

Signal Selection for Sleep Apnea Classification

Yashar Maali and Adel Al-Jumaily

University of Technology, Sydney (UTS)

Faculty of Engineering and IT, Sydney, Australia

Yashar.Maali@student.uts.edu.au, Adel.Al-Jumaily@uts.edu.au

Abstract. This paper presents a method for signals and features selection when classifying sleep apnea. This study uses a novel hierarchical parallel particle swarm optimization structure as proposed by the authors previously. In this structure, the swarms are separated into ‘masters’ and ‘slaves’ and access to global information is restricted according to their types. This method is used to classify sleep apneic events into apnea or hypopnea. In this study, ten different biosignals are used as the inputs for the system albeit with different features. The most important signals are subsequently determined based on their contribution to classification of the sleep apneic events. The classification method consists of three main parts which are: feature generation, signal selection, and data reduction based on PSO-SVM, and the final classifier. This study can be useful for selecting the best subset of input signals for classification of sleep apneic events, by attention to the trade of between more accuracy of higher number of input signals and more comfortable of less signals for the patient.

Keywords: Sleep apnea, Particle swarm optimization, Support vector machines, Parallel processing.

1 Introduction

Sleep Apnea (SA) is one of the most common and critical components of sleep disorders. SA is characterized by the repeated temporary cessation of breathing during sleep [1]. Clinically, apnea is defined as the total or near-total absence of airflow. This becomes significant once the reduction of the breathing signal amplitude is at least around 75% with respect to the normal respiration and occurs for a period of 10 seconds or longer. A hypopnea is an event of comparatively less intensity; it is defined as a reduction in baseline of the breathing signal amplitude to around 30–50%, also lasting 10 seconds in adults [2].

The SA has several short term and long term side effects[3]. Short-term effects lead to impaired attention and concentration, reduce quality of life, increased rates of absenteeism with reduced productivity, and increased the possibility of accidents at work, home or on the road. Long-term consequences of sleep deprivation include increased morbidity and mortality from increasing automobile accidents, coronary artery disease, heart failure, high blood pressure, obesity, type 2 diabetes mellitus, stroke and memory impairment as well as depression. Long-term consequences, however, remain open[4].

Unfortunately, as many patients are asymptomatic, sleep apnea may go undiagnosed for years [5, 6]. Usually it is patients’ spouses, roommates, or family members who

report the apnea periods alternating with arousals and accompanied by loud snoring [7, 8]. Symptomatic patients with SA are usually assessed by Sleep Medicine Specialists and diagnosed through an overnight sleep study in a sleep clinic. SA is diagnosed by a manual analysis of a polysomnographic record, an integrated device comprising of the EEG, EMG, EOG, ECG, and oxygen saturation (SPO2) [9]. The polysomnography also records the airflow through the mouth and nose, along with the thoracic and abdominal effort signals [10], and the position of the body during sleep. A sleep apnea event can also be classified into three groups as: central sleep apnea, obstructive sleep apnea, and mixed sleep apnea. In case of the first, sleep apnea is originated by the central nervous system. In the case of the second, the reason for the pauses in the breathing lie in a respiratory tract obstruction, while in the third case, both of these reasons may be present.

The conventional scoring of the polysomnographic recording is costly and time-consuming. Therefore, many efforts have been made to develop systems that score the records automatically [11-13]. For this reason several Artificial Intelligent (AI) algorithms are used in this area such as; fuzzy rule-based system [14], genetic SVM [15], and PSO-SVM [16] which have been proposed in our previous works.

Classification of apneic events to apnea or hypopnea is also so important for severity calculation of the sleep disorder. The classification of apneic events is also considered in many studies, such as [17-19]. But, up to the authors' knowledge, there has no publish work on impact and importance of different signals in sleep apnea classification. In this paper, we rank most important signals for the classification of sleep disorder events into apnea or hypopnea by a hierarchical parallel PSO-SVM[16]. The second section of this work covers details of the parallel model. We also introduce the PSO-SVM model in the third section of this paper, which is followed by experimental results in section four, and the conclusion in section five.

2 Parallel PSO

Particle Swarm Optimization (PSO), was introduced by Kennedy and Eberhart in 1995 [20, 21] based on the movement of swarms and inspired by the social behaviors of birds or fishes. Similar to the genetic algorithm, PSO is a population-based stochastic optimization technique. In the PSO, each member is named particle, and each particle is a potential solution to the problem. In comparison with the genetic algorithms, PSO updates the population of particles by considering their internal velocity and position, which are obtained by the experience of all the particles.

2.1 Proposed Parallel Structure

In this work, given the enormous size of the search space, a single PSO cannot perform well and may lead to the local optimum with low accuracy. Therefore, we have used a new PSO parallel structure[16], to perform better explorations and exploitations in the search space.

In the traditional multi-PSO models, all of the swarms are at the same level and exchanging information is just based on the definition of neighborhood [22]. But, in this

structure, swarms are classified at two different levels as ‘masters’ and one (or more) ‘slave(s)’. Master swarms have access to the best particle of other swarms but the slave swarms have no access to others’ information, and actually just provide information for others. Sending the best local particle information among the masters and from the slave(s) to the masters can entail performing each iteration by a specified frequency.

In this hierarchical model, one of the master swarms is considered to be the centre swarm. All of the swarms, masters or slaves, send the local best particle to the centre swarm. The centre swarm computes the global best particle and sends it to the other master swarms. So all of the master swarms update their particles using the global best particle, but the slave swarms only use their own local best particles for updating themselves. The pseudo code for the hierarchical multi-swarm PSO is as follows.

Begin

Select the number of master and slave swarms, number of the particles for each sub-swarm along with the frequency for the sending of the information. Select one of the master swarms as the centre.

Initialize the position and velocity of each particle

Do in parallel until the maximum number of iterations has reached {

Evaluate the fitness value of each particle

Find out the local best particle in each sub-swarm

If meet sending condition

Sending the local best particle (lp_{best}) from each swarm to the centre swarm.

Updating global best particle (Gp_{best}) in the centre swarm.

Sending the global best particle to the master swarms.

End If

Calculate the new velocity of each particle in each sub-swarm

Update the position of each particle in each swarm

End Do}

Return the best solution (the global best particle) of the algorithm

End

To provide more information how the model work; figure 1 illustrate a sample of the proposed parallel structure with 4 masters and 2 slave swarm. In this implementation, master 1 is selected as the center swarm, so all of the swarms send their local

best particles to this swarm. After computing the global best swarm, center swarm sends it to all of the master swarms. So in this structure, slave swarm provides information for other swarms, but not gets benefit from the other's information.

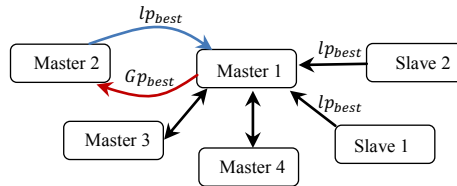


Fig. 1. Proposed parallel structure with 4 masters and two slaves

Indeed, fast convergence is one of the disadvantages of the PSO, which is heightened in the parallel structure. This hierarchical model tries to expand both of the exploration and exploitation abilities of the parallel PSO, by integrating the isolated swarms and the linked swarms. Slave swarm(s) in this model help prevent premature convergence. PSO parameters for the slaves and masters can be different. Therefore, it helps to have greater exploration abilities for the slaves and a greater local search for the masters.

3 Approach and Method

In this section, we present the proposed algorithm for the ranking of input signals by attention to their impact on classification of the sleep apneic events into apnea or hypopnea. The proposed methodology is as follows:

- Feature generation: this stage generates several statistical features for each event from the wavelet packet coefficients.
- Parallel PSO-SVM: this stage uses the PSO algorithm to select a best signal and features subset interactively with the SVM. PSO also is used for tuning the parameters of the SVM. In the process, SVM is used as the fitness evaluator. This PSO-SVM combination is applied in parallel by the new architecture to achieve better performance and avoid of the local optimal solutions.
- Final classification: the selected pattern is used for classification of the unseen validation data in this stage. The accuracy of this step is assumed as the final performance of the algorithm.

The details of these steps are as follows:

3.1 Features Generation

As mentioned previously, we used 10 biosignal from the polysomnographic record. Table 1 lists these signals and their frequencies.

Table 1. List of signals that are used in this study

	Signal	Frequency
#1	EEG C3A2	100
#2	EEG C4A1	100
#3	EEG A1A2	100
#4	EOG	100
#5	EMG chin	200
#6	ECG	200
#7	Air flow	10
#8	Thoracic movement	10
#9	Abdominal movement	10
#10	SPO2	1

For feature generation, First of all, 10 seconds after of each apneic event is selected from each signal. Then, we used wavelet package coefficients to generate features for each of these signals. For example for airflow, abdominal, thoracic movement, and SPO2, we generated the features by applying statistical measures from Table 2 to each coefficient of the 4 levels Daubechies wavelet packet with order 3 [23].

Table 2. List of statistical features, x is coefficients of wavelet packet

$\log(\text{mean}(x^2))$	$\text{kurtosis}(x^2)$	$\text{geomean}(x)$
$\text{std}(x^2)$	$\text{var}(x^2)$	$\text{mad}(x)$
$\text{skewness}(x^2)$	$\text{mean}(x)$	$\text{mean}(x^2)$
$\text{skewness}(x)$	$\text{kurtosis}(x)$	$\text{var}(x)$
$\text{geomean}(x^2)$	$\text{mad}(x^2)$	$\text{std}(x)$

The details of feature generation for other signals are as follows:

A. EEG

Based on [24], we applied the Daubechies order 2 (db2) wavelet transform. The frequency ranges of the EEG signal are broken down into Delta (below 3.5 Hz), Theta (4-7 Hz), Alpha (8-13 Hz), and Beta (14-30 Hz) bands [24]. Then, the following

statistical features were used to represent the time–frequency distribution of the EEG signals:

1. Mean quadratic value or Energy of wavelet packet (WP) coefficients for each of the sub bands,
2. Total Energy,
3. Ratio of different Energy values,
4. Mean of the absolute values of the coefficients in each sub band, and
5. Standard deviation of the coefficients in each sub band.

Also more features are generated for EEG signals based on [25]. Given this reason, we apply a level 4 wavelet packet of Daubechies by order 20 and then, generate the features for each coefficient displayed in Table 2.

B. EOG

To generate the features from the EOG signal, we apply an 8 level Daubechies with order 3 wavelet packet. Then, we generate the features for each coefficient according to Table 2, for each of the left and right EOG signals. Also, we calculate the normalized correlation coefficient between the two EOG channels [26].

C. EMG

For the EMG signal based on [23], we generate features by applying statistical measures from Table 2 to the coefficients of level 2 Daubechies with the order 2.

D. ECG

For feature generation of ECG signals based on [27], first of all we apply an order 3 Daubechies wavelet packet with 8 level. Then, for each event by attention to the wavelet coefficients features are generated by the statistical measures from table 2.

By attention to these ten signals and described features, 205 features are generated for each apneic event. In next step the parallel PSO-SVM is used for feature selection to eliminate the irrelevant features and therefore improving the accuracy of the classification.

3.2 Parallel PSO-SVM Algorithm

After generating the features, we need to classify events as apnea or hypopnea. For this reason, we use the described hierarchical parallel PSO-SVM algorithm to initiate pattern selection and tune the parameters of the SVM.

SVM is used to evaluate each particle of the PSO. The fitness of each particle is computed as the accuracy of the SVM when classifying the test set with the selected pattern related to the particle.

In the first step, total data from all of the samples are integrated as Meta data. Then, the Meta data is separated into the train and the validation. The hierarchical parallel PSO-SVM is used using the 5*10 CV paradigms to select the best signals and features sets from the training data. Then, the selected signals and features are used to classify the validation set. Accuracy of classification of validation set by the tuned SVM and selected patters is considered as performance of the system.

3.3 PSO Structure

In this study, we use the constriction coefficient PSO [28]. In this approach, the velocity update equation is as (1),

$$v_{ij}(t+1) = \chi \left[v_{ij}(t) + \phi_1 (y_{ij}(t) - x_{ij}(t)) + \phi_2 (\hat{y}_{ij}(t) - x_{ij}(t)) \right], \quad (1)$$

where y_{ij} is the particle best and $\hat{y}_{ij}$ is the global best particles. And,

$$\chi = \frac{2k}{|2 - \phi - \sqrt{\phi(\phi - 4)}|}, \quad (2)$$

with,

$$\phi = \phi_1 + \phi_2, \quad \phi_i = c_i r_i \quad i = 1, 2.$$

Equation (2) is used under the constraints that $\phi \geq 4$ and $k, r_i \in [0, 1]$.

The parameter k in the equation (2) controls the exploration and exploitation. For $k \sim 0$, fast convergence is expected and for $k \sim 1$ we can expect slow convergence with a high degree of exploration [28].

The constriction approach has several advantages over traditional PSO model such as; we do not need velocity clamping for constriction model and this model guarantees convergence under the given constraints[29].

By paying attention to the proposed parallel structure, for the slave swarms, k considered as 0.8 and $c_1 = 2, c_2 = 4$, and for the master swarms k considered as 0.2 and $c_1 = 4, c_2 = 2$.

3.4 Particle Representation

In this study, each particle consists of three arrays; the length of the first array is equal to the number of signals, which equal ten. Each cell can get a real number between 0 and 1 as importance of the relevant signal. Rank of each signal is computed by the rank of the corresponding cell.

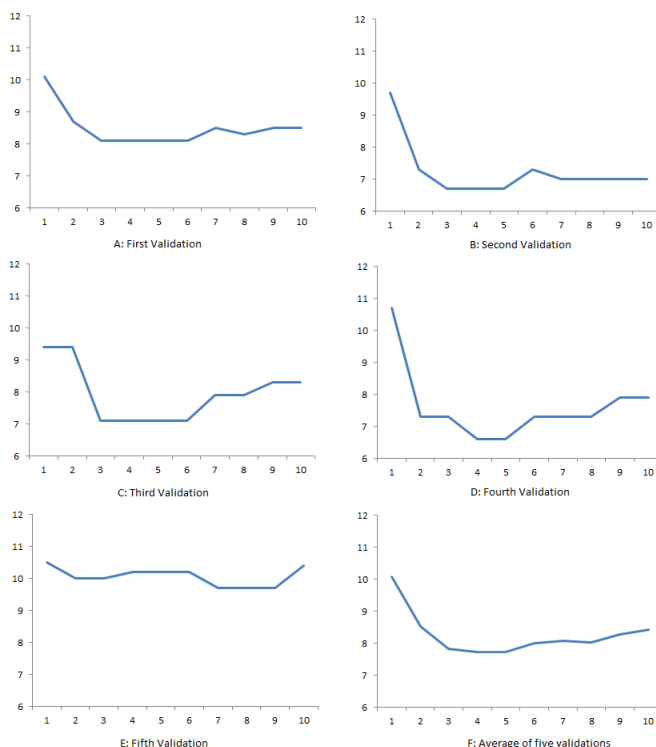
The length of the second array is equal to the number of the features. Again each cell can get a real number between 0 and 1. Features, which their corresponding cells have values higher than 0.5, are selected for classification. The third array is related to the gamma and cost as parameters of the SVM, which can get a value between 2^{-5} to 2^5 .

4 Results and Discussion

Experimental data consist of polysomnographic records from 20 subjects, which events of them are annotated by experts were provided by concord hospital in Sydney. In the parallel structure, 4 slaves and two masters are selected and each swarm contain 20 particles. The RBF kernel is selected for the SVM in both of the master and slave swarms. Frequency for changing information between swarms is set as 5 iterations.

Table 3. Diversity of classes in different runs

	Validation		Training	
	Apnea	Hypopnea	Apnea	Hypopnea
#1	833	636	1632	1309
#2	780	683	1685	1262
#3	865	701	1600	1244
#4	874	610	1591	1335
#5	787	601	1678	1344

**Fig. 2.** Classification error curve using the signal selection method

In this experiment records from 5 subjects are considered as the validation and records from 15 subjects make the training data. It also take in consideration to overcome the impact of the validation sets on the final result, 5 independent experiments had validated. Table 3 tabulates the number of sleep apnea or hypopnea events in each of the validation set and the training in these 5 runs.

We investigated whether the best way to select signals for apnea or hypopnea classification. First we consider the best size for the input signals' set. The results are depicted in Figure 2. The first five plots, A to E, show the individual classification

error for the five validations against the different numbers of signals chosen. The sixth plot, figure F, shows the classification error averaged over the five validations.

As can be seen in Figure 2, using irrelevant signals increased the classification error. For example, using 3 channels for validation 1, figure A, yields the least error. The result averaged over the five validations plotted in the last figure, figure F. This figure also shows that when removing channels iteratively the classification error decreases slightly until all irrelevant channels are removed. The average error rate (taken over all validations) of 7.84% using 3 signals is very close to the error of the baseline which is 7.74% using 5 signals. This can show that increasing the number of input signals from 3 to 5 has not high impact on the performance. Therefore, by attention to the cost of using more signals and on the other side slight improvement in the performance we can decide the preferred size of the input signals.

Table 4. Ranking of 10 signals

Rank	Validations				
	1	2	3	4	5
1	Airflow	Airflow	Airflow	Airflow	Airflow
2	Abdominal	Abdominal	Abdominal	Abdominal	Abdominal
3	<u>Thoracic</u>	<u>EMG chin</u>	<u>Thoracic</u>	Thoracic	Thoracic
4	EMG chin	Thoracic	EMG chin	<u>EMG chin</u>	EMG chin
5	EEG C4A2	ECG	EEG C4A2	ECG	EEG C4A2
6	EEG C4A1	EEG C4A2	ECG	EEG C4A1	ECG
7	ECG	EEG C4A1	EEG C4A1	EEG C4A2	<u>EEG C4A1</u>
8	EOG	EOG	SPO2	SPO2	EOG
9	SPO2	SPO2	EOG	EOG	SPO2
10	EEG C3A2	EEG C3A2	EEG C3A2	EEG C3A2	EEG C3A2

Table 4 contains signal rankings; as we have the three most common signals are airflow, abdominal and thoracic movement. For each validation we can obtain a heuristic estimate on the number of irrelevant signals from the classification error curves in Figure 1. We underlined one entry in each column of Table 4. The row number of that entry is an estimate for the ranks position of the best set of signals for classification of that validation set. For example in validation 4, the local minimum of the classification error happened by 4 signals. Thus, the best 4 selected signals can be used instead of the full data without increasing the error. Selecting 4 signals instead of the whole polysomnographic records is important because less signals means less recording cost and also more comfortable for the patients.

5 Conclusion

In this study we ranked the polysomnographic signals by attention to their impact on classification of sleep apneic events into apnea or hypopnea. We used a hierarchical parallel PSO-SVM for the signal selection. This study shows that it is possible to reduce the number of signals for a robust classification without an increase of

classification error. Average of results from 4 different experiments shows, using 4 signals can provide the least classification error. But, the difference between classification error of the best 3 signals and the best 4 signals is less than 0.10%. This comparison is important as using less signals means more comfortable for the patients. Therefore, trade-off between selecting more signals and its cost must be considered. The same study for detecting apnea events and also classification of apnea events to central, obstructive or mixed will be subject to future research.

References

1. Guilleminault, C., van den Hoed, J., Mitler, M.: Overview of the sleep apnea syndromes. In: Guilleminault, C., Dement, W.C. (eds.) *Sleep Apnea Syndromes*, pp. 1–12. Alan R Liss, New York (1978)
2. Flemons, W.W., et al.: Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. *Sleep* 22(5), 667–689 (1999)
3. Chokroverty, S., et al.: Sleep deprivation and sleepiness. In: *Sleep Disorders Medicine*, 3rd edn., pp. 22–28. W.B. Saunders, Philadelphia (2009)
4. Chokroverty, S.: Overview of sleep & sleep disorders. *Indian Journal of Medical Research* 131(2), 126–140 (2010)
5. Ball, E.M., et al.: Diagnosis and treatment of sleep apnea within the community - The Walla Walla project. *Archives of Internal Medicine* 157(4), 419–424 (1997)
6. Kryger, M.H., et al.: Utilization of health care services in patients with severe obstructive sleep apnea. *Sleep* 19(9), S111–S116 (1996)
7. Stradling, J.R., Crosby, J.H.: Relation between systemic hypertension and sleep hypoxemia or snoring- analysis in 748 men drawn from general-practice. *British Medical Journal* 300(6717), 75–78 (1990)
8. Hoffstein, V.: Snoring. In: Kryger, M.H., Roth, T., Dement, W.C. (eds.) *Principles and Practice of Sleep Medicine*, pp. 813–826. Saunders, Philadelphia (2000)
9. Penzel, T., et al.: Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. *Medical & Biological Engineering & Computing* 40(4), 402–407 (2002)
10. Kryger, M.H.: Management of obstructive sleep-apnea. *Clinics in Chest Medicine* 13(3), 481–492 (1992)
11. Cabrero-Canosa, M., Hernandez-Pereira, E., Moret-Bonillo, V.: Intelligent diagnosis of sleep apnea syndrome. *IEEE Engineering in Medicine and Biology Magazine* 23(2), 72–81 (2004)
12. Cabrero-Canosa, M., et al.: An intelligent system for the detection and interpretation of sleep apneas. *Expert Systems with Applications* 24(4), 335–349 (2003)
13. de Chazal, P., et al.: Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea. *IEEE Transactions on Biomedical Engineering* 50(6), 686–696 (2003)
14. Maali, Y., Al-Jumaily, A.: Genetic Fuzzy Approach for detecting Sleep Apnea/Hypopnea Syndrome. In: 2011 3rd International Conference on Machine Learning and Computing, ICMLC 2011 (2011)
15. Maali, Y., Al-Jumaily, A.: Automated detecting sleep apnea syndrome: A novel system based on genetic SVM. In: 2011 11th International Conference on Hybrid Intelligent Systems, HIS (2011)

16. Yashar, M., Adel, A.-J.: A Novel Partially Connected Cooperative Parallel PSO-SVM Algorithm Study Based on Sleep Apnea Detection. IEEE Congress on Evolutionary Computation 2012, Brisbane, Australia (Accepted in, 2012)
17. Schluter, T., Conrad, S.: An approach for automatic sleep stage scoring and apnea-hypopnea detection. *Frontiers of Computer Science* 6(2), 230–241 (2012)
18. Aksahin, M., et al.: Artificial Apnea Classification with Quantitative Sleep EEG Synchronization. *Journal of Medical Systems* 36(1), 139–144 (2012)
19. Guijarro-Berdinas, B., Hernandez-Pereira, E., Peteiro-Barral, D.: A mixture of experts for classifying sleep apneas. *Expert Systems with Applications* 39(8), 7084–7092 (2012)
20. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: *Proceedings of the Sixth International Symposium on Micro Machine and Human Science, MHS 1995* (1995)
21. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: *1995 IEEE International Conference on Neural Networks Proceedings*, vol. 1-6, pp. 1942–1948 (1995)
22. Fan, S.K.S., Chang, J.M.: Dynamic multi-swarm particle swarm optimizer using parallel PC cluster systems for global optimization of large-scale multimodal functions. *Engineering Optimization* 42(5), 431–451 (2010)
23. Kiatpanichagij, K., Afzulpurkar, N.: Use of supervised discretization with PCA in wavelet packet transformation-based surface electromyogram classification. *Biomedical Signal Processing and Control* 4(2), 127–138 (2009)
24. Ebrahimi, F., et al.: Automatic Sleep Stage Classification Based on EEG Signals by Using Neural Networks and Wavelet Packet Coefficients. In: *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1-8, pp. 1151–1154 (2008)
25. Ebrahimi, F., et al.: Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients. In: *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2008* (2008)
26. Kempfner, J., et al.: Automatic REM Sleep Detection Associated with Idiopathic REM Sleep Behavior Disorder. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6063–6066 (2011)
27. Gubbi, J., Khandoker, A., Palaniswami, M.: Classification of sleep apnea types using wavelet packet analysis of short-term ECG signals. *Journal of Clinical Monitoring and Computing* 26(1), 1–11 (2012)
28. Clerc, M.: The swarm and the queen: towards a deterministic and adaptive particle swarm optimization. In: *Proceedings of the 1999 Congress on Evolutionary Computation, CEC 1999* (1999)
29. Engelbrecht, A.P.: *Fundamentals of Computational Swarm Intelligence*. Wiley (2005)