons. In this paper, feed-forward ANNs trained using backpropagation were used. Three topologies were used to classify stress in reading. Each of the ANNs was provided stress features as inputs based on a selection method. Therefore, the ANNs differed only on the number of inputs. The ANN based stress classification models were:

- **ANN:** an artificial neural network classification model that was provided with all the features in the stress feature set as input to recognize stress patterns
- **PISA+ANN:** ANN with inputs as features produced by PISA
- **GA+ANN:** ANN with inputs as features produced by a GA

The MATLAB adapt function was used for training the ANNs on an incremental basis. Each network was trained for 1000 epochs using the Levenberg-Marquardt algorithm. The network had 7 hidden neurons and one neuron in the output layer. Future work could investigate optimizing the topology of the ANN for stress classification on the reading data set.

4.2 Support Vector Machine Based Stress Classifiers

SVMs have been widely used in literature for classification problems including classifications based on physiological data [18]. Provided a set of training samples, a SVM transforms the data samples using a nonlinear mapping to a higher dimension with the aim to determine a *hyperplane* that partitions data by class or labels. A hyperplane is chosen based on *support vectors*, which are training data samples that define maximum *margins* from the support vectors to the hyperplane to form the best decision boundary. This contributes to the resistance to data overfitting and helps to generalize classifications well.

Despite the useful characteristics, SVMs are still not robust to feature sets with redundant and irrelevant features. As a consequence, hybrids of SVM with PISA or GA were used to deal with ineffective features to investigate whether the hybrids with feature selection methods improved the quality of the classification.

The SVM based stress classification models developed were:

- **SVM:** a support vector machine classification model that was provided with all the features in the stress feature set as input similar to the ANN
- **PISA+SVM:** SVM with inputs as features produced by PISA
- **GA+SVM:** SVM with inputs as features produced by a GA

5 Results and Discussion

The six ANN and SVM based techniques were implemented and tested on the reading data set for stress recognition. Classification results were obtained using 10-fold cross-validation. The results are presented in Table 2. Classifiers with feature selection methods performed better than classifiers that used all stress features to model stress patterns. The hybrid classifiers had stress recognition rates and F-score values that were at least 8% and 12% better respectively. Classifiers with a GA as the feature selection method produced the highest stress recognition rates with GA+SVM as the best performing technique.

Performance measures for the classification techniques show that it was beneficial to use the feature selection methods to model stress patterns in reading. The stress features would have had redundant features and PISA would have reduced it. PISA compared every feature to every other feature in a pair-wise fashion and took a greedy approach in selecting features. Unlike PISA, the GA took a global view of the features and would have managed to reduce more redundant, irrelevant and corrupted features.

In terms of execution time, the GA based approaches, GA+ANN and GA+SVM, took longer times than the other techniques to produce solutions. It took PISA less than one second to select the features for the classification models whereas the execution times for the GA based approaches were in the order of hours. Classification without a feature selection method or with PISA took relatively a similar amount of time. Empirical execution times for the different approaches are shown in Table 3.

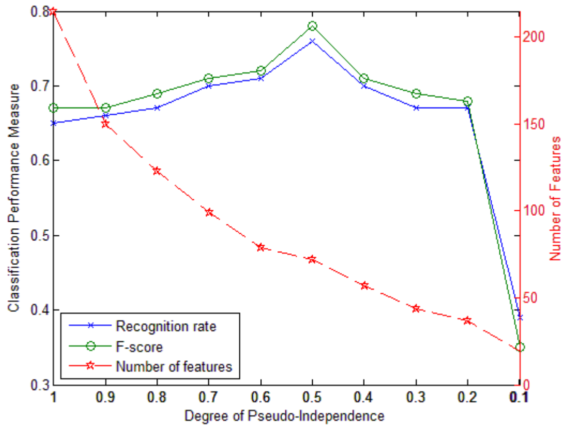
The execution times for GA based approaches were recorded after the search reached convergence except for GA+ANN, which took a lot longer to execute. GA+ANN took at least 3 days and it took the other techniques not more than a few hours to produce a solution. Therefore, it was not practical to let the GA+ANN search execute for longer. Table 2 and Table 3 have a * with the values to show results at the point when the GA+ANN search was terminated.

Table 2. Performance measures for stress recognition using the different approaches based on 10-fold cross-validation

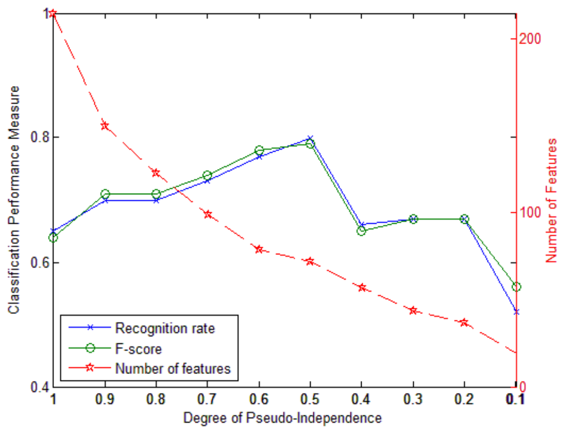
Classification Performance Measure	ANN	PISA+ANN	GA+ANN	SVM	PISA+SVM	GA+SVM
Accuracy	0.68	0.76	0.82*	0.67	0.80	**0.98**
F-score	0.67	0.79	0.82*	0.67	0.79	**0.98**

Table 3. Relative execution times for the different classification techniques

	ANN	PISA+ANN	GA+ANN	SVM	PISA+SVM	GA+SVM
Execution time (minutes)	5.1	5.1	4582*	0.5	0.5	275



(a)



(b)

Fig. 5. Performances for stress recognition using different degrees of pseudo-independence and PISA as the feature selection method for the classification (a) ANN based classification (b) SVM based classification

Due to the relative long execution times for the GA based approaches to search for a better stress classification result, other feature selection approaches can be investigated in the future. With shorter execution times and classification performance results for classifiers using PISA, PISA has the potential to increase the performance for classifications. In future, PISA based classifiers could be extended to have a more complex definition for pseudo-independence.

ANN and SVM classification results using PISA over a range of degrees of pseudo-independence were also obtained to determine the effect of the different degrees of pseudo-independence on the classification results. The results are displayed in graphs shown in Fig. 5. The graphs show the stress recognition rates and F-score along with the number of features for the different degrees of pseudo-independence. For both, ANN and SVM, the plot for the classification results show a positive overall rate of change and then it becomes negative after 0.5 degree of pseudo-independence. The best classification results are produced when the degree of pseudo-independence is 0.5. At this degree, every feature is pseudo-independent at the degree of 0.5 to every other feature in the set of features used as input to develop the ANN and SVM classification models.

6 Conclusion and Future Work

Classification models were developed to recognize individual-independent stress patterns in physiological and physical data for reading. The use of a feature selection method that dealt with redundant features improved the quality of the classification. However, classification models based on a genetic algorithm provided better recognition rates for stress than a correlation based feature selection method. On the other hand, genetic algorithm based approaches required much longer execution times but the correlation based feature selection method hardly had any impact on the execution time. In future, the correlation based feature selection method could be extended to have a wider definition for feature independence. Further, a hybrid of the two feature selection methods presented in this work could be developed to reduce the execution time for the genetic algorithm based search. Moreover in this work, features were selected from an individual-independent viewpoint. Future work could investigate feature selection based on relationships of features for each individual and analyze their effect of the approach on classifications.

References

[1] Selye, H.: The stress syndrome. The American Journal of Nursing 65, 97–99 (1965)

[2] Hoffman-Goetz, L., Pedersen, B.K.: Exercise and the immune system: a model of the stress response? Immunology Today 15, 382–387 (1994)

[3] The-American-Institute-of-Stress. America's No. 1 Health Problem - Why is there more stress today? (August 05, 2010), `http://www.stress.org/americas.htm`

[4] Lifeline-Australia, Stress Costs Taxpayer $300K Every Day (2009), `http://www.lifeline.org.au`

[5] Liao, W., Zhang, W., Zhu, Z., Ji, Q.: A real-time human stress monitoring system using dynamic bayesian network. In: Computer Vision and Pattern Recognition - Workshops, CVPR Workshops (2005)
[6] Zhai, J., Barreto, A.: Stress recognition using non-invasive technology. In: Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference, FLAIRS, pp. 395–400 (2006)
[7] Dou, Q.: An SVM ranking approach to stress assignment. University of Alberta (2009)
[8] Labbé, E., Schmidt, N., Babin, J., Pharr, M.: Coping with stress: the effectiveness of different types of music. Applied Psychophysiology and Biofeedback 32, 163–168 (2007)
[9] Healey, J.A., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on Intelligent Transportation Systems 6, 156–166 (2005)
[10] Yu, L., Liu, H.: Feature selection for high-dimensional data: A fast correlation-based filter solution. In: 12th International Conference on Machine Learning, pp. 856–863 (2003)
[11] Goldberg, D.E.: Genetic algorithms in search, optimization, and machine learning. Addison-wesley (1989)
[12] Park, B.J., Jang, E.H., Kim, S.H., Huh, C., Sohn, J.H.: Feature selection on multi-physiological signals for emotion recognition. In: 2011 International Conference on Engineering and Industries (ICEI), Korea, pp. 1–6 (2011)
[13] Niu, X., Chen, L., Chen, Q.: Research on genetic algorithm based on emotion recognition using physiological signals. In: 2011 International Conference on Computational Problem-Solving (ICCP), pp. 614–618 (2011)
[14] Hill, J.D., Boyle, L.N.: Driver stress as influenced by driving maneuvers and roadway conditions. Transportation Research Part F: Traffic Psychology and Behaviour 10, 177–186 (2007)
[15] Ferreira, P., Sanches, P., Höök, K., Jaensson, T.: License to chill!: how to empower users to cope with stress. In: Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges, pp. 123–132 (2008)
[16] Dishman, R.K., Nakamura, Y., Garcia, M.E., Thompson, R.W., Dunn, A.L., Blair, S.N.: Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. International Journal of Psychophysiology 37, 121–133 (2000)
[17] Siedlecki, W., Sklansky, J.: A note on genetic algorithms for large-scale feature selection. Pattern Recognition Letters 10, 335–347 (1989)
[18] Cheng, B.: Emotion recognition from physiological signals using support vector machine. In: Wu, Y. (ed.) Software Engineering and Knowledge Engineering. AISC, vol. 114, pp. 49–52. Springer, Heidelberg (2012)