

# Multiple-Parameter Radar Signal Sorting Using Support Vector Clustering and Similitude Entropy Index

Zhanling Wang · Dengfu Zhang · Duyan Bi · Shiqiang Wang

Received: 12 March 2013 / Revised: 26 November 2013 / Published online: 18 December 2013  
© Springer Science+Business Media New York 2013

**Abstract** The radar signal sorting method based on traditional support vector clustering (SVC) algorithm takes a high time complexity, and the traditional validity index cannot efficiently indicate the best sorting result. Aiming at solving the problem, we study a new sorting method based on cone cluster labeling (CCL) method. The CCL method relies on the theory of approximate coverings both in feature space and data space. Also a new cluster validity index, similitude entropy (SE), is proposed. It can be used to evaluate the compactness and separation of clusters with information entropy theory. Simulations including the performance comparison between the proposed method and the conventional methods are presented. Results show that while maintaining the sorting accuracy, the proposed method can reduce the computing complexity effectively in sorting the signals.

**Keywords** Signal sorting · Clustering algorithms · Support vector clustering · Cone cluster labeling · Similitude entropy

## 1 Introduction

The sorting of radar signal is to separate the interleaved radar pulse signal flow from different sources. As an important part of electronic intelligence (ELINT) and

---

Z. Wang · D. Zhang · D. Bi · S. Wang (✉)  
School of Aeronautics and Astronautics, Air Force Engineering University, No. 1 Baling Road,  
Baqiao District, Xi'an City 710038, P.R. China  
e-mail: [wunsicon@163.com](mailto:wunsicon@163.com)

Z. Wang  
e-mail: [wangzhanling1988@163.com](mailto:wangzhanling1988@163.com)

D. Zhang  
e-mail: [zhang68\\_94@163.com](mailto:zhang68_94@163.com)

D. Bi  
e-mail: [biduyan@126.com](mailto:biduyan@126.com)

electronic support measure (ESM) system, signal sorting affects the performance of electronic reconnaissance equipments directly, and is a key technology to campaign decision-making [22].

In general, parameters used in sorting algorithms are extracted from the pulses and mainly composed of direction of arrival (DOA), radio frequency (RF), pulse amplitude (PA), pulse width (PW) and time of arrival (TOA), which together form pulse descriptor word (PDW) [20]. Radar signal sorting approaches are mostly based on single parameter, parameter for parameter or multiple-parameter sorting, namely including three aspects as follows:

- (1) The sorting method relying on single parameter mainly focuses on the parameter of TOA by which the pulse repetition interval (PRI) can be obtained. Then some different algorithms could be used to perform the signal sorting with PRI [6, 28]. However, with more and more complex waveform design and electromagnetic environment, sorting with single parameter has some limitations and the sorting rate is not high either.
- (2) The parameter for the parameter sorting method is a serial regulation detection system. Pulse parameters such as RF, PA, PW and DOA, are compared with pulse groups that are confirmed in advance, and it should be detected that whether each parameter drops into the existent cell with certain tolerance range or not [25]. This method not only has a slow sorting speed but also is inefficient for sorting the parameters which are incomplete or under the influence of jamming or noise.
- (3) The multiple-parameter sorting method is usually performed with clustering method using several parameters in PDW [7, 8, 12]. This approach can acquire a better sorting result compared to those two sorting methods mentioned above.

As with the above aspects, the radar signal sorting method should mainly focus on the clustering sorting method. As the most efficient clustering method, the support vector clustering (SVC) algorithm [1] has been widely researched in both theoretical developments and practical applications due to its outstanding features such as in [9, 13, 15, 16, 18, 19, 21, 24]. Zhang et al. [27] deal with an application of SVC to radar emitter signal (RES) recognition. And in [7] the algorithm is used to perform the radar signal sorting and gets a good sorting result. However, as a common agreement the SVC algorithm is time-consuming (as shown in experiment result) because of its optimization problem and cluster labeling method [17].

This paper presents a new sorting algorithm: firstly using SVC algorithm to map the radar signal data points to a high dimensional feature space and form some separate clusters of points; then differentiating between points that belong to different clusters with cone cluster labeling (CCL) method, we regard these two steps as cone mapping support vector clustering (CMSVC); finally indicating the best sorting result efficiently with the proposed validity index of similitude entropy (SE), which assesses the compactness and separation of clusters with information entropy theory.

The paper is organized as follows. In Sect. 2 the signal sorting model based on SVC with the cluster labeling method of CCL is introduced. The proposed similitude entropy (SE) index for indicating the correct sorting clusters is derived, and then the radar signal sorting system based on CMSVC and SE index is presented in Sect. 3. Section 4 implements the sorting algorithm by adjusting the parameters of CMSVC

and updating the dynamic library of radar signal clusters. In Sect. 5, several experimental results are presented and discussed. The paper is concluded in Sect. 6.

## 2 Cone Mapping Support Vector Clustering-Based Signal Sorting

In this section, we introduce the signal sorting model based on SVC with the cluster labeling method of CCL, namely based on the cone mapping support vector clustering (CMSVC) algorithm.

### 2.1 Sorting Model Based on Support Vector Clustering

While using the sorting methods based on clustering, the greatest obstacle of clustering sorting is that features used to cluster cannot obtain satisfactory separation of clusters. This drawback hinders the practical effects of clustering sorting. But support vector clustering (SVC) hold the edge to resolve the above issue.

The SVC is an unsupervised and non-parametric clustering algorithm, presented by Ben-Hur et al. [1] based on the support vector domain description (SVDD) algorithm. This method takes the support vector machine (SVM) as a tool to perform clustering and its basic idea is as follows: first, mapping the data points by means of a *Gaussian* kernel to a high dimensional feature space from data space; then looking for the optimum hypersphere that can enclose the image of the data in this new space; and finally mapping this hypersphere back to data space and forming a set of contours which enclose the data points. Theses contours are interpreted as cluster boundaries. Points enclosed by each separate contour are associated with the same cluster.

The SVC algorithm is the most effective method for unsupervised classification and has several unique advantages—generating cluster boundaries of arbitrary shape, enabling analyzing noisy data points and separating between overlapping clusters. This is what other clustering algorithms cannot accomplish. And the linear separable probability of data points is enhanced with the algorithm by a nonlinear transformation. As a result, the useful feature is probably recognized, extracted and magnified relatively [2, 26]. Meanwhile, by employing slack variables the SVC is able to deal with outliers and get rid of the influence of noise.

Let  $G \subseteq \mathfrak{R}^3$  be a subset of a PDW, here  $G = \{RF, DOA, PW\}$ , and  $\{\mathbf{g}_i\} \subseteq G$  be the radar signal pulse parameters vector, here  $i = 1, 2, \dots, N$ ,  $N$  denoting the number of vector. By mapping the parameters, namely data points to feature space we get a smallest closed sphere with the center  $\mathbf{a}$  and the radius  $R$ . Thus the signal sorting problem can be converted to solve the optimization problem described by the constraints [1]

$$\min R^2 + C \sum_i \xi_i \quad \text{s.t.} \quad \|\Phi(\mathbf{g}_i) - \mathbf{a}\|^2 \leq R^2 + \xi_i \quad \xi_i \geq 0 \quad (1)$$

where  $C$  is the constant penalty factor,  $\Phi$  is a nonlinear transformation,  $\Phi(\mathbf{g})$  denotes the image of  $\mathbf{g}$ ,  $\|\bullet\|$  is the *Euclidean* norm and  $\xi_i$  the slack variable.

To solve the problem, the *Lagrangian* is introduced as

$$L = R^2 - \sum_i (R^2 + \xi_i - \|\Phi(\mathbf{g}_i) - \mathbf{a}\|^2) \beta_i - \sum_i \xi_i \mu_i + C \sum_i \xi_i \quad (2)$$

where the  $\beta_i$  and  $\mu_i$  are Lagrange multipliers.

Setting to zero the derivate of formula (2), we can deduce the equivalent problem of formula (1) from the *Wolfe* dual form as

$$\begin{aligned} & \max_{\beta} \left( \sum_i \beta_i \Phi^T(\mathbf{g}_i) \Phi(\mathbf{g}_i) - \sum_{i,j} \beta_i \beta_j \Phi^T(\mathbf{g}_i) \Phi(\mathbf{g}_j) \right) \\ & \text{s.t. } 0 \leq \beta_i \leq C \ \& \ \sum_i \beta_i = 1 \end{aligned} \quad (3)$$

In this paper we use the *Gaussian* kernel  $K(\mathbf{g}_i, \mathbf{g}_j) = \Phi^T(\mathbf{g}_i) \Phi(\mathbf{g}_j) = e^{-q \|\mathbf{g}_i - \mathbf{g}_j\|^2}$  with width parameter  $q$  to present the dot products as a *Mercer* kernel [1]. Incorporating the constraints of formula (2), the distance between image points and the center of sphere is defined as

$$R^2(\mathbf{g}) = \|\Phi(\mathbf{g}) - \mathbf{a}\|^2 = 1 - 2 \sum_j \beta_j K(\mathbf{g}_j, \mathbf{g}) + \sum_{i,j} \beta_i \beta_j K(\mathbf{g}_i, \mathbf{g}_j) \quad (4)$$

In view of (3) and (4) we get the optimum  $\beta_i$  and the optimum hypersphere radius  $R$ . The contours defined by the set  $\{\mathbf{g} : R(\mathbf{g}) = r\}$  are formed in the data space.

Combining with the Karush–Kuhn–Tucker (KKT) complementarity condition [14] we conclude:

- (1) A point  $\Phi(\mathbf{g}_i)$  only with  $\beta_i \neq 0$  can be used to define the optimum center, then we have:
  - (1) A point  $\Phi(\mathbf{g}_i)$  with  $0 < \beta_i < C$  is mapped to the surface of the feature space sphere. The points as  $\mathbf{g}_i$  will be called Support Vectors (SVs) which satisfy  $R(\mathbf{v}_i) = R$  and lie on cluster boundaries. The points lying on the same boundary contour in data space confirm the cluster boundary of the same type radar signal instances, and the boundary shape depends on kernel function.
  - (2) A point  $\Phi(\mathbf{g}_i)$  with  $\beta_i = C$  is mapped to the outside of the feature space sphere. The points as  $\mathbf{g}_i$  will be called Bounded Support Vectors (BSVs), which imply  $r > R$  and lie on external contours. These points denote the outliers or noise.
- (2) A point  $\Phi(\mathbf{g}_i)$  with  $\beta_i = 0$  is mapped to the inside of the feature space sphere. The points as  $\mathbf{g}_i$  lie on internal contours. And the same internal contours confirm the same type radar signal instances.

## 2.2 Support Vector Clustering with Cone Clustering Labeling Method

In order to solve the problem that contours cannot differentiate between points that belong to different clusters, the complete graph (CG) and support vector graph (SVG)

cluster labeling methods are discussed in [1]. The improved method proximity graph (PG) and gradient descent (GD) are proposed in [11, 23]. But these methods still take on a high time complexity because they sample a line segment to decide whether a pair of data points is in the same cluster. So these algorithms cannot meet the real-time and accuracy requirements of sorting radar signals in the dense and complex electromagnetic environment.

For decreasing the consuming time of the cluster labeling method ulteriorly, the cone clustering labeling (CCL) algorithm in [10] is applied to label clusters. This algorithm avoids the clustering inaccuracy caused by calculating the adjacent matrix with sampling a path between two data points which are in the same cluster while decreasing the time complexity.

The main idea of this algorithm is to find cones which include the image of SV, namely  $\Phi(\mathbf{v}_i)$ , and cover a key portion of the minimal hypersphere in feature space. In data space these cones correspond to hyperspheres. And a data point  $\mathbf{g}_j$  with the image of  $\Phi(\mathbf{g}_j)$  which is inside of the cones lies inside of the hyperspheres. The data points as  $\mathbf{g}_j$  possess strong intra-class compactness and satisfy  $\|\mathbf{v}_i - \mathbf{g}_j\| \leq \|\mathbf{v}\|_i - \Phi^{-1}(\mathbf{a})\|$ . The union of the hyperspheres forms an approximate covering of the data space contours. All the SVs are clustered and then the data points corresponded to SVs in the same class form a single cluster using the approximate covering. Thus the remaining data points can be easily clustered.

The formula  $\|\mathbf{v}\|_i - \Phi^{-1}(\mathbf{a})\| = x^{1/2}$  is deduced in [10], where  $x = -\ln(\sqrt{1 - R^2})/q$ . Setting  $Z = \|\mathbf{v}\|_i - \Phi^{-1}(\mathbf{a})\|$ , then the main CCL algorithm can be described as follows:

- Step 1. Compute  $Z$  for  $q$ .
- Step 2. Compute the *Euclid* distance between pairs of SVs. If the distance is less than  $2Z$ , the data points pairs lie in the same cluster.
- Step 3. Repeat the step 2 until finishing the SVs clustering.
- Step 4. Compute the distance, namely  $d$ , between the rest of data  $\mathbf{g}_j$  and SV, and then rank the  $\mathbf{g}_j$  to the class which contains the nearest SV from  $\mathbf{g}_j$ .
- Step 5. Repeat the step 4 until all the remaining data points are clustered.

### 3 Radar Signal Sorting System Based on CMSVC and SE Index

Traditional cluster validity index cannot evaluate the clustering results well to find the partitioning that best fits the radar data as the pulse signal flow is badly interleaved. So a new-type validity index is required to fit the complicated signal environment. The radar signals with period repetition from one emitter strongly resemble themselves, and the ones from different emitters have poor resemblance. Based on this, the resemblance degree can be used to depict the resemblance by calculating the resemblance coefficient between the signal flows. The resemblance coefficient is defined as

$$S(\mathbf{g}_1, \mathbf{g}_2) = \frac{\mathbf{g}_1^T \mathbf{g}_2}{\|\mathbf{g}_1\| \|\mathbf{g}_2\|} \quad (5)$$

where  $\{\mathbf{g}_{1,2}(u), u = 1, 2, 3\}$  denotes the sample vector of two emitter signals,  $\|\mathbf{g}\| = (\sum \mathbf{g}^2(u))^{1/2}$ ,  $S \in (0, 1]$ .  $S$  can be viewed as the resemblance probability of

two sample vectors. Under this consideration, in order to get better clustering results, we always expect that the within-cluster resemblance should be as strong as possible after clustering and between-cluster resemblance should be as poor as possible. According to the physical significance of information entropy and the characteristics of radar pulse signal, the information entropy can be combined with the resemblance among signals. The poorer resemblance the sample vectors have, the bigger value the entropy is, and vice versa. Thus, we try to introduce the information entropy index to represent the resemblance level of clustering signals. The index is proportional to within-cluster resemblance and inversely proportional to between-cluster resemblance. Within-cluster resemblance is denoted by the within-cluster similitude entropy  $H_{\text{comp}}(C)$  and between-cluster resemblance by the between-cluster similitude entropy  $H_{\text{sep}}(C)$ . The index is called the similitude entropy (SE), which is defined as

$$\text{SE} = H_{\text{sep}}(C)/H_{\text{comp}}(C) \quad (6)$$

where  $H_{\text{sep}}(C)$  denotes the between-cluster similitude entropy and  $H_{\text{comp}}(C)$  denotes within-cluster one. These two variables are expressed as follows:

$$H_{\text{sep}}(C) = \sum_{k=1}^c \min_{l=1, \dots, c, l \neq k} \{H_{lk}\}, \quad H_{\text{comp}}(C) = \frac{1}{c} \sum_{k=1}^c H_{kk} \quad (7)$$

where

$$H_{lk} = -\frac{1}{n_k} \sum_{i=1}^{n_k} S_{l,ki} \log S_{l,ki} \quad (8)$$

$$S_{l,ki} = S(\mathbf{m}_l, \mathbf{g}_{ki}) = \frac{\mathbf{m}_l^T \mathbf{g}_{ki}}{\|\mathbf{m}_l\| \|\mathbf{g}_{ki}\|} \quad (9)$$

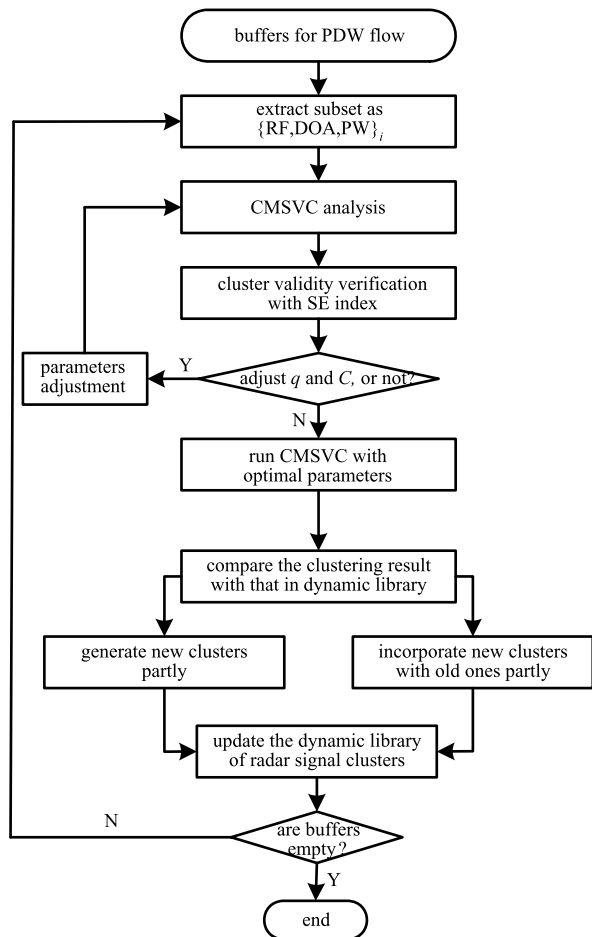
where  $\mathbf{m}_l$  denotes the center of cluster  $C_l$ ,  $\mathbf{g}_{ki}$  denotes the  $i$ th sample vector in cluster  $C_k$ , and  $S_{l,ki}$  denotes the resemblance coefficient between  $\mathbf{m}_l$  and  $\mathbf{g}_{ki}$ , respectively. When  $l = k$ ,  $H_{lk}$  denotes the within-cluster similitude entropy of  $C_l$ . When  $l \neq k$ ,  $H_{lk}$  denotes the within-cluster similitude entropy between  $C_l$  and  $C_k$ .

Hence, a greater value of SE predicates a better compactness and separation of the clusters. So by maximizing the value of SE the correct clustering result could be determined. Then we run CMSVC algorithm with different value of  $q$  and get different class numbers of  $c$ ,  $c \in [2, N - 1]$ . In the spirit of the analysis mentioned above we conclude that the similitude entropy value should satisfy  $\text{SE} = \max\{\text{SE}_c, 2 \leq c \leq N - 1\}$ .

It can be viewed as a cluster validity index whereas the SE index can depict the within-cluster compactness and between-cluster separation efficiently while the PDW subset  $G$  is partitioned to  $c$  classes under the term SE index. Meanwhile, in view of the low time complexity of CCL algorithm, we propose a sorting method for radar signal aggregating CMSVC algorithm and SE index as shown in Fig. 1. The main steps are as follows:

Step 1. Extract the subset  $G$  of PDW by a real-time processing method in subsection, where  $i$  denotes the  $i$ th extracting operation and  $G = \{RF, DOA, PW\}_i$ .

**Fig. 1** Framework of radar signal sorting system based on the combination of CMSVC and SE index



- Step 2. Pre-sort the sample vectors constituted with  $RF$ ,  $DOA$  and  $PW$  via running CMSVC algorithm and adjust the parameters of  $q$ , the Gaussian kernel, and  $C$ , the soft margin constant with SE index.
- Step 3. Sort the sample vectors of  $G$  using optimum parameters for final output which would be processed according to SE index once more and update the dynamic library of radar signal clusters.

#### 4 Algorithm Implementation

In this section we introduce the parameters adjustment algorithm with the proposed validity index SE and perform the CMSVC algorithm with these optimal parameters. Finally, the dynamic library of radar signal clusters is updated using the SE index.

#### 4.1 Adjust Parameters Using SE Index

In SVC sorting algorithm, as  $q$  increases the cluster boundaries become more and more rough, and forming an increasing number of clusters, the same as in CMSVC. Thus, The SVC algorithm runs from a small value of  $q$  which increases by using the heuristic algorithm. And the heuristic algorithm is used to get the optimal value of  $q$ . In the beginning starting out with the penalty factor  $C = 1$ , we do not allow for any outliers or need to handle the BSVs. If the number of SVs is excessive, or a number of singleton clusters form when we heuristically increase the value of  $q$ , the clusters should contain BSVs. One should decrease the value of  $C$  to smooth cluster boundaries and avoid the impact on clustering accuracy caused by BSVs [1]. Thus the correct radar signal sorting result is observed. The parameter adjustment algorithm with SE index is typically implemented by the following steps:

- Step 1. Initialize the value of  $q$  as  $q = 1 / \max_{ij} \|\mathbf{g}_i - \mathbf{g}_j\|^2$ .
- Step 2. Perform the CMSVC algorithm with the parameter of  $q$  and obtain a temporary clustering result.
- Step 3. Evaluate the following two conditions according to the clustering result in Step 2:
  - (1) whether the number of SVs is excessive;
  - (2) whether a number of singleton clusters form.
 If any one of conditions (1) or (2) occurs, we heuristically decrease the value of  $C$  and go to Step 2, otherwise go to Step 4.
- Step 4. Compute the SE validity measure according to the clustering result obtained in Step 2.
- Step 5. If the maximum SE is obtained, go to Step 6. Otherwise heuristically increase the value of  $q$  and go to Step 4.
- Step 6. Identify the clustering parameters of  $q$  and  $C$  and perform the clustering algorithm with these two parameters again. Then stop and output the best clustering result.

#### 4.2 Update the Dynamic Library of Radar Signal Clusters

When the real-time radar signal sorting method is implemented, it is necessary to compare the data clusters between existing ones in dynamic library and new obtained ones. Via this operation it can decide whether one cluster in the new clusters or in existing ones belongs to the same class. We solve this problem using the following method. Assuming that the new clusters set obtained in Sect. 4.1 is  $U' = \{C_l, l = 1, 2, \dots, c'\}$ , and the existing clusters set in dynamic library is  $U = \{C_k, k = 1, 2, \dots, c\}$ . Then the algorithm of updating the dynamic library of radar signal clusters is as follows:

- Step 1. Initialize  $l = 1$ .
- Step 2. Calculate the SE validity measure between new cluster  $C_l$  and clusters set  $U$  according to (8), and these values are expressed as  $\{H_{lk}, k = 1, 2, \dots, c\}$ .
- Step 3. Sort the  $H_{lk}$  ascendingly, namely  $H_{lt_1} < H_{lt_2} < \dots < H_{lt_p} \dots < H_{lt_c}$ , where  $\{t_1, t_2, \dots, t_p, \dots, t_c\}$  denotes a rank of  $k$ .



**Table 1** The radar pulse signal parameters information

Radars	$PRF/KHz$		$RF/MHz$		$PW/us$	$DOA/^\circ$	$N_s$
	Type	Range	Type	Range			
1	Jittered	0.52–0.58	Agile	2400–2550	1.15–1.25	28–40	1988
2	Staggered	0.8–1	Jumping	2550–2650	1.15–1.2	48–60	3264
3	Jittered	0.76–0.84	Agile	2855–3050	1.25–1.35	35–45	2891
4	Staggered	0.5–0.6	Agile	3050–3150	0.95–1.05	40–60	1973

Where  $N_s$  means number of pulses

Step 4. Initialize  $p = 1$ .

Step 5. Calculate the  $SE_{old}$  of clusters set  $U$  firstly, then incorporate the new cluster  $C_l$  into the existing cluster  $C_{t_p}$ , and form a cluster of  $C'_{t_p}$ , and calculate the  $SE_{new}$  of clusters set  $U$  after replacing  $C_{t_p}$  with  $C'_{t_p}$ . If  $SE_{new} \geq SE_{old}$ , the incorporation is confirmed and go to step 6, otherwise retract the incorporation. When the incorporation is retracted, if  $p < c$ , set  $p = p + 1$  and repeat step 5, otherwise add the new cluster  $C_l$  to clusters set  $U$  and set  $c = c + 1$ .

Step 6. If all the new clusters have been processed, stop the algorithm; otherwise set  $l = l + 1$  and go to step 2.

## 5 Simulation Experiment Result

In order to verify the efficacy of the proposed SE index and the low time complexity of the CCL algorithm, a series of radar pulse signal data is simulated. The experiments data used in this paper include 10116 radar pulse signals. The simulated data after preprocessing is as in Table 1. Here, we compare the validity measures, the DB [4], Dunn [5] and PS index [3] and the proposed SE validity measure to illustrate the performances. We adopt the datasets of data1, data2 and data3 from the simulated pulse datasets of the sample vectors of 1–200, 201–400 and 401–600 to compare the performances. And the result is as in Table 2. Meanwhile more experiments are conducted in Table 3. The accuracy and consuming time are compared in Table 4 using the different cluster labeling method of CG, SVG, PG, GD and CCL under the guidance of SE index. In Table 4 the accuracy is equal to  $(n - n_1 - n_2)/n$ , where  $n$  denotes the number of all sample vectors,  $n_1$  denotes that of the missing vectors when the datasets are clustered and  $n_2$  denotes that of vectors clustered incorrectly. The worst-case asymptotic time complexity means sorting signals without BSVs.

It is proved that when the DB and PS index is minimized, the crisp and compact clusters in the data can be obtained in [4] and [3] using the clustering method. As an indication of the optimal clustering scheme, the Dunn presented in [5] is the point at which it takes its maximum value.

However, from the above experimental comparison, we affirm that using these three indices we cannot well determine the optimal value of  $q$  or the correct number of clusters that exist in our dataset of data1, data2 and data3, which means the dense

**Table 2** The optimal number of clusters and value of  $q$  comparison using several validity indices

	$q$	$c$	DB	Dunn	PS	SE
Data1 (2 classes)	<b>220</b>	<b>2</b>	<b>0.157</b>	1.010	<b>0.001</b>	<b>3.014</b>
	310	3	0.431	<b>1.014</b>	0.010	0.706
	330	4	0.385	0.657	0.035	0.014
	345	5	0.743	0.033	0.019	0.170
	346	6	0.654	0.033	0.021	0.140
Data2 (4 classes)	6	2	0.654	<b>0.222</b>	0.022	1.778
	11	3	<b>0.126</b>	0.170	0.007	0.707
	<b>41</b>	<b>4</b>	0.131	0.177	<b>0.005</b>	<b>1.948</b>
	110	5	0.176	0.183	0.007	0.665
	130	6	0.384	0.119	0.022	0.431
Data3 (4 classes)	0.8	2	<b>0.107</b>	0.005	0.064	0.524
	16	3	1.382	0.144	0.010	1.279
	<b>66</b>	<b>4</b>	0.285	<b>0.204</b>	0.003	<b>1.388</b>
	100	5	0.269	0.027	<b>0.002</b>	0.723
	140	6	0.570	0.022	0.029	0.324

Where  $q$  is the width parameter of *Gaussian* kernel and  $c$  is the clusters number obtained when the signal is sorted by SVC with CCL labeling method under different  $q$

radar signal pulse flow cannot be sorted efficiently with these three indices. For instance, Index Dunn proposes the partitioning of data1 into three clusters as the best partitioning while the correct number of clusters fitting the dataset is two. Moreover, the Indices DB and Dunn select the clustering scheme of three and clusters, respectively, for data2 which actually contains four clusters. In the case of data3, the Indices DB and PS select two and five clusters, respectively, as the optimal scheme while the correct number is four.

On the contrary to DB, Dunn and PS indices, SE finds the correct number of clusters for all these three datasets. It means that the SE validity index can indicate the best sorting result efficiently for radar signal under the complicated electromagnetic environments.

While in order to verify further that the strategy can improve efficiency without sacrificing sorting accuracy, we introduce Q. Guo's pulse data which include 5687 radar pulse signals. These supplementary data are as in Table 3. Based on this, more experiments are conducted. These signals are sorted with the method of Multiple-parameter Radar Signal Sorting. And the parameters are adjusted using the SE index.

Here the first group data including 10116 radar pulse signals is called "G 1<sub>st</sub>", and the second group data including 5687 radar pulse signals "G 2<sub>nd</sub>". After several statistics about the sorting results, accuracy, and runtime comparisons with other cluster labeling methods are as in Table 4.

Table 4 illustrates that the computing efficiency is enhanced extremely when we sort the signals with CCL cluster labeling method compared to CG, SVG, PG and GD method. Meanwhile it can keep the sorting accuracy when the CCL algorithm

**Table 3** The radar pulse signal parameters information from Q. Guo

Radars	PRF/KHz		RF/MHz		PW/us	DOA/°	Ns
	Type	Range	Type	Range			
1	Slippery	0.3–0.4	Agile	2080–2250	1.2–1.3	48–60	824
2	Jittered	0.3–0.4	Agile	2750–2850	1–1.1	60–80	823
3	Staggered	0.8–1	Jumping	2250–2350	1.2–1.25	68–80	2149
4	Jittered	0.7–0.9	Agile	2550–2750	1.3–1.4	56–64	1891

Where  $N_s$  means number of pulses

**Table 4** Accuracy and runtime comparison using several cluster labeling methods under SE

Cluster labeling method	Accuracy (%)		Worst-case asymptotic time complexity	Total execution time (s)	
	G 1 <sub>st</sub>	G 2 <sub>nd</sub>		G 1 <sub>st</sub>	G 2 <sub>nd</sub>
CG	98.92	98.12	$O(mN^2N_{sv})$	87.646	53.283
SVG	98.92	98.12	$O(mNN_{sv}^2)$	72.967	43.121
PG	97.93	97.61	$O(N^2 + mNN_{sv})$	41.419	25.785
GD	97.89	97.56	$O(mN^2(k + N_{sv}))$	144.614	83.599
<b>CCL</b>	<b>98.92</b>	<b>98.12</b>	<b><math>O(NN_{sv})</math></b>	<b>6.013</b>	<b>4.380</b>

Where  $k$  is the number of iterations for GD to converge to a stable equilibrium points,  $m$  is the number of sample points along a line segment typically adopted as 10 to 20 and  $N_{sv}$  is the number of SV

enhances the computing efficiency compared to CG cluster labeling method used in [7]. In addition, compared to the others CCL method decreases the consuming time a lot as the dataset become large. To sum up, the proposed method keeps the advantage of sorting signal with SVC; and that it is more suitable for the radar signal processing system.

## 6 Conclusions

For the purpose of reducing the processing time of the radar signal sorting, enhancing the sorting accuracy and meeting the real-time and accuracy requirements of electronic warfare, we introduce the CMSVC algorithm into the radar signal sorting in complicated electromagnetic environments. Meanwhile we use information entropy theory to verify the clustering validity and update the library of radar signal clusters dynamically. And we conduct performance simulations for both typical methods and our proposed method. Experimental results show that the proposed method can reduce the computing complexity to decrease the consumed time in course of sorting the signals and also it can keep the sorting accuracy.

**Acknowledgements** This paper was supported in part by the National Natural Science Foundation of China under the Grant nos. 61203268 and 61202339, by the National Defense Technology Keystone Laboratory Foundation, China under the Grant nos. 9140C610301080C6106 and 9140C6001070801, by the Aviation Science Foundation, China under the Grant nos. 20095596014 and 20101996009, and supported by the Shaanxi Natural Science Foundation of China under the Grant no. 2012JQ8034.

## References

1. A. Ben-Hur, D. Horn, H.T. Siegelmann, V. Vapnik, Support vector clustering. *J. Mach. Learn. Res.* **2**, 125–137 (2001)
2. F. Camastra, A. Verri, A novel kernel method for clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(5), 801–805 (2005)
3. C.H. Chow, M.C. Su, E. Lai, Symmetry as a new measure for cluster validity, in *2nd WSEAS International Conference on Scientific Computation and Soft Computing* (2002), pp. 209–213
4. D.L. Davies, D.W. Bouldin, A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1**(2), 224–227 (1979)
5. J.C. Dunn, Well separated clusters and optimal fuzzy partitions. *J. Cybern.* **4**, 95–104 (1974)
6. G.H. Guo, Y. Mao, J. Han, X. Qing, An improved algorithm of PRI transform. *Intelligent systems, in WRI Global Congress on. IEEE*, vol. 3 (2009), pp. 145–149
7. Q. Guo, C.H. Wang, Z. Li, Support vector clustering and type-entropy based radar signal sorting method. *J. Xian Jiaotong Univ.* **44**(8), 63–67 (2010)
8. J. Han, M.H. He, Z.B. Zhu, J. Wang, Sorting unknown radar emitter signal based on the complexity characteristics. *J. Electron. Inf. Technol.* **31**(11), 2552–2556 (2009)
9. M.S. Hansen, K. Sjöstrand, H. Ólafsdóttir, H.B.W. Larsson, M.B. Stegmann, R. Larsen, Robust pseudo-hierarchical support vector clustering. *Image Anal.* **4522**, 808–817 (2007)
10. S.H. Lee, K.M. Daniels, Cone cluster labeling for support vector clustering, in *Proceedings of 6th SIAM Conference on Data Mining* (2006), pp. 484–488
11. J. Lee, D. Lee, An improved cluster labeling method for support vector clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(3), 461–464 (2005)
12. D.W. Lee, J.W. Han, W.D. Lee, Adaptive radar/pulses clustering based on density cluster window, available. [http://www.ieice.org/proceedings/ITC-CSCC2008/pdf/p1377\\_P1-104.pdf](http://www.ieice.org/proceedings/ITC-CSCC2008/pdf/p1377_P1-104.pdf). Accessed 16 June 2012
13. P. Ling, C.G. Zhou, X. Zhou, Improved support vector clustering. *Eng. Appl. Artif. Intell.* **23**(4), 552–559 (2010)
14. T. Liu, S. Shamai, A note on the secrecy capacity of the multiple-antenna wiretap channel. *IEEE Trans. Inf. Theory* **55**(6), 2547–2553 (2009)
15. C.H. Wang, Separation of composite defect patterns on wafer bin map using support vector clustering. *Expert Syst. Appl.* **36**(2), 2554–2561 (2009)
16. J.S. Wang, J.C. Chiang, A cluster validity measure with a hybrid parameter search method for the support vector clustering algorithm. *Pattern Recognit.* **41**(2), 506–520 (2008)
17. J.S. Wang, J.C. Chiang, An efficient data preprocessing procedure for support vector clustering. *J. Univers. Comput. Sci.* **15**(4), 705–721 (2009)
18. B.Y. Wang, F. Wan, Classification of single-trial EEG based on support vector clustering during finger movement. *Lect. Notes Comput. Sci.* **5552**(1), 354–363 (2009)
19. L.H. Wang, H.G. Ma, X.Y. Zhang, G.X. Zhang, Multi-component linear FM signal detection based on support vector clustering. *J. Electron. Inf. Technol.* **29**(11), 2661–2664 (2007)
20. J. Wang, P. Lei, D. Yang, W. Li, X.Y. Yan, A novel deinterleaving algorithm of radar pulse signal based on DSP, in *IEEE International Symposium on Industrial Electronics* (2009), pp. 1899–1903
21. M.R. Widyanto, H. Hartono, Angle decrement based Gaussian kernel width generator for support vector clustering. *Asian J. Inf. Technol.* **7**(8), 388–393 (2008)
22. R.G. Wiley, *ELINT: The Interception and Analysis of Radar Signals*, 2nd edn. (MA, Artech House, Boston, 2006), pp. 317–356
23. J. Yang, V. Estivill-Castro, S.K. Chalup, Support vector clustering through proximity graph modeling, in *Proceedings. 9th International Conference on Neural Information Processing* vol. 2 (2002), pp. 898–903
24. C.Y. Yeh, C.W. Huang, S.J. Lee, Multi-kernel support vector clustering for multi-class classification. *Int. J. Innov. Comput. Inf. Control* **6**(5), 2245–2262 (2010)
25. Z.F. Yu, F. Ye, J.P. Luo, A multi-parameter synthetic signal sorting algorithm based on clustering, in *8th International Conference on Electronic Measurement and Instruments* (2007), pp. 2363–2366
26. L. Zhang, W.D. Zhou, L.C. Jiao, Kernel clustering algorithm. *Chin. J. Comput.* **25**(6), 587–590 (2002)
27. G.X. Zhang, H.N. Rong, W.D. Jin, Intra-pulse modulation recognition of unknown radar emitter signals using support vector clustering. *Lect. Notes Comput. Sci.* **4223**(1), 420–429 (2006)
28. C.H. Zhao, G.Q. Zhao, D.X. Liu, The deinterleaving of radar pulse trains with stagger PRI. *J. Xidian Univ.* **30**(3), 381–385 (2003)