ASIAN JOURNAL OF
INFORMATION TECHNOLOGY

MEDWELL PUBLICATIONS SINCE 2000

Optimizing Classification of Spread Spectrum Signals Based on Features Extraction

[1]Haidy S. Fouad, [2]Hend A. Elsayed and [1]Shawkat K. Girgis

[1]*Department of Information Technology, Institute of Graduate Studies and Research, Alexandria University, Alexandria, Egypt*

[2]*Department of Communication and Computer Engineering, Delta University for Science and Technology, Mansoura, Egypt*

Key words: Binary optimization, Gray Level Co-occurrence Matrix (GLCM), Direct Sequence Spread Spectrum (DSSS), Frequency Hopped Spread Spectrum (FHSS), clustering, sequential selection, Whale Optimization (WOA)

Corresponding Author:
Haidy S. Fouad
Department of Information Technology, Institute of Graduate Studies and Research, Alexandria University, Alexandria, Egypt

Page No.: 1-11
Volume: 19, Issue 1, 2020
ISSN: 1682-3915
Asian Journal of Information Technology

Abstract: Spread Spectrum techniques (SS) attract the attention they are widely used in wireless communications and radar. Direct sequence spread spectrum and frequency hopped spread spectrum are the two most used techniques. In this research, the clustering of these signals is performed by feature-based model. Features are extracted by Gray Level Co-occurrence Matrix (GLCM), gray level run-length matrix, cumulants, moments, PCA, KPCA and Fast-ICA features. Clustering by GLCM features gets the best result which is one of the common textures features extraction techniques. The selection of relevant features is the big challenge. Therefore, the main contribution is to optimize the SS identification based on clustering techniques by decreasing the number of features without accuracy degradation which is based on filters, sequential forward selection and binary metaheuristic search strategies such as Binary Particle Swarm Optimization (BPSO), Genetic Algorithm (GA), bat feature selection (BBA) and hybrid Whale Optimization Algorithm with Simulated Annealing and Tournament mechanism (WOASAT). BPSO as a wrapper method is proposed to the optimization as it outperforms the other techniques in terms of accuracy and selected features with k-means, k-medoids or HAC.

INTRODUCTION

In spread spectrum techniques, the utilization of spreading sequences have the upside of security by its randomness. Low probability of intercept, anti-jamming, anti-interference and modulation by Code Division Multiple Access (CDMA) all these advantages make SS signals the favorable choice to wireless systems[1]. The expanding interest for wireless systems raised a challenge of efficient spectrum utilization. Additionally, there is a new generation of promising solutions that implement wireless radio communication signals like Software-Defined Radio (SDR), Cognitive Radio (CR) and Global Positioning Systems (GPS). The first concern for successful CR and GPS is the signal identification[2], the optimal use of the available radio-frequency is considered an optimization problem. The most used and practical Spread Spectrum signals are Direct Sequence (DSSS) and Frequency Hopping Systems (FHSS)[3]). Some practical use of DSSS is in IEEE 802.11b and IEEE 802.15.4 while FHSS is used in Bluetooth.

Feature selection: Dealing with large number of features generates a common problem which is the curse of dimensionality[4]. Therefore, features selection is a hot research topic[5]. GLCM features are employed in satellite images such as energy, entropy, correlation and contrast based on their irrelevant[6]. Many ways are deployed to construct a subset of features. Filters have a variety of evaluation measures such as distance, information, dependency or consistency. Some of the most used are mutual information, fisher score, the coefficients such as pearson's correlation kendall correlation, spearman correlation, Chi-square test, relief-f, lasso, minimum redundancy maximum relevance MRMR, laplacian score and spectral feature selection SPEC[7, 8]. Recently filter method such as LW-index which is proposed as filter with wrapper method using Sequence Forward Search algorithm (SFS-LW)[9]. There are many research directions for reducing the GLCM features dimension using filters. Energy, correlation, sum entropy and sum variance are used in mammograms, based on t-test[10] and based on ROC angular second moment entropy and sum entropy are selected with an AUC threshold.

With wrapper method, a variety of features subset selection techniques are used with classifiers as performed by Vanaja and Kumar[11]. k-means is the most unsupervised classification methods used in the wrapper methods[12]. Accuracy is the most common evaluation criterion, beside F-measure, Area Under Curve (AUC), recall, entropy, stability, internal validity measures[13] and average mean squared error of neural networks[14]. Many considerations control the used algorithm like simplicity, stability, number of reduced features, classification accuracy, storage and computational requirements[15].

The problem considered in this research is the optimization of spread spectrum signals clustering, based on features extraction and achieving better clustering performance than using all features. So, the final objective of this research aims for selecting the optimal number of features with highest clustering accuracy. Optimization is done by filters, wrapper methods as SFFS and a bundle of metaheuristic methods which have never been applied to SS signal identification before.

MATERIALS AND METHODS

Signals model: Both signals are simulated in MATLAB R2017a, a random number generator is used to generate binary data stream. The spreading is done by PN-sequence and the modulation by Binary Phase-Shift Keying (BPSK). Because of nowadays complex electromagnetic environment, the signals transmit over the Additive White Gaussian Noise (AWGN) channel, variant Signal-to-Noise Ratio (SNR) are added from -15 to 10 dB (decibel). The original DSSS signal d(t) is generated using 10 bits of -1 and 1, each having a duration t which equals 100 chip patterns, multiplied by a pseudorandom sequence which spreads the bandwidth. The spreading is applied by PN-sequence c(t) which is defined as a sequence of 1 and 0s. Then, the output signal s(t) is modulated by Binary Phase-Shift Keying (BPSK) using sinusoidal carrier wave w_c. The modulation is done by using two sinusoidal carrier waves. The simplest signals modulation is BPSK. The equation is in Eq. 1:

$$s(t) = d(t)c(t)\cos(w_c t) \qquad (1)$$

The original FHSS signals is simulated by a binary data stream and obtained from a random number generator, from ten bits of -1 or 1, then the signal is modulated onto a [illegible] sinusoidal carrier by BPSK to get the modulated signal m(t). The frequency of the carrier is switched between many frequency channels in a pseudorandom manner and at a specified time interval both signals are performed with Additive White Gaussian Noise (AWGN) with SNR added from -15 to 10 dB. In both SS signals, basic signals are generated in a time domain with time-series lengths N = 1000 samples. For comparison perspective, a default random number generator in MATLAB is used.

Signals normalization: The modulated signals are normalized with a scaling factor based on average power using "modnorm" function in MATLAB, finally, the signal constellation is performed by multiplying the modulated signal by the scaling factor. Normalizing the signals gives them the comparative ability and does not make the signal power change with the modulation scheme.

Features extraction: In case of no prior parameter's knowledge, many ways become useless and blind identification or clustering are taken into the consideration. The research proposed many statistical texture features as shown in Fig. 1.

Gray level co-occurrence matrix features extraction Gray Level Co-occurrence Matrix (GLCM): It is a method of extracting texture features by second order statistical. It demonstrates the spatial relationship between pixels. GLCM achieved better results than other texture methods[6]. The 22 features were proposed in previous

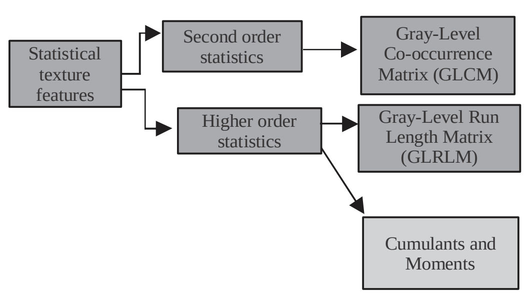


Fig. 1: Proposed textures feature extraction methods

Table 1: List of the 22 GLCM-based features

Features	Parameters
f1	Autocorrelation Amraoui *et al.*[2]
f2	Contrast: MATLAB[14, 2]
f3	Correlation: MATLAB
f4	Correlation[14, 2]
f5	Cluster prominence:[2]
f6	Cluster shade[2]
f7	Dissimilarity[2]
f8	Energy: MATLAB[14, 12]
f9	Entropy[12]
f10	Homogeneity: MATLAB
f11	Homogeneity[2]
f12	Maximum probability[2]
f13	Sum of squares variance[14]
f14	Sum average[14]
f15	Sum variance[14]
f16	Sum entropy[14]
f17	Difference variance[14]
f18	Difference entropy[14]
f19	Information measure of correlation 1[14]
f20	Information measure of correlation 2[14]
f21	Inverse difference Normalized (INN)[2]
f22	Inverse difference moment normalized[14]

Jovic *et al.*[14]; Haralick[16]; Amraoui *et al.*[2]; Soh[17]; Clausi[18]

work, based on GLCM[19]. GLCM describes how frequently a pixel as the reference pixel has intensity value i is repeated in a specific relationship to the neighbor pixel with the intensity value j, separated by a definite pixel distance (Δx, Δy). So, each element (i, j) of the matrix is the number of occurrences of the pair of pixel with value i and a pixel with value j which are at a distance d relative to each other. As input matrix is given, gray level co-occurrence matrix is calculated with definite gray level, direction and distance between the pixel that has focus and its neighbor. Then, the matrix is normalized to obtain the probability matrix. Finally, features derived from the GLCM and calculated by its own equation. To obtain the GLCM features, first convert the input into eight gray levels to capture the signal patterns and create the GLCM, then construct a framework matrix from input window by finding the spatial relation between the reference and neighbor pixel, add the matrix to its transpose to make it symmetrical, normalize the matrix to turn it into probabilities and finally, 22 features are derived and calculated using it. So, the features vector contains 22 features, Table 1 illustrate these features. Although, some features have the same names, the definitions and equations are not identical.

Gray Level Run Length Method (GLRLM): Using higher order statistical 11 features are extracted. First, 7 features are extracted using Gray-Level Run Length Matrix (GLRLM)[20] as shown in Table 2. Then, 4 features extracted which are cumulants (3st, 4st order) and moments (3st, 4st order). GLCM characterizes the textures by continuous pixels having same gray level and contrary GLRLM counts how many times that length occurs for each gray level as shown in Fig. 2.

Table 2: List of Gray-Level Run Length Matrix (GLRLM)

Features	Parameters
f1	Short Run Emphasis (SRE)
f2	Long Run Emphasis (LRE)
f3	Gray Level Non-uniformity (GLN)
f4	Run Percentage (RP)
f5	Run Length Non-uniformity (RLN)
f6	Low Gray Level Run Emphasis (LGRE)
f7	High Gray level Run Emphasis (HGRE)

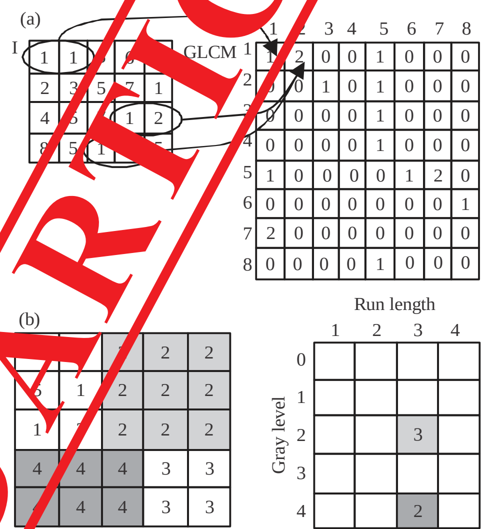


Fig. 2(a, b): Main concept of GLCM and GLRLM

Cumulants (3st, 4st order), Moments (3st, 4st order): Higher order moments and cumulants provide a statistical way to describe the shape of distribution function of a signal[21].

Principal Components Analysis (PCA), Kernel-based, Principal Components Analysis (KPCA) and Fast and Independent Components Analysis (Fast-ICA): They are the most common feature extraction methods in blind source separation. PCA attempts to find uncorrelated sources, conversely ICA attempts to find independent sources. ICA is an algorithm that finds directions corresponding to projections with high non-Gaussian, on the other hand PCA finds directions accounting for maximum variance. FastICA is a fixed-point algorithm.

Feature scaling is an affective step prior clustering to avoid incorrect impact and to simplify the values in the distance-based methods. Z-score as standardization method is important because it gives the same importance to all features and leads to better quality. (Suarez-Alvarez *et al.*,[22]. It scales the features to have a standard normal distribution with $\mu = 0$ and $\sigma = 1$, representing the mean and standard deviation from the mean, respectively. All features are scaled before clustering. The z-score is calculated with the following Eq. 2:

$$Z = (X - \mu)/\sigma \tag{2}$$

Table 3: Clustering techniques and its parameters

Clustering model	Names	Parameters
Centroid based methods	k-means, k-medoids, C-means, (FCM), Possibilistic C-Means (PCM), Fuzzy Possibilistic C-Means (FPCM)	Initial center is points of data chosen randomly
	C-means	Exponent for the fuzzy partition matrix = 10
	Subtractive	Range of influence is equal 0.9
Hierarchical	Agglomerative Clustering (HAC)	Euclidean distances and ward linkage
Density-based spatial	DBSCAN	Eps: analytical function k: 5
Model-based	A mixture of Gaussians clustering (GMM)	Diagonal covariance matrix Components number k = 2

Table 4: Performance of the candidate feature extraction techniques

Variables	Avg extraction time	First level of accurate clustering (SNR)	Top average accuracy	Best clustering time
GLCM	0.107	0	**84.4**	**0.004**
GLRLM	0.178	-1	80.27	0.008
HOS (Moment, Cumulant)	0.138	2	78.40	0.005
PCA	0.137	2	67.64	**0.002**
KPCA	0.185	-1	67.73	0.005
FAST_ICA	0.104	10	53.88	0.003

Bold values are significant

Clustering techniques: Unsupervised method has many advantages proposed it in SS identification. As, no need for complex training step and the process directed by the structure of the data and investigates its characteristics. Therefore, no need to ground-truth labels to separate the data into clusters. This research applied ten of clustering techniques based on GLCM features which they are belonging to multi clustering concept as centroid based methods, hierarchical clustering, density-based spatial clustering and model-based clustering.

Features selection techniques: The selection of the most effective features presents big challenge for the recognition of these signals[23]. The main objective of this study is to optimize the SS identification based on clustering techniques. Features selection is categorized in three main techniques, filter, wrapper and embedded, it is possible to use hybrid approaches[14]. Each features selection techniques category has pros and cons. Filters usually used to analyze intrinsic properties of data without any classifier assistance, the key of filters is rank and sort then defined number of features are tested using the classifiers. On the contrary wrappers often build on classifier performance or predictive power. In embedded method learning algorithm plays a basic role in feature selection process, it injects the feature selection with the learning algorithm. The hybrid model used the benefits of filters as criteria and wrapper model[24]. The proposed model adopts filter and wrapper-based method to find the optimal subset of features.

Clustering: Table 3 presents the used ten clustering techniques and its parameters. For all clustering techniques iterations are equals 100. For assessment reason, the overall accuracy rate is used. It measures the quality of cluster labels produced by the algorithm compared with the class labels. It is an external measure provides the success rate Cs which obtains by dividing the number of samples correctly classified S by the total number of samples N. The complement of this is called the misclassification rate. It can calculate using in Eq. 3:

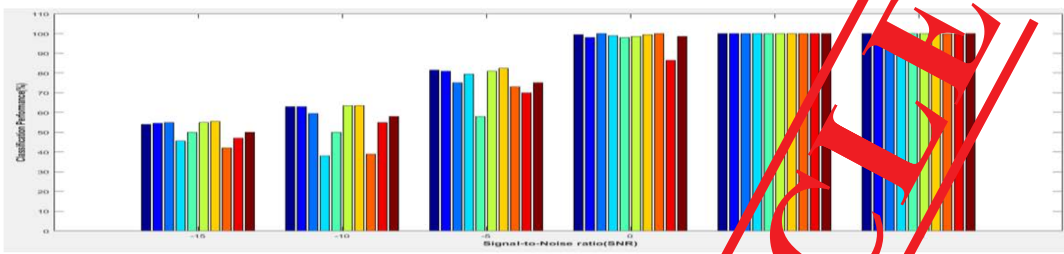
$$C_S = S/N \tag{3}$$

Table 4 shows the performance of each method. Last study identifies the signals with 100% accuracy over 4 SNR using k-means based on two features which are carrier frequency and estimated bandwidth[25].

RESULTS AND DISCUSSION

To find out the appropriate features for accurate clustering, many feature extraction techniques were used, although, GLRLM got better level of signals identification in -1 SNR, it neglected as it consumes much time and has unbalanced results. KPCA, KPCA and FAST_ICA all have low average accuracy. GLCM gets the appropriate result compares with the other techniques. Figure 3 gives the accuracy of the 10 clustering methods using all the 22 features based on GLCM, in points -15, 10, 5, 0, 5 and 10. From Table 5, it is noticeable that affinity propagation takes a long time compared with others therefore it will be neglected.

Features selection: Selection of the most effective features are considered as the objective of this researcher. The feature selection problem generally can be defined as multi-objective problem that maximize the classification performance and minimize the number of features[26].



	-15	-10	-5	0	5	10
Kmeans	54.3	63	**81.5**	99.5	100	100
kmedoids	54.5	63	81	98	[illegible]	100
Hierarchical	55	59.5	75	100	100	100
C-Means	45.5	39	79.5	100	[illegible]	100
Possibilistic C-Means	54	62.5	78.5	98.5	100	100
Fuzzy-Possibilistic C-Means	50	50	58	99	100	100
Affinity	55	63.5	81	98.5	100	100
Gaussian Mixture	47	55	70	[illegible]	100	100
BDSCAN	50	58	75	98.5	100	100

Fig. 3: Clustering accuracy using 22 GLCM-based features at different SNR levels

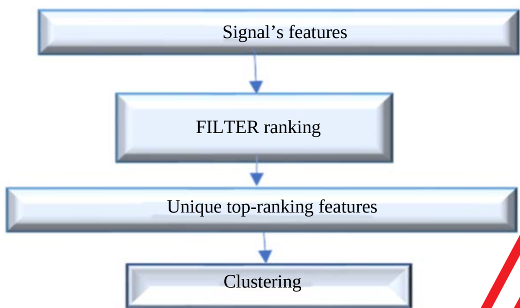


Fig 4: Subset clustering steps using filters

Table 5: Average time of clustering with all features

Extraction method/Techniques	Time (sec)
GLCM features	
k-means	0.005
Affinity	2.130
Hierarchical	0.016
GMM	0.064
FPCM	0.012
C-means	0.004
Subtractive	0.011
FCM	0.006
k-medoids	0.011
PCM	0.017
DBSCAN	0.048

With absence of the required number of important features, the experiment goes into many phases summerized in Table 6:

- Isolated features test: test each feature clustering accuracy
- Filter: construct a collective features set using the unique first top score feature over the 26 noise level
- Two features performance: test Subset Features with Forward Sequential selection (SFFS) and Binary Particle Swarm Optimization (BPSO)
- Undefined number of features: test stochastic number of features with selective heuristic algorithms

Test each features performance: Each feature is tested by overall accuracy of the 10 clustering techniques. Table 6 shows average accuracy of each features compared with the results of clustering using all features. The best features considered as the one that establishing higher average clustering accuracy with early signals detected within the 26 levels of SNR. Selecting isolated feature cannot give the absolute accurate best result as there is a probability to find relation between the features subset.

Filter: When using filter method in features selection, the desired number of features can be defined by the researcher. The signal corruption with additive white Gaussian noise at different SNR, therefore, the subset constructs from the unique top-ranking feature during all 26 SNR levels. Finally, the subset performance is tested with different clustering techniques as shown in Fig. 4.

Thirteen filters methods are used which are Relief F, laplacian score, mutual information (mutinffs), local learning-based clustering (llcfs), Correlation based feature delection (Cfs), Unsupervised discriminative feature selection (Udfs), adaptive structure learning (fsvFS), concave minimization, Infinite Latent Feature Selection (ILFS), Adaptive Structure Learning (FSASL) (lasso[27, 28], Dependence Guided Unsupervised Feature Selection (DGUFS) and Unsupervised Feature Selection with Ordinal Locality (UFSwithOL)[28, 13]. Although, ILFS achieves the highest accuracy, it ranked 10 features as the top one as shown in Fig. 5. As the fitness measured with multi objectives optimization problem, the best result is obtained by best classifier accuracy and the minimum number of features, the best performance was by LLCFS which ranked only one feature during all noise levels with accuracy equals 82.77 as shown in Fig. 6. Table 7 shows the performance of each filter method. Filter is a fast method but do not consider the relationships between variables.

Table 6: Result of each feature accuracy with the ten clustering techniques

Features	k-means	k-medoids	HAC	C-means	PCM	FPCM	Affinity	GMM	DB	Subtractive	Top accuracy	Level of accuracy 100%
Avg. accuracy of all features	**84**	**84.4**	**82.02**	**73.21**	**74.71**	**83.77**	**81.19**	**77.96**	**65.38**	**69.23**	**84.4**	0
f1	62.58	68.92	57.33	62.60	64.92	62.19	61.98	66.6	61.13	[illegible]	**68.92**	5
f2	82.38	82.46	72.06	59.73	74.35	75.58	78.06	54.54	67.27	82.38	**82.46**	1
f3	82.94	82.98	71.60	51.73	74.13	74.77	80.37	55.75	[illegible]	82.94	**82.98**	2
f4	82.94	82.98	71.60	51.73	74.13	74.77	80.37	55.75	65.25	82.94	**82.98**	2
f5	82.10	82.23	60.98	56.13	73.23	76.54	79.96	69.29	[illegible]	82.10	**82.23**	2
f6	82.75	82.77	74.79	63.75	76.92	78.27	82.23	61.35	67.27	[illegible]	**82.77**	**0**
f7	**83.44**	**83.48**	**77.48**	**55.00**	**75.35**	**77.35**	**79.58**	**50.42**	[illegible]	**83.44**	**83.48**	**1**
f8	72.75	71.50	67.27	52.40	66.75	66.48	64.92	67.94	57.40	72.75	**72.75**	6
f9	68.42	66.65	56.08	57.83	58.96	64.92	56.94	[illegible]	50.00	[illegible]	**68.42**	0
f10	82.38	82.50	63.25	44.44	73.48	73.96	78.85	[illegible]	64.63	[illegible]	**82.50**	2
f11	82.37	82.42	69.13	44.54	73.21	77.04	77.10	[illegible]	[illegible]	82.37	**82.42**	2
f12	72.29	72.63	66.44	50.29	67.12	70.92	64.23	66.58	[illegible]	72.29	**72.63**	7
f13	77.33	77.37	61.29	73.46	73.12	75.00	65.83	64.67	67.21	77.33	**77.37**	2
f14	77.31	77.37	63.44	71.44	72.83	72.81	65.56	[illegible]	67.25	77.31	**77.37**	1
f15	67.62	67.62	53.27	67.71	67.42	64.48	66.08	66.75	[illegible]	67.62	**67.71**	5
f16	69.56	69.21	60.54	57.75	63.96	65.08	63.62	[illegible]	57.54	69.56	**69.56**	7
f17	82.38	82.46	71.67	59.73	74.35	75.52	78.06	54.54	[illegible]	82.38	**82.46**	1
f18	76.73	75.62	64.46	51.52	70.31	71.94	69.17	60.46	61.50	76.73	**76.73**	4
f19	77.81	76.71	67.50	57.56	72.92	74.15	[illegible]	[illegible]	63.38	77.81	**77.81**	3
f20	77.88	77.60	63.75	57.46	73.29	72.00	76.60	65.37	63.38	77.88	**77.88**	3
f21	**83.67**	**83.58**	**73.54**	**47.19**	**76.54**	**80.00**	**80.06**	**50.94**	**67.15**	**83.67**	**83.67**	**1**
f22	82.63	82.63	69.79	55.79	75.25	[illegible]	78.38	[illegible]	67.27	82.63	**82.63**	1

Bold values are significant

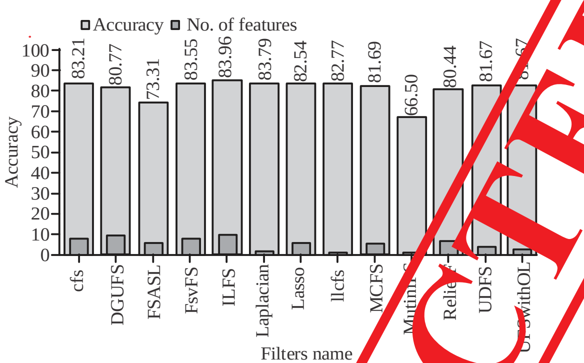


Fig. 5: Accuracy of each filter with the number of selected features

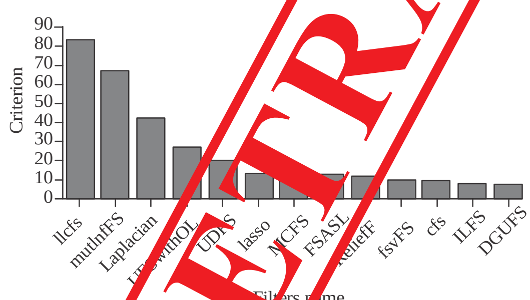


Fig. 6: Fitness of filter methods

Wrapper method: This reasonably method as classifier accuracy represents the criterion measure and the final goal. It selects the features that work best with a given classifier and runs clustering algorithm on one data set and compute the criterion value on the same data set as shown in Fig. 7.

Wrapper methods are used in searching for the best two features, in addition to searching about undefined number of features. Three factors should be specified: classifier, feature subset evaluation criteria and a searching technique. The used techniques are summarized below:

Classifier: The research uses clustering techniques direction as listed before.

Feature subset evaluation criteria: As an multi objectives optimization problem, the best result is obtained by best classifier accuracy and the minumum number of feature.

Searching technique: Have many directions varying between sequential and randomized methods. The research adopts two direction, sequential and binary metaheuristic search algorithm.

Find the best two features accuracy

Sequential Forward Floating Selection (SFFS): Trying all possible subsets with greedy stepwise approaches are considered as a computationally expensive task. Therefore, SFFS is proposed. The SFFS starts with an empty subset, then it adds feature that minimizes the criterion value. Stop when find the best defined number of features which achieve the criterion. Misclassification rate is used as a SFFS criterion without separating test and

Table 7: Performance of filters methods

Filter acronym	Unique top features	k-means	k-medoids	HAC	C-means	PCM	FPCM	Affinity	GMM	DB	Subtractive
ILFS	12	**83.960**	83.88	74.500	80.210	83.62	73.27	[illegible]	67.31	[illegible]	83.88
RELIEFF	7	80.080	79.96	78.350	77.100	79.06	71.21	[illegible]	66.96	[illegible]	79.96
MUTINFFS	1	62.620	62.56	57.330	62.600	64.92	62.21	**66.50**	[illegible]	58.269	61.13
FSVFS	8	83.340	**83.56**	70.236	80.256	55.43	76.06	[illegible]	[illegible]	53.670	55.77
LAPLACIAN	2	83.730	**83.79**	62.260	79.120	55.33	77.77	83.52	78.60	82.690	70.63
UDFS	4	81.150	**81.67**	54.400	80.500	60.60	[illegible]	79.87	[illegible]	58.980	71.56
CFS	9	82.750	**83.21**	67.520	80.400	55.09	[illegible]	[illegible]	77.71	52.340	62.77
LLCFS	1	82.750	**82.77**	74.790	82.230	64.15	[illegible]	[illegible]	[illegible]	70.480	65.36
FSASL	6	72.350	66.48	60.600	64.810	68.94	[illegible]	72.27	69.31	72.300	66.48
LASSO	**6**	82.310	**82.17**	70.880	77.690	48.37	77.04	[illegible]	78.40	80.350	53.85
UFSWITHOL	3	81.150	**81.67**	54.400	80.500	60.60	73.21	79.87	77.92	58.980	71.56
DGUFS	10	80.480	79.69	53.600	79.170	56.98	[illegible]	80.77	79.85	63.150	61.04
MCFS	6	80.420	**81.69**	56.370	80.210	58.25	[illegible]	[illegible]	79.52	70.830	57.69

Bold values are significant

Table 8: Parameters of BPSO

Parameter	Values
An inertial weight	0.2
Gloubal learning cofficient	2
Personal learning cofficient	**2**
Population	**30**

Bold values are significant

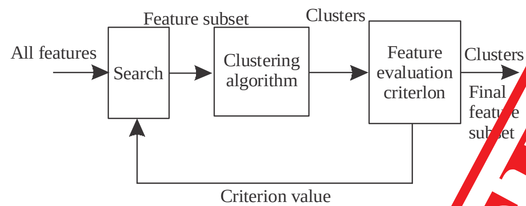


Fig. 7: Wrapper method for clustering

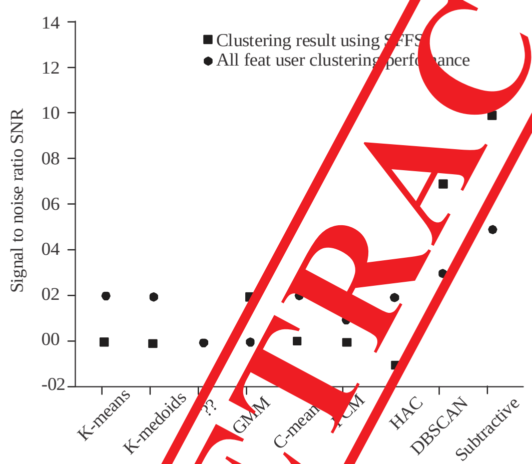


Fig. 8: All features and SFFS first accurate SNR level

Searching technique: Have many directions varying between sequential and randomized methods. The research adopts two direction, sequential and binary metaheuristic search algorithm.

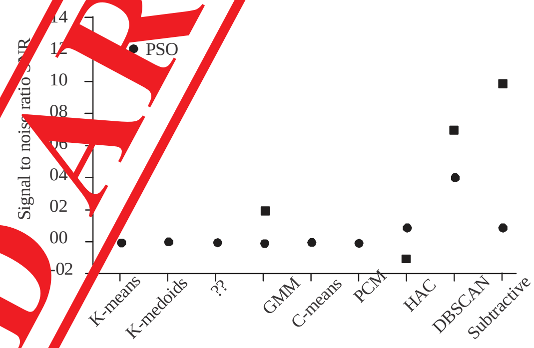


Fig. 9: PSO and SFFS first accurate SNR level

training sets. SFFS improves the first level of identification using FPC and gets 100% accuracy in -1 SNR. SFFS improves k-means, k-medoids, C-means in 0 SNR as shown in Fig. 8.

Binary Particle Swarm Opimization (BPSO): Argued to be computationally less expensive than other evolutionary computation methods[29]. In this research the accuracy of the 10 clustering methods used as a fitness measure. Parameters are listed in Table 8.

Comparison Between BPSO and SFFS: Figure 9 presents first level of accurate identification of PSO and GA. PSO can only enhance DBSCAN to 4 SNR and subtractive clustering to 0 SNR. The two has same performance in k-means, k- medoids, C-means and PCM in 0 SNR.

As the two mathods construct the subset with two features, the comparision can be made by their accuracy, Fig. 10 shows thier avarage accuracy. But the performance of two features does not presents an equalsifent improvement in accuracy. Table 9 shows the SFFS and PSO accuracy compared with all and top one features accuracy. Each of the two methods have disadvantages, SFFS suffers from stacking in local

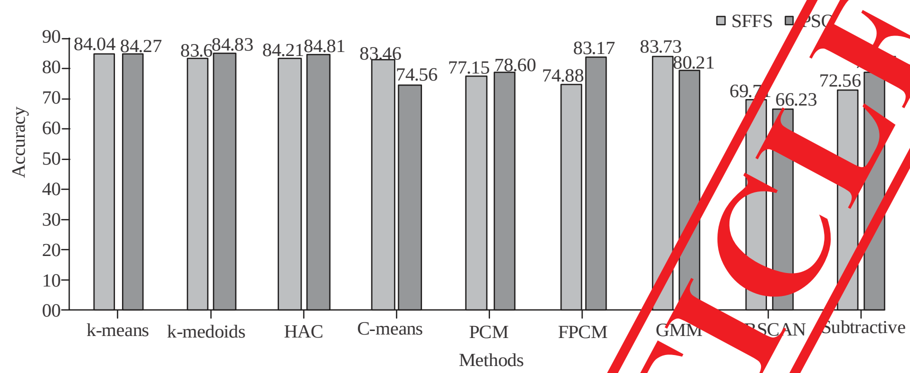


Fig. 10: Comparison between SFFS and PSO method in terms of clustering accuracy

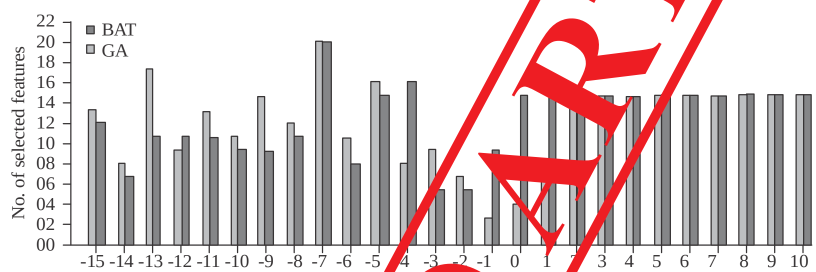


Fig. 11: Comparison between GA and BAT in term of selected features using hieratical accuracy

Table 9: Comparative between one and two feature average accuracy

Variables	k-means	k-medoids	HAC	C-means	PCM	FPCM	GM M	DBSANE	Subtractives
All	**83.81**	83.40	82.02	[illegible]	74.71	**83.77**	77.96	65.38	69.23
Top accuracy with one feature	83.67	83.58	77.48	73.46	76.92	80.00	82.23	49.58	67.27
Two features (SFFS)	84.04	83.60	84.21	**83.46**	72.15	74.88	**83.73**	**69.71**	72.56
Two features (PSO)	**84.27**	**84.83**	**84.81**	74.56	**78.60**	83.17	80.21	66.23	**78.77**

Bold values are significant

Table 10: Parameters of GA

Parameters	Values
Size of population	50
Replacement rate	0.8
Crossover rate	[illegible]
Mutation rate	0.01
Selection method	Uniform selection
Elitism	**2**
No. of generation	**100**
Selection metjod	Tournament size 2

Bold values are significant

optima, high computational time and unable to remove features that become neglected after the addition of other features and BPSO easily suffers from less exact at the regulation of its speed and direction.

Then GA, Binary Bat (BBA) and Hybrid Whale Optimization Algorithm with Simulated Annealing (WOSAT) are used in optimization of the clustering performance with undefined number of features. With the previous fittness measures which aimes to find the minumum number of selected features with minumum misclassification error.

Random methods evolutionary feature subset selection binary GA: One of the most advanced algorithms for feature selection is the genetic algorithm. This is a stochastic method for function optimization based on the mechanics of natural genetics and biological evolution. GA has some advantages such as the ability to manage data sets with many features and to solve the complex and non-linear problems and don't need specific knowledge about the problem under study. Parameters of the experiment is showed in Table 10.

Binary bat feature selection: The idea of the Bat Algorithm (BA) is to mimic the behaviors of bats when catching their prey. BA was found to outperform PSO, GA and others heuristic algorithms[31, 30], using k-Nearest Neighbor (KNN), Naive Bayes (NB), Decision Tree (DT) and the Optimum-Path Forest (OPF) classifiers. BA is based on echolocation behavior of micro bats with varying pulse rate of emission and loudness. The bats communicate with each other through the global best solution and move towards the global best solution. The initially number of bats is preferred

Table 11: Parameters of BBA

Parameters	Values
No. of bat	20
Terminated	100 iterations
decrease sound loudness and increase pulse rate	1

Table 12: Parameters of WOASAT

Parameters	Values
No. of search agents	10
b in WOA	1
Max iteration	100
Tournament selection probability	0.5

Table 13: Average accuracy and selected features values obtained from the different optimizers

	WOASAT		GA		BBA		Accuracy (2 features)	
Methods	Avg. features num.	Avg. accuracy	Avg. features num.	Avg. accuracy	Avg. features num.	Avg. accuracy	SFFS	BPSO
k-means	4.17	**83.83**	9.230	**83.65**	[illegible].080	79.98	**84.04**	**84.27**
k-medoids	3.83	77.25	9.080	**82.50**	8.8[illegible]	79.00	83.60	**84.83**
HAC	5.67	76.08	11.27	68.29	11.00	70.69	**84.21**	**84.81**
C-means	5.67	75.92	9.230	**83.90**	10.27	67.38	83.46	74.56
FPCM	4.83	74.75	8.000	83.35	[illegible]	73.67	74.88	83.17
GMM	4.33	76.08	9.920	78.90	9.420	76.73	83.73	80.21
DBSCAN	6.33	75.67	9.650	67.2[illegible]	11.12	55.48	69.71	66.23
Subtractive	2.50	77.00	11.50	59.[illegible]2	13.15	50.00	72.56	78.77

Bold values are significant

to be as half of the number of features in the dataset. Random number were used to initial the value of pulse and loudness. Parameters of BBA is showed in Table 11.

Comparison between GA and BBA: From the selected features perspective and as example Fig. 11 shows the variation in the selected features numbers using hierarchical clustering accuracy.

GA and BBA drawback: GA seems to be very expensive in computational terms, since, evaluation of each individual requires building a predictive model and can take a long time to converge, since, they have a stochastic nature. GA consumed a double time of BBA. The implementation of BBA is more complicated than many other metaheuristic algorithms because each agent (bat) is assigned a set of interacting parameters such as position, velocity, pulse rate, loudness and frequencies.

Hybrid whale optimization algorithm with simulated Annealing and tournament: WOA is an optimization algorithm that mimics the behavior of the humpback whales bats is preferred[32]. Humpback whales know the location of prey and encircle them. They consider the current best candidate solution is best obtained solution and near the optimal solution. After assigning the best candidate solution the other agents try to update their positions towards the best search agent. WOA parameters listed in Table 12. WOA has good properties such as:

- High flexibility and implementation simplicity
- Less dependency on parameters: fewer number of parameters to control, since, it includes only two main internal parameters to be adjusted
- WOA algorithm smoothly transit between exploration (search for pray) and exploitation (encircling prey/bubble-net attacking method) depending on only one parameter
- Provide good balance between exploration (local optima avoidance) and exploitation (convergence) enhance the performance of the searching algorithm

SA enhances the exploitation in WOA algorithm by searching the most promising regions located by WOA algorithm (Table 13). Then, tournament selection employment enhanced the exploration in WOA algorithm which complemented the role of SA (Mafarja and Mirjalili, 2017). SA is a single-solution metaheuristic algorithm based on the hill climbing. It beat the problem of stuck in local optima, SA utilizes a certain probability to accept a worse solution. Table 13 shows the performance superiority of WOASAT compared with GA and BBA in terms of the selected features.

CONCLUSION

GLCM features can be used in SS signal identification using clustering techniques as they got the best result. They are the most appropriate technique compare to another feature's methods. Dissimilarity and Inverse difference Normalized (INN) extracted by GLCM are the two best features compared to other features with an average overall accuracy of 83.67, 83.48. They can be used in identification of the two signals even with a low SNR as they can accurately clustering the signals in 1 SNR. The best results are obtained with k-means, k-medoids, C-means and FPCM. Using subset of GLCM features, filter method cannot reach the appropriate features for optimization. PSO can propose to the optimization using HAC and k-means as it outperforms SFFS, WOASAT, GA and BBA in terms accuracy and number of features.

REFERENCES

01. Hasan, M., J.M. Thakur and P. Podder, 2016. Design and implementation of FHSS and DSSS for secure data transmission. Int. J. Signal Process. Syst., 4: 144-149.
02. Amraoui, A., B. Benmammar, F. Krief and F.T. Bendimerad, 2012. Intelligent wireless communication system using cognitive radio. Intl. J. Distrib. Parallel Syst., 3: 91-104.
03. Torrieri, D., 2005. Principles of Spread-Spectrum Communication Systems. Springer, Boston, USA.
04. Cheng, Z. and Z. Lu, 2018. A novel efficient feature dimensionality reduction method and its application in engineering. Complexity, Vol. 2018, 10.1155/2018/2879640
05. Li, J., K. Cheng, S. Wang, F. Morstatter, R.P. Trevino, J. Tang and H. Liu, 2017. Feature selection: A data perspective. ACM Comput. Surv. (CSUR.), Vol. 50,
06. Zhang, X., J. Cui, W. Wang and C. Lin, 2017. A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm. Sensors, Vol. 17, 10.3390/s17071474
07. Sen, B., M. Peker, A. Cavusoglu and F.V. Celebi, 2014. A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms. J. Med. Syst., Vol. 38, 10.1007/s10916-014-0018-0
08. Krishnaveni N. and V. Radha, 2019. Feature selection algorithms for data mining classification: A survey. Indian J. Sci. Technol., Vol. 12, No. 6. 10.17485/ijst/2019/v12i6/139581
09. Liu, C., W. Wang, Q. Zhao, X. Shen and M. Konan, 2017. A new feature selection method based on a validity index of feature subset. Pattern Recognit. Lett., 92: 1-8.
10. Pratiwi, M., Alexander, J. Harefa and S. Nanda, 2015. Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network. Procedia Comput. Sci., 59: 83-91.
11. Vanaja, S. and K.R. Kumar, 2014. Analysis of feature selection algorithms on classification: A survey. Int. J. Comput. Applic., 96: 29-35.
12. Rajalakshmi and S. Jayanthi, 2017. Feature selection using wrapper methods for writer identification of Tamil handwritten documents. Int. J. Comput. Technol. Appl., 8: 666-671.
13. Roffo, G., 2017. Ranking to learn and learning to rank: On the role of ranking in pattern recognition applications. Ph.D. Thesis, Department of Computer Science, University of Verona (Università degli Studi di Verona), Verona, Italy.
14. Jovic, A., K. Brkic and N. Bogunovic, 2015. A review of feature selection methods with applications. Proceedings of the 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO'15), May 25-29, 2015, IEEE, Opatija, Croatia, pp: 1200-1205.
15. Chandrashekar, G. and F. Sahin, 2014. A survey on feature selection methods. Comput. Electr. Eng., 40: 16-28.
16. Haralick, R.M., K. Shanmugam and I.H. Dinstein, 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern., SMC-3: 610-621.
17. Soh, L.K. and C. Tsatsoulis, 1999. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans. Geosci. Remote Sens., 37: 780-795.
18. Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens., 28: 45-62.
19. Fouad, H.S., H.A. Elsayed and S.K. Guirguis, 2019. Direct sequence and frequency hopping signals classification based on co-occurrence matrix and clustering techniques. J. Comp. Sci., 15: 78-91.
20. Wei, X., 2007. Gray level run length matrix toolbox v1.0, software. Beijing Institute of Technology Aeronautics Astronautics Engineering Experiment Teaching Center, Beijing, China.
21. Sanaullah, M., 2013. A review of higher order statistics and spectra in communication systems. Global J. Sci. Front. Res. Phys. Space Sci., 13: 31-50.
22. Suarez-Alvarez, M.M., D.T. Pham, M.Y. Prostov and Y.I. Prostov, 2012. Statistical approach to normalization of feature vectors and clustering of mixed datasets. Proc. R. Soc. A: Math. Phys. Eng. Sci., 468: 2630-2651.
23. Hamed, H.A., A.K. Abdullah and S. Al-Waisawy, 2018. Frequency hopping spread spectrum recognition based on discrete fourier transform and skewness and kurtosis. Int. J. Appl. Eng. Res., 13: 7081-7085.
24. Jaskowiak, P.A. and R.J. Campello, 2015. A cluster based hybrid feature selection approach. Proceedings of the 2015 Brazilian Conference on Intelligent Systems (BRACIS'15), November 4-7, 2015, IEEE, Natal, Brazil, pp: 43-48.

25. Jayaweera, S., 2015. Signal Classification in Wideband Cognitive Radios. In: Signal processing for Cognitive Radios, Jayaweera, S. (Ed.). Wiley, New Jersey, USA., ISBN: 9781118824931, pp: 429-471.
26. Kumbhar, P. and M. Mali, 2016. A survey on feature selection techniques and classification algorithms for efficient text classification. Int. J. Sci. Res., 5: 1267-1275.
27. Du, L. and Y.D. Shen, 2015. Unsupervised feature selection with adaptive structure learning. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'15), August 10-13, 2015, Sydney, Australian, pp: 209-218.
28. Guo, J., Y. Quo, X. Kong and R. He, 2017. Unsupervised feature selection with ordinal locality. Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME'17), July 10-14, 2017, IEEE, Hong Kong, China, pp: 1213-1218.
29. Normandin, M.E., S.D. Mohanty and T.S. Weerathunga, 2018. Particle swarm optimization based search for gravitational waves from compact binary coalescences: Performance improvements. Phys. Rev. D., Vol. 98. 10.1103/PhysRevD.98.044029
30. Nakamura, R.Y.M., L.A.M. Pereira, K.A. Costa, D. Rodrigues and J.P. Papa, 2012. BBA: A binary bat algorithm for feature selection. Proceedings of the 25th SIBGRAPI Conference on Graphics, Patterns and Images, August 22-25, 2012, Ouro Preto, pp: 291-297.
31. Mohamed, R., M.M. Yusof and N. Wahidi, 2018. A comparative study of feature selection techniques for bat algorithm in various applications. MATEC Web Conf., Vol. 150.
32. Mafarja, M.M. and S. Mirjalili, 2017. Hybrid whale optimization algorithm with simulated annealing for feature selection. Neurocomputing, 260: 302-312.