

MULTI-OBJECTIVE GENETIC ALGORITHMS AS AN EFFECTIVE TOOL FOR FEATURE SELECTION IN THE SPEECH-BASED EMOTION RECOGNITION PROBLEM

Ch.Yu. Brester¹, O.E. Semenkina², M.Yu. Sidorov³

^{1,2} Siberian State Aerospace University

31 Krasnoyarsky Rabochy ave., Krasnoyarsk, 660014, Russia

E-mail: ¹christina.bre@mail.ru, ²semenkina.olga@mail.ru

³Ulm University, Albert-Einstein-Allee, 43, Ulm, 89081, Germany

E-mail: maxim.sidorov@uni-ulm.de

Feature selection is a quite important step in data analysis. Extracting relevant attributes may not only decrease the dimensionality of the dataset and, consequently, reduce time costs spent on the next stages, but also contribute to the quality of the final solution. In this paper we demonstrate some positive effects of the usage of a heuristic feature selection scheme which is based on a two-criterion optimization model. The approach proposed is applied to the speech-based emotion recognition problem, which is currently one of the most important issues in human-machine interactions. A number of high-dimensional multilingual (English, German, Japanese) databases are involved to investigate the effectiveness of the technique presented. Three different multi-objective genetic algorithms and their cooperative modifications are applied as optimizers in combination with classification models such as a Multilayer Perceptron, a Support Vector Machine and Logistic Regression. In most cases we may observe not only a dimensionality reduction, but also an improvement in the recognition quality. To avoid choosing the most effective multi-objective genetic algorithm and the best classifier, we suggest applying a heterogeneous genetic algorithm based on several heuristics and an ensemble of diverse classification models.

Keywords: feature selection, multi-objective genetic algorithm, island model, speech-based emotion recognition.

ГЕНЕТИЧЕСКИЕ АЛГОРИТМЫ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ КАК ЭФФЕКТИВНЫЙ ИНСТРУМЕНТ ОТБОРА ПРИЗНАКОВ В ЗАДАЧЕ РАСПОЗНАВАНИЯ ЭМОЦИЙ ПО РЕЧИ

К.Ю. Брестер¹, О.Э. Семенкина², М.Ю. Сидоров³

^{1,2} Сибирский государственный аэрокосмический университет,

Российская Федерация, 660014, г. Красноярск, пр. имени газеты «Красноярский рабочий», 31

E-mail: ¹christina.bre@mail.ru, ²semenkina.olga@mail.ru

³Ульмский университет, Германия, 89081, г. Ульм, аллея им. Альберта Эйнштейна, 43

E-mail: maxim.sidorov@uni-ulm.de

Отбор информативных признаков является одним из важных этапов анализа данных. Извлечение релевантных атрибутов может не только снизить размерность набора данных, а, следовательно, сократить временные затраты на последующих стадиях, но и улучшить качество финального решения. В данной работе демонстрируются положительные эффекты использования эвристической схемы отбора информативных признаков, основанной на двухкритериальной оптимизационной модели. Предлагаемый подход применяется к задаче распознавания эмоций человека по речи, что в настоящее время является одним из ключевых вопросов

в сфере человеко-машинных коммуникаций. Для исследования эффективности представленной технологии были привлечены базы данных высокой размерности: они содержат акустические характеристики голосовых записей на английском, немецком и японском языках. Три различных генетических алгоритма многокритериальной оптимизации и их кооперативные модификации были использованы в сочетании с рядом классификаторов (полносвязный перцептрон, машины опорных векторов, логистическая регрессия). В большинстве случаев можно обнаружить не только сокращение размерности вектора признаков, но и улучшение качества распознавания эмоций. Чтобы избежать выбора наиболее эффективного генетического алгоритма и классификатора, предлагается использовать гетерогенный кооперативный алгоритм, сочетающий несколько эвристик, и ансамбль классификаторов различной природы.

Ключевые слова: отбор признаков, генетический алгоритм многокритериальной оптимизации, опорная модель, распознавание эмоций по речи.

Introduction

Nowadays due to the tremendous capacity of data storage it has become possible to collect a huge amount of information. But the crucial question which needs to be answered by researchers is how to use all this data in an effective way? In the international scientific community there are plenty of discussions about proper solutions. Generally, two common directions might be distinguished. The first way is extensive, which implies a constant increase in computer power and the usage of as much data as possible. The second one is intensive, which is based on the subtle methodology of choosing the most relevant data such as instance or feature selection. From our perspective, the intensive approach seems to be more preferable because it refers to the thorough pre-processing of the source data which is considered to be an essential stage before the application of any mathematical model.

In this paper we discuss one of the possible approaches for selecting informative attributes from datasets in the framework of classification problems. We propose a heuristic feature selection scheme based on a two-criterion optimization model as an alternative to conventional methods (Principal Component Analysis (PCA) or Factor Analysis). Various multi-objective genetic algorithms (MOGAs) and their modifications are used as optimizers; a Multilayer Perceptron, a Support Vector Machine and Logistic Regression are involved as classification models.

In the human-machine communication sphere lots of high-dimensional feature sets are gathered too. Speech records are now used to train machines to reveal the speaker state: there is a wide spectrum of different classification problems such as gender, age or emotion recognition. However, the number of acoustic characteristics which might be extracted from the speech signal varies from several hundred to thousands. Therefore, it is necessary to select a subsystem of informative features related to each problem. And in this study we investigate the effectiveness of the approach proposed on the emotion recognition problem using a set of multilingual corpora (English, German, and Japanese).

The rest of the paper is organized as follows: the next section contains a brief description of the approach proposed. Then, some details of MOGAs and their modifications are presented. Further, we introduce the speech-based emotion recognition problem and the corpora used. Next, we continue with the experiments conducted and the results obtained. The conclusion includes some inferences and future plans.

Heuristic feature selection

Previously, it was demonstrated [1] that for the speech-based emotion recognition problem selecting informative features with the conventional PCA led to a tremendous decrease in the

classifier performance. Therefore, based on the experimental results [2], we highlight the necessity of developing some effective alternative methods.

In recent times there has been a growing interest in the sphere of *Evolutionary Machine Learning*. The integration of an evolutionary search into the machine learning field allows researchers to develop more universal algorithmic schemes which might be applied for high-dimensional problems with different types of variables in the dynamic environment. Taking into account these positive effects, we decided to engage genetic algorithms (GAs) in the feature selection procedure.

Firstly, a two-criterion optimization model was designed based on the feature selection scheme called *filter* [3]. This approach is referred to the pre-processing stage because it uses information extracted from datasets and reduces the number of attributes, taking into consideration such measures as consistency, dependency and distance. Possible criteria which characterize the dataset relevance are *Attribute Class Correlation*, *Inter- and Intra- Class Distances*, *Laplacian Score*, *Representation Entropy* and *the Inconsistent Example Pair measure* [4]. We should also emphasize that in the framework of this approach the information about classifier performance and a learning algorithm is ignored totally, therefore, feature selection procedures based on the filter scheme might be effectively used in combination with an ensemble of diverse classifiers, which is quite reasonable in such a case when we do not know one particular reliable and effective model.

The two-criterion model, which we propose using, contains *the Intra-class distance (IA)* and *the Inter-class distance (IE)* as optimized criteria:

$$IA = \frac{1}{n} \sum_{r=1}^k \sum_{j=1}^{n_r} d(p_j^r, p_r) \rightarrow \min, \quad (1)$$

$$IE = \frac{1}{n} \sum_{r=1}^k n_r d(p_r, p) \rightarrow \max, \quad (2)$$

where p_j^r is the j -th example from the r -th class, p is the central example of the dataset, $d(\dots)$ denotes the Euclidian distance, p_r and n_r represent the central example and the number of examples in the r -th class.

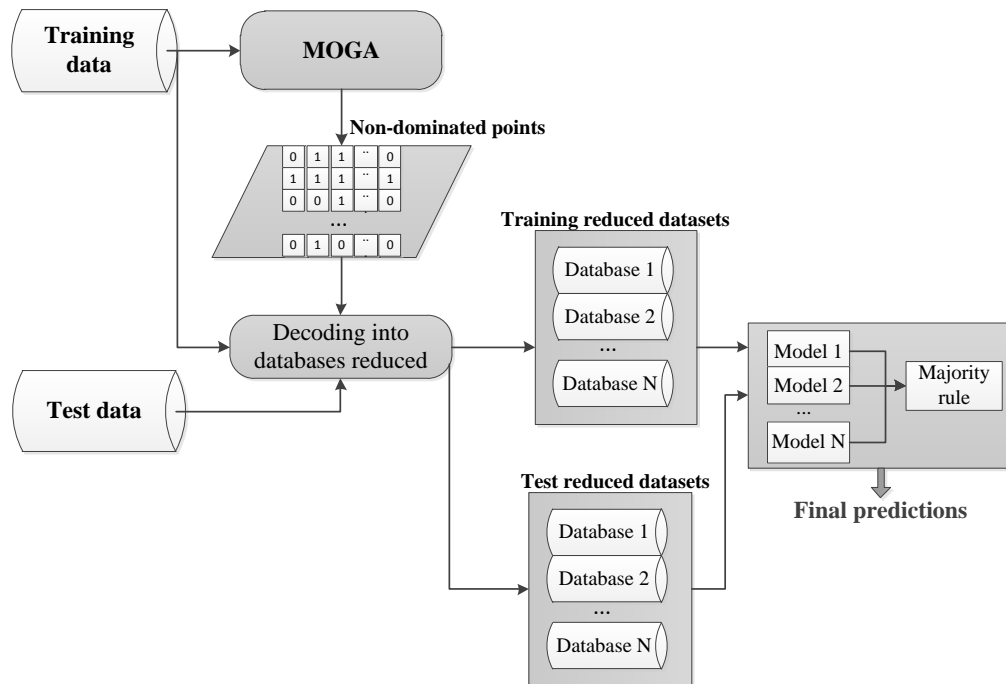


Fig. 1. The general scheme of the approach proposed

To define possible solutions, we suggest applying a MOGA which operates with binary strings, where *unit* and *zero* correspond to a relevant attribute and an irrelevant one respectively. In contrast to one-criterion GAs, the outcome of MOGAs is a set of non-dominated points which form the Pareto set approximation. Non-dominated candidate-solutions cannot be preferred to each other and, taking into account this fact, we propose a way to derive the final solution based on all the points from the Pareto set (Figure 1).

In the experiments the sample should be divided into the training and test parts. It is assumed that the outcome of a MOGA, which is executed on the training examples, is N binary strings (the set of non-dominated solutions). Each chromosome should be decoded into reduced databases (training and test parts), according to the rule: if a gene is equal to '0' then eliminate the corresponding attribute, and if a gene is equal to '1' then include the respective feature in the reduced database. In short, we obtain N different sets of features and train N various classifiers based on this data. For each test example the engaged models vote for different classes according to their own predictions. The final decision is defined as a collective choice based on the majority rule.

Taking into consideration the predictions of several classifiers is a good alternative to choosing one particular solution from the set of non-dominated points. In fact, candidates, which demonstrate high effectiveness on the training data, might often be the worst on the test data. Therefore, to avoid such cases, we use the scheme described.

In most cases the quality of the final solution, found with a MOGA, depends on the algorithm settings. For the problem considered the effectiveness of different heuristics might vary significantly. Therefore, in this study we apply a number of MOGAs which are based on diverse heuristic mechanisms. Moreover, we are also trying to improve the performance of conventional MOGAs by implementing their cooperative modifications. The next section provides a concise description of the algorithms used.

Multi-objective genetic algorithms and their cooperative modifications

The common scheme of any MOGA includes the same steps as any conventional one-criterion GA:

```

Generate the initial population
Evaluate criteria values
While (stop-criterion!=true), do:
    {Estimate fitness-values;
    Choose the most appropriate individuals with the mating selection
    operator based on their fitness-values;
    Produce new candidate solutions with recombination;
    Modify the obtained individuals with mutation;
    Evaluate criteria values for new candidate solutions;
    Compose the new population (environmental selection);
    }

```

Designing a MOGA, researchers are faced with some issues which are related to fitness assignment strategies, diversity preservation techniques and ways of elitism implementation. Therefore, in this paper we investigate the effectiveness of MOGAs, which are based on various heuristics, from the perspective of the feature selection procedure. *Non-Sorting Genetic Algorithm II (NSGA-II)* [5], *Preference-Inspired Co-Evolutionary Algorithm with goal vectors (PICEA-g)* [6] and *Strength Pareto Evolutionary Algorithm 2 (SPEA2)* [7] are used as tools to optimize the introduced criteria (1), (2). In Table 1 there are the basic features of each method.

Basic features of the MOGA used

MOGA	Fitness Assignment	Diversity Preservation	Elitism
NSGA-II	Pareto-dominance (<i>niching mechanism</i>) and diversity estimation (<i>crowding distance</i>)	Crowding distance	Combination of the previous population and the offspring
PICEA-g	Pareto-dominance (<i>with generating goal vectors</i>)	Nearest neighbour technique	The archive set and combination of the previous population and the offspring
SPEA2	Pareto-dominance (<i>niching mechanism</i>) and density estimation (<i>the distance to the k-th nearest neighbour in the objective space</i>)	Nearest neighbour technique	The archive set

However, it is almost impossible to know in advance which algorithm is the most effective for the current problem. On the one hand, a series of experiments might be conducted to find the best MOGA, which is quite a time-consuming way. On the other hand, different algorithms might be combined in a cooperation to avoid the choice of the most effective one. Actually, this kind of modification is easily implemented based on an island model.

The *island model* [8] of a GA implies the parallel work of several algorithms: they might be the same or different. The initial number of individuals M is spread across L subpopulations: $M_i=M/L, i=1, \dots, L$. At each T -th generation algorithms exchange the best solutions (*migration*). There are two parameters: *migration size*, the number of candidates for migration, and *migration interval*, the number of generations between migrations. It is also necessary to define the island model topology, in other words, the scheme of migration. We use the fully connected topology that means each island shares its best solutions with all other islands included in the model. This multi-agent model is expected to preserve a higher level of genetic diversity.

Firstly, conventional NSGA-II, PICEA-g, and SPEA2 have been implemented to be used as optimizers in the feature selection procedure.

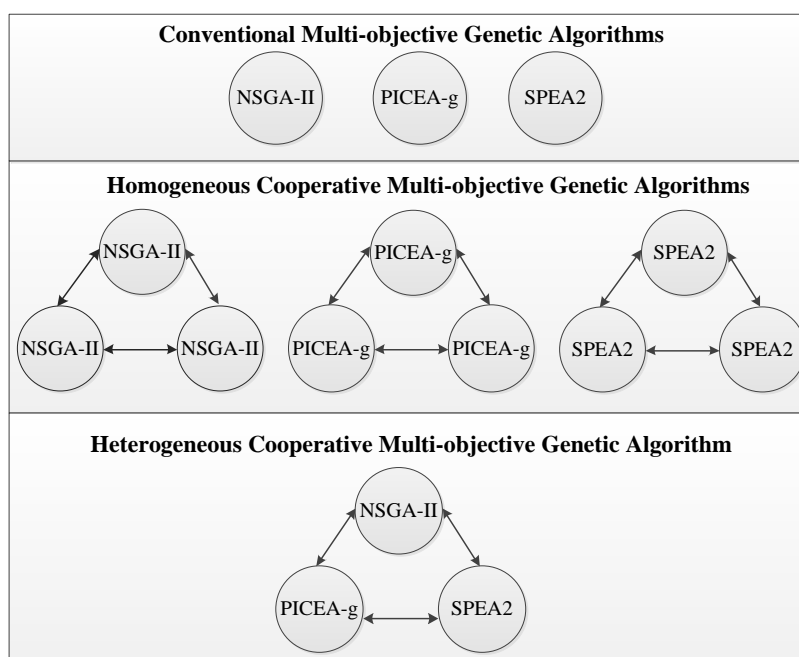


Fig. 2. The three categories of algorithms used

Secondly, we have achieved a number of *homogeneous* cooperative algorithms: in each case the island model has the same three components: they are NSGA-II, PICEA-g or SPEA2. In addition to diversity preservation, another benefit of this model is the possibility to reduce the computational time due to the parallel work of islands.

Finally, a *heterogeneous* cooperative algorithm has been developed. Three different MOGAs (NSGA-II, PICEA-g and SPEA2) have been included in this model as its components at once. The benefits of the particular algorithm (NSGA-II, PICEA-g or SPEA2) could be advantageous at different stages of optimization [9].

To sum up, there are three main categories of MOGAs which are used in this study and they are portrayed in Figure 2.

Speech-based emotion recognition and corpora description

One of the obvious ways to improve the intellectual abilities of spoken dialogue systems is related to their personalization. While communicating, machines should perceive the qualities of the user (as people usually do) such as age, gender and emotions to adapt their answers for the particular speaker.

In this paper we consider one particular aspect of the personalization process that is *speech-based emotion recognition*. Generally, any approach used to solve this recognition problem consists of three main stages.

At first, it is necessary to extract acoustic characteristics from the collected utterances. At the «*INTERSPEECH 2009 Emotion Challenge*» an appropriate set of acoustic characteristics representing any speech signals was introduced. This set of features includes attributes such as power, mean, root mean square, jitter, shimmer, 12 MFCCs and 5 formants. The mean, minimum, maximum, range and deviation of the following features are also used: pitch, intensity and harmonicity. The number of characteristics is 384. To get the conventional feature set introduced at *INTERSPEECH 2009*, the Praat [10] or OpenSMILE [11] systems might be used. Secondly, all extracted attributes or the most relevant of them should be involved in the supervised learning process to adjust a classifier. At the final stage, the signal that has to be analysed is transformed into an unlabelled feature vector (also with the usage of the Praat or OpenSMILE systems) and then the trained classification model receives it as the input data to make a prediction.

In the study a number of multilingual speech databases have been used and here we provide their brief description.

Table 2

Statistical description of the used corpora

Database	Language	Full length (min.)	Number of emotions	File level duration		Notes
				Mean (sec.)	Std. (sec.)	
Emo-DB	German	24.7	7	2.7	1.02	Acted
SAVEE	English	30.7	7	3.8	1.07	Acted
UUDB	Japanese	113.4	4	1.4	1.7	Non-acted

The *Emo-DB* emotional database (German) [12] was recorded at the Technical University of Berlin and consists of labelled emotional German utterances which were spoken by 10 actors (5 female). Each utterance has one of the following emotional labels: neutral, anger, fear, joy, sadness, boredom or disgust.

The *SAVEE* (Surrey Audio-Visual Expressed Emotion) corpus (English) [13] was recorded as a part of an investigation into audio-visual emotion classification from four native English male speakers. The emotional label for each utterance is one of the standard set of emotions (anger, disgust, fear, happiness, sadness, surprise and neutral).

The *UUDB* (The Utsunomiya University Spoken Dialogue Database for Paralinguistic Information Studies) database (Japanese) [14] consists of spontaneous Japanese human-

human speech. The task-oriented dialogue produced by seven pairs of speakers (12 female) resulted in 4,737 utterances in total. Emotional labels for each utterance were created by three annotators on a five-dimensional emotional basis (interest, credibility, dominance, arousal and pleasantness). For this work, only the pleasantness and arousal axes are used. The corresponding quadrant (anticlockwise, starting in the positive quadrant, and assuming arousal as abscissa) can also be assigned emotional labels: happy-exciting, angry-anxious, sad-bored and relaxed-serene.

There is a statistical description of the used corpora in Table 2.

Performance assessment

The effectiveness of the proposed feature selection technique was estimated in combination with three classification models [15]:

- 1) **Support Vector Machine (SMO)**. To design a hyperplane separating sets of examples Sequential Minimal Optimization (SMO) is used for solving the large scale quadratic programming problem.
- 2) **Multilayer Perceptron (MLP)**. A feedforward neural network with one hidden layer is trained with the error backpropagation algorithm (BP).
- 3) **Linear Logistic Regression (Logit)**. This linear model describes the relationship between labels and independent variables using probability scores.

In all experiments the F-score metric [16] was assessed to compare the quality of classification (the more effective the classifier used, the higher F-score value obtained). To derive more statistically significant results, the 6-fold cross-validation procedure was implemented for each database.

Firstly, all these classifiers were applied without any feature selection procedure at all. We obtained their predictions separately to assess the F-score value for each model. Then these classifiers were included in the ensemble of models and the final prediction was formed based on the majority rule.

Secondly, conventional MOGAs were engaged for feature selection. All algorithms were provided with the same amount of resources (90 generations and 150 individuals in populations). For each MOGA the following settings were defined: binary tournament selection, uniform recombination and the mutation probability $p_m=1/n$, where n is the length of the chromosome. In the final set of non-dominated points we had 30 binary strings. After the feature selection stage all of the classifiers mentioned above and their ensembles were applied.

In this paper we do not compare the proposed feature selection technique with some conventional methods because the main purpose is to accomplish the thorough investigation of MOGAs and their cooperative modifications as optimizers in the framework of the discussed approach. Some results related to the comparison with a number of conventional feature selection techniques might be found in [2].

Then, in the next experiments we used cooperative modifications of MOGAs to select informative features. We applied three homogeneous algorithms (NSGA-II – NSGA-II – NSGA-II; PICEA-g – PICEA-g – PICEA-g; SPEA2 – SPEA2 – SPEA2) and the heterogeneous one (NSGA-II – PICEA-g – SPEA2). For each MOGA all islands had an equal amount of resources (90 generations and $150/3 = 50$ individuals in populations), the migration size was equal to 10 (in total each island got 20 points from two others), and the migration interval was equal to 10 generations. The genetic operators were the same as in the previous experiment.

The results obtained are presented below in the diagrams (Figure 3-5).

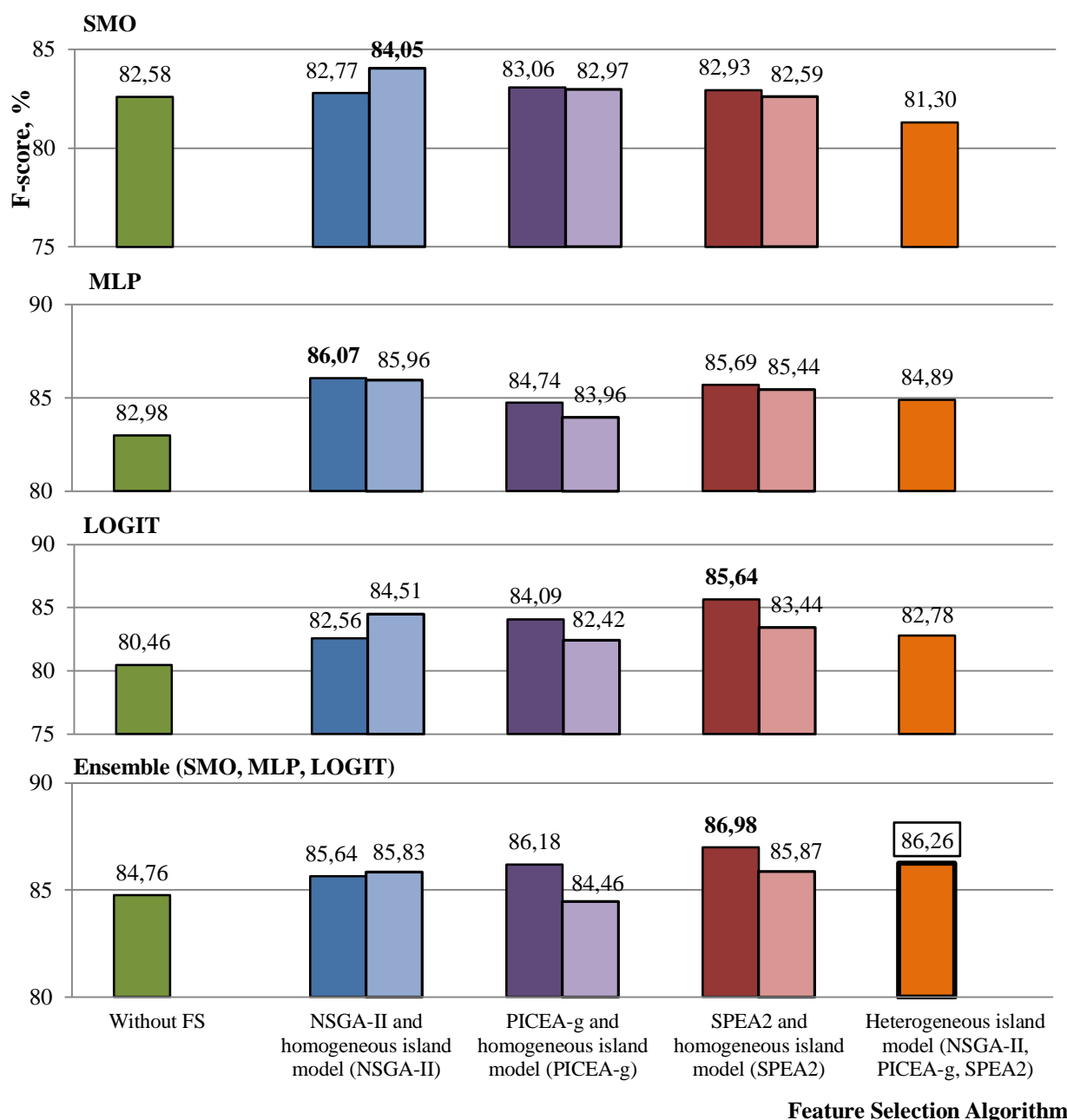


Fig. 3. The experimental results for the EMO-DB database

Firstly, these diagrams demonstrate that on full datasets the most effective classifiers for EMO-DB, SAVEE and UUDB are not the same (when they are applied separately). Moreover, we may find that the ensemble of classifiers (SMO, MLP, and LOGIT) outperforms any of these models for all of the corpora. Therefore, the problem of choosing the most effective model might be solved with the application of classifier ensembles.

Predominantly for EMO-DB and SAVEE the usage of MOGAs for selecting informative features leads to an improvement in F-score values. For UUDB in some cases we may observe a minor decrease in this metric, but it is not statistically significant (a t-test with a significance level $p=0.05$).

Talking about various categories of MOGAs (conventional, homogeneous and heterogeneous), we may note:

- for different corpora various combinations of classifiers (SMO, MLP, LOGIT) and conventional MOGAs are the best;

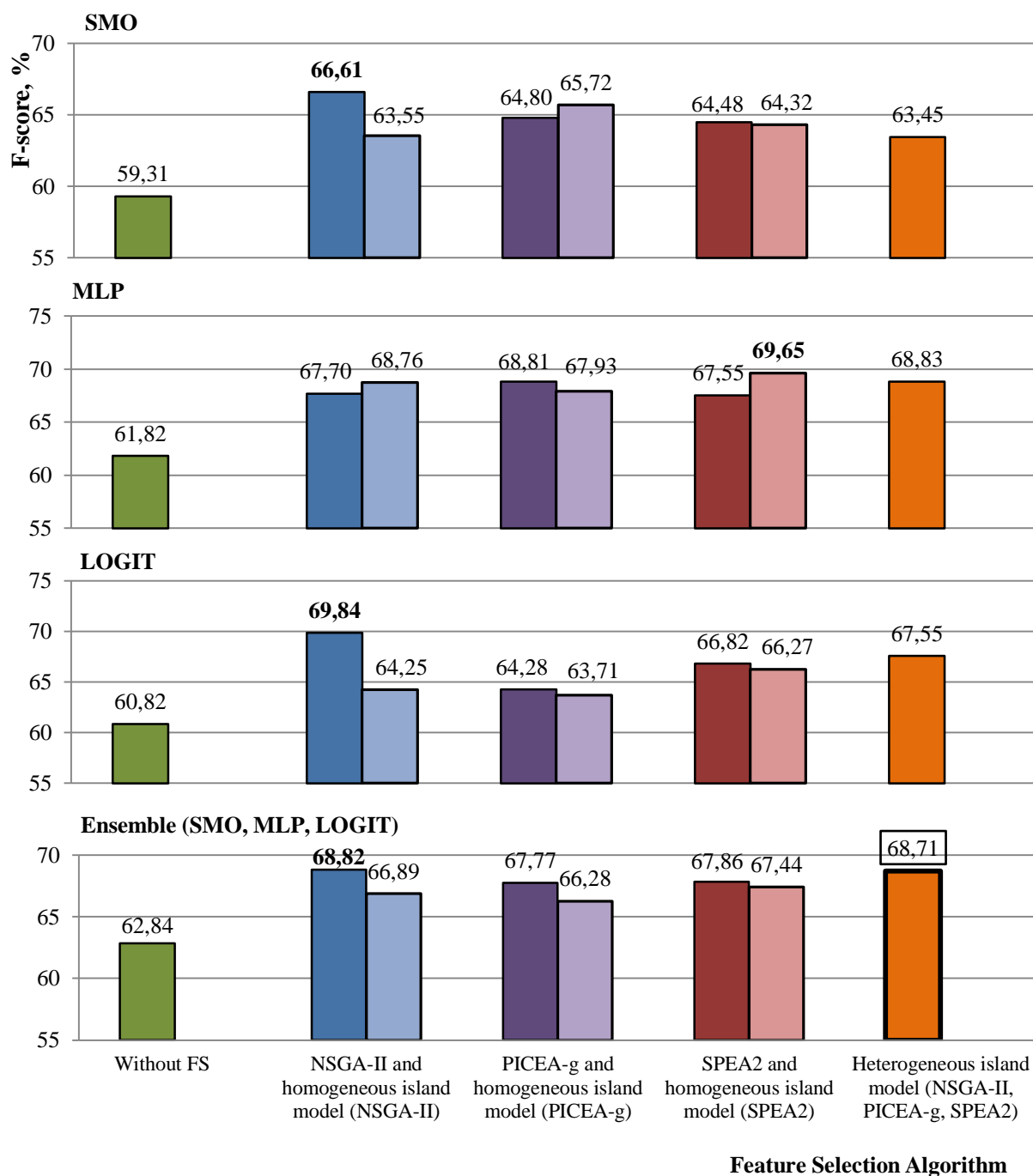


Fig. 4. The experimental results for the SAVEE database

- the homogeneous modification of a MOGA often demonstrates worse results in terms of F-score in comparison with the same conventional MOGA. However, due to the parallel structure, these algorithms require less time spent on their execution. On average, for all of the corpora time costs for feature selection are decreased roughly by a factor of 2.55 (the additional time is spent on the migration process);

- the heterogeneous MOGA might be effectively used in combination with the ensemble of classifiers, which allows us to avoid the choice of the most effective MOGA and the best classification model for the problem considered. For each corpus we compared the F-score value achieved by the heterogeneous MOGA and the ensemble of classifiers with the best value of this metric obtained with any MOGA and any classifier. As a result, we revealed that this difference is not significant statistically (a t-test with a significance level $p=0.05$).

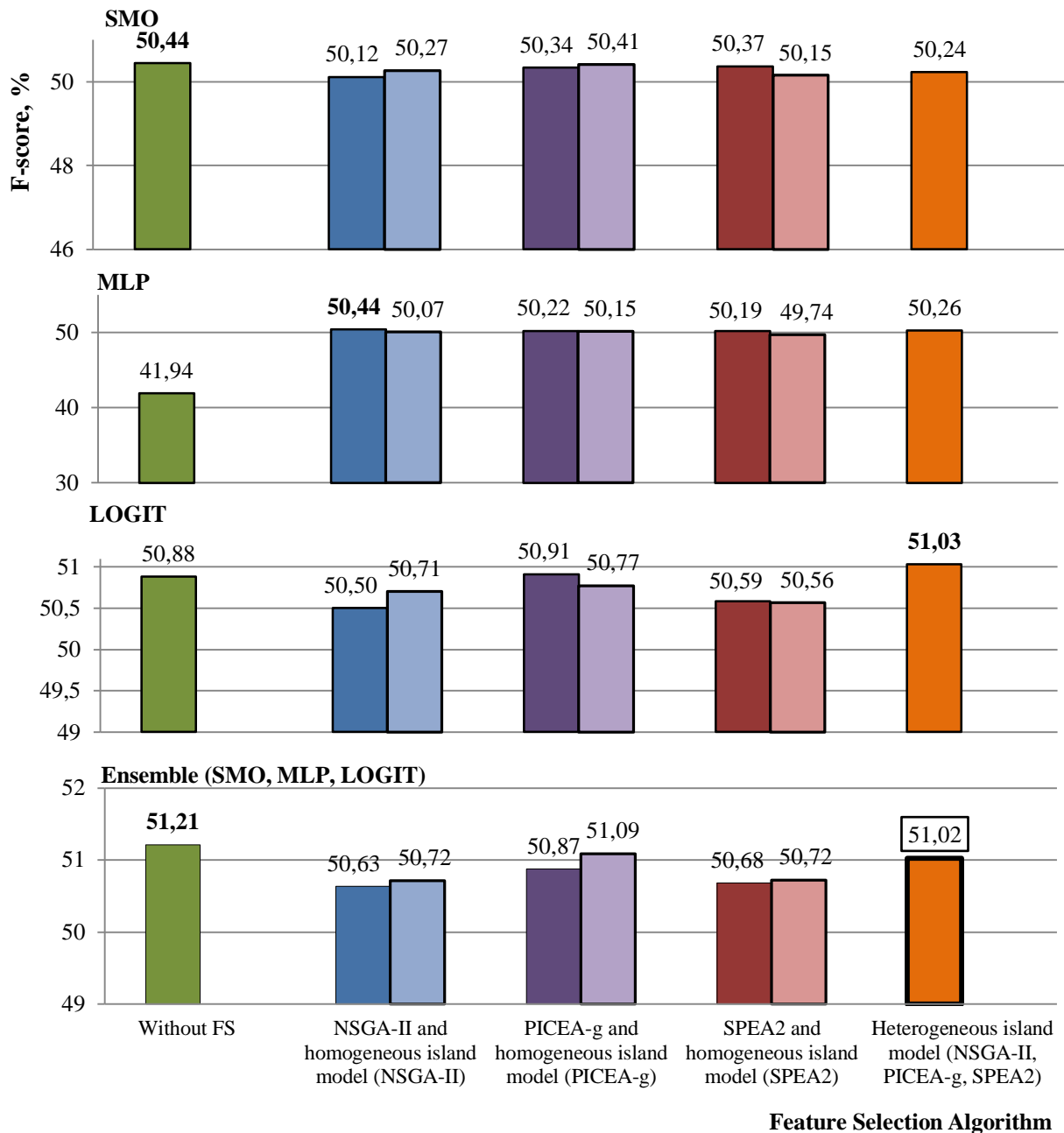


Fig. 5. The experimental results for the UUDB database

Generally, for diverse MOGAs the average number of selected features in the reduced databases varies: for EMO-DB from 159.5 to 180.9, for SAVEE from 162.0 to 186.1, for UUDB from 139.1 to 167.5 (initially there are 384 features).

Conclusion

Based on the results obtained, we may conclude that the proposed feature selection technique is an effective alternative to the conventional PCA because in most cases the application of any MOGA leads to an improvement in the classifier performance and a significant dimensionality reduction.

We investigated the effectiveness of different MOGAs and their modifications in the framework of the approach presented and found that it was reasonable to use the heterogeneous MOGA and the ensemble of classifiers to eliminate the choice of the most

effective heuristic algorithm and the best model without detriment to the classification quality.

In comparison with conventional MOGAs, homogeneous modifications are often preferable only in the sense of time costs, whereas the heterogeneous one shows high F-score values, especially in combination with the ensemble of classifiers.

Finally, the promising results prove that the proposed algorithmic scheme might be applied to solve some other problems related to the speech-based recognition of human qualities such as gender or speaker identification.

References

1. Brester C., Semenkin E., Sidorov M., Kovalev I., Zelenkov P. Evolutionary feature selection for emotion recognition in multilingual speech analysis. *Proceedings of the IEEE Congress on Evolutionary Computation (CEC2015)*, Sendai, Japan, 2015, p. 2406–2411.
2. Sidorov M., Brester Ch., Schmitt A. Contemporary stochastic feature selection algorithms for speech-based emotion recognition. *Proceedings of INTERSPEECH 2015*, Dresden, Germany, *in press*.
3. Kohavi R., John G.H. Wrappers for feature subset selection. *Artificial Intelligence*, 97, 1997, p. 273–324.
4. Venkatadri M., Srinivasa Rao K. A multiobjective genetic algorithm for feature selection in data mining. *International Journal of Computer Science and Information Technologies*, vol. 1, no. 5, 2010, p. 443–448.
5. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6 (2), 2002, p. 182–197.
6. Wang R. Preference-inspired co-evolutionary algorithms. A thesis submitted in partial fulfillment for the degree of the Doctor of Philosophy, University of Sheffield, 2013, p. 231.
7. Zitzler E., Laumanns M., Thiele L. SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization. *Evolutionary Methods for Design Optimisation and Control with Application to Industrial Problems EUROGEN 2001*, 3242 (103), 2002, p. 95–100.
8. Whitley D., Rana S., and Heckendorn R. Island model genetic algorithms and linearly separable problems. *Proceedings of AISB Workshop on Evolutionary Computation*, Manchester, UK. Springer, volume 1305 of LNCS, 1997, p. 109–125.
9. Brester Ch., Semenkin E. Cooperative Multi-objective Genetic Algorithm with Parallel Implementation. *Advances in Swarm and Computational Intelligence*, LNCS 9140, 2015, p. 471–478.
10. Boersma P. Praat, a system for doing phonetics by computer. *Glott international*, vol. 5, no. 9/10, 2002, p. 341–345.
11. Eyben F., Wöllmer M., and Schuller B. Opensmile: the munich versatile and fast opensource audio feature extractor. *Proceedings of the international conference on Multimedia*, 2010. ACM, p. 1459–1462.
12. Burkhardt F., Paeschke A., Rolfes M., Sendlmeier W.F., and Weiss B. A database of German emotional speech. *In Interspeech*, 2005, p. 1517–1520.
13. Haq S., Jackson P. Machine Audition: Principles, Algorithms and Systems. *Chapter Multimodal Emotion Recognition*, IGI Global, Hershey PA, Aug. 2010, p. 398–423.
14. Mori H., Satake T., Nakamura M., and Kasuya H. Constructing a spoken dialogue corpus for studying paralinguistic information in expressive conversation and analyzing its statistical/acoustic characteristics. *Speech Communication*, 53, 2011.
15. Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I. H. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, Vol. 11, Issue 1, 2009.

16. Goutte C., Gaussier E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. *ECIR'05 Proceedings of the 27th European conference on Advances in Information Retrieval Research*, 2005, p. 345–359.

Библиографические ссылки

1. Brester C., Semenkin E., Sidorov M., Kovalev I., Zelenkov P. Evolutionary feature selection for emotion recognition in multilingual speech analysis // *Proceedings of the IEEE Congress on Evolutionary Computation (CEC2015)*, Sendai, Japan. 2015. P. 2406–2411.

2. Sidorov M., Brester Ch., Schmitt A. Contemporary stochastic feature selection algorithms for speech-based emotion recognition // *Proceedings of INTERSPEECH 2015*, Dresden, Germany, in press.

3. Kohavi R., John G.H. Wrappers for feature subset selection // *Artificial Intelligence*, 97. 1997. P. 273-324.

4. Venkatadri M., Srinivasa Rao K. A multiobjective genetic algorithm for feature selection in data mining // *International Journal of Computer Science and Information Technologies*, vol. 1, no. 5. 2010. P. 443–448.

5. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II // *IEEE Transactions on Evolutionary Computation* 6 (2). 2002. P. 182–197.

6. Wang R. Preference-inspired co-evolutionary algorithms. A thesis submitted in partial fulfillment for the degree of the Doctor of Philosophy, University of Sheffield. 2013. P. 231.

7. Zitzler E., Laumanns M., Thiele L. SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization // *Evolutionary Methods for Design Optimisation and Control with Application to Industrial Problems EUROGEN 2001*, 3242 (103). 2002. P. 95–100.

8. Whitley D., Rana S., and Heckendorn R. Island model genetic algorithms and linearly separable problems // *Proceedings of AISB Workshop on Evolutionary Computation*, Manchester, UK. Springer, volume 1305 of LNCS. 1997. P. 109-125.

9. Brester Ch., Semenkin E. Cooperative Multi-objective Genetic Algorithm with Parallel Implementation // *Advances in Swarm and Computational Intelligence*, LNCS 9140. 2015. P. 471–478.

10. Boersma P. Praat, a system for doing phonetics by computer // *Glott international*, vol. 5, no. 9/10. 2002. P. 341–345.

11. Eyben F., Wöllmer M., and Schuller B. Opensmile: the munich versatile and fast opensource audio feature extractor // *Proceedings of the international conference on Multimedia*. 2010. ACM, P. 1459–1462.

12. Burkhardt F., Paeschke A., Rolfes M., Sendlmeier W.F., and Weiss B. A database of German emotional speech // *In Interspeech*. 2005. P. 1517–1520.

13. Haq S., Jackson P. Machine Audition: Principles, Algorithms and Systems. Chapter Multimodal Emotion Recognition // *IGI Global*, Hershey PA. Aug. 2010. P. 398–423.

14. Mori H., Satake T., Nakamura M., and Kasuya H. Constructing a spoken dialogue corpus for studying paralinguistic information in expressive conversation and analyzing its statistical/acoustic characteristics // *Speech Communication*, 53. 2011.

15. Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I. H. The WEKA Data Mining Software: An Update // *SIGKDD Explorations*, Vol. 11, Issue 1. 2009.

16. Goutte C., Gaussier E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation // *ECIR'05 Proceedings of the 27th European conference on Advances in Information Retrieval Research*. 2005. P. 345–359.