

RESEARCH

Open Access



# A novel sparse representation algorithm for AIS real-time signals

Shuaiheng Huai<sup>\*</sup>  and Shufang Zhang

## Abstract

Sparse representation of signals based on a redundant dictionary is a new signal representation theory. Recent research activities in this field have concentrated mainly on the study of dictionary design and sparse decomposition algorithms. Currently, the application of sparse representation on an Automatic Identification System (AIS) signal still requires further investigations. In this paper, a novel sparse representation of the AIS signal is proposed based on an adaptive redundant dictionary. Considering the characteristics of the AIS signal, an adaptive redundant dictionary is constructed using the  $K$  singular value decomposition (K-SVD) algorithm. Furthermore, an effective pursuit algorithm is proposed to obtain the sparse representation of AIS signal using the adaptive dictionary. The binary AIS message is demodulated from the sparse representation of AIS signal. The experimental results indicate that the sparse representation of the AIS signal has high accuracy and the reconstructive error rate can be under 10%; thus, the reconstructive precision is simultaneously guaranteed. The processing time of the proposed sparse representation algorithm is less than 26.7 ms which satisfies the requirements of AIS real-time signal processing. It shows that introducing the signal sparse representation in a real-time signal system obtains a satisfactory result.

**Keywords:** Automatic Identification System (AIS) signal, Sparse representation,  $K$  singular value decomposition (K-SVD), Real-time signal

## 1 Introduction

The AIS is composed of shore-based stations and shipborne equipment. It is a digital vessel navigation aid system that can identify vessels, assist in tracking targets, streamline information exchange and provide other auxiliary information to avoid collisions, etc. [1, 2]. The International Maritime Organization (IMO) has already required that all vessels should be equipped with both space- and land-based dual-positioning system in the future. To achieve the urgent requirements of IMO, the utilisation of the existing AIS shore stations has been studied for positioning. Therefore, the AIS can serve as a land-based wireless positioning system for coastal vessels, and the shipborne AIS equipment can perform both communication and positioning functions, which is referred to as the AIS autonomous positioning system (AAPS) [3]. The AAPS achieves autonomous positioning by measuring very high frequency (VHF) radio signals from AIS base stations which are established all over the world. However,

the existing AIS is essentially a communication system, and the positioning function has not been considered in the initial design and construction of the system. There are many technical problems exist that need to be solved using the AIS shore station for positioning, for example, how to extract location information from the AIS [4, 5] and develop the secondary phase correction model [6, 7] and the new carrier phase-measurement technology. The carriers currently used in the positioning system are all dual-phase-modulated carriers with the same frequency, and the equipment can be accurately positioned by carrier phase-measurement technology. However, according to the demands of communication, the carrier of an AIS signal is dual-frequency Gaussian-filtered minimum shift keying (GMSK) modulated, which is much more complicated than the dual-phase modulation. The AIS signal cannot be applied to the carrier tracking and measurement method in traditional positioning system because of its characteristics in terms of frequency difference and non-periodicity. Therefore, a new carrier measurement method of AIS signal is needed, which is referred as the AIS signal holographic correlation detection. Since the

<sup>\*</sup> Correspondence: [huaishuaiheng@dlmu.edu.cn](mailto:huaishuaiheng@dlmu.edu.cn)

Information Science and Technology College, Dalian Maritime University, Dalian 116026, China

AIS is a real-time system, in the method, obtaining the data of AIS signal within a certain time interval is the primary problem to be addressed. To solve this problem, the sparse representation theory is introduced into the AIS in the present study.

In the field of signal processing, an important issue is how to utilise the space transform to effectively express the signal and improve the compression efficiency. Conventional signal representation is usually based on 'base expansion' such as the classical Fourier transform and the wavelet transform. However, this type of signal expansion, which is based on an orthogonal basis, suffers from some limitations. For example, it cannot always achieve a desired result, especially for signals that have a wide variation range in the time and frequency. An ideal method of signal expansion should be based on the characteristics of the processed signal and adaptively select the appropriate base function to complete the decomposition of the signal. Mallat and Zhang proposed the cogitation of the signal sparse representation on a redundant dictionary in 1993 [8]. The so-called redundant dictionary was an ultra-complete redundant function library. The result of the signal sparse decomposition on the redundant dictionary was that the coefficients of most of the base functions in the signal expansion were zero, and only a few base functions had large non-zero coefficients. The base functions in the present study are the elements in the dictionary that are called atoms. Therefore, the main features of the signal can be expressed by a small number of atoms. The compressed sensing theory proposed by Candes, Romberg, Tao and Donoho in 2004 raised the signal sparse representation to a new level [9–11]. The compressed sensing theory was different from the traditional signal processing which was based on the Nyquist sampling theory. It indicated that as long as the signal contains sparsity in a certain space, it can be reconstructed with high probability at a sampling frequency that is much lower than the Nyquist sampling frequency. In this theoretical framework, the sampling rate depends on the structure and content of the signal and not on the bandwidth of the signal. Therefore, further research on the signal sparse representation is extremely significant in terms of theoretical and practical applications. The primary issue of signal sparse representation is the construction of a redundant dictionary. The existing redundant dictionary schemes cannot always guarantee sparsity of the signal [12]. In recent years, methods to obtain a redundant dictionary through learning and training have been vigorously developed, one of which is the  $K$  singular value decomposition (K-SVD) [13].

In the present study, we construct the redundant dictionary adapted to AIS signal based on the K-SVD algorithm and carry out the sparse representation of AIS signal using the orthogonal matching pursuit (OMP)

algorithm. The suitable parameters of the sparse representation are calculated considering the noise. The binary codes demodulated from the sparse AIS signal are obtained. The experimental results show that the whole sparse representation method exhibits good performance. The adaptive redundant dictionary based on K-SVD makes the number of atoms required for the sparse decomposition small, and the accuracy of the AIS signal sparse representation is simultaneously guaranteed. The sparse representation of the AIS signal is achieved with high accuracy within the required time, and the error rate of the demodulated sparse AIS signal is within the required range. We can obtain the data we need from the sparse representation of AIS signals, and the amount of data that needs to be processed reduces greatly as well as the time when compared with the original AIS signals. The work presented in this paper lays the foundation for realising the AIS signal holographic correlation detection mathematical mode. It is an important part of realising the autonomous positioning function of the AIS. The proposed method provides a reference to signal processing systems that need to obtain or save information from real-time signals.

The remainder of this paper is organised as follows. Section 2 introduces the principle of sparse representation, including the AIS signal model and the theory of signal sparse representation. Section 3 presents the sparse representation method including the adaptive redundant dictionary construction algorithm based on K-SVD and the sparse decomposition algorithm. Section 4 provides and discusses the experimental results. Section 5 concludes this paper.

## 2 Principle of sparse representation

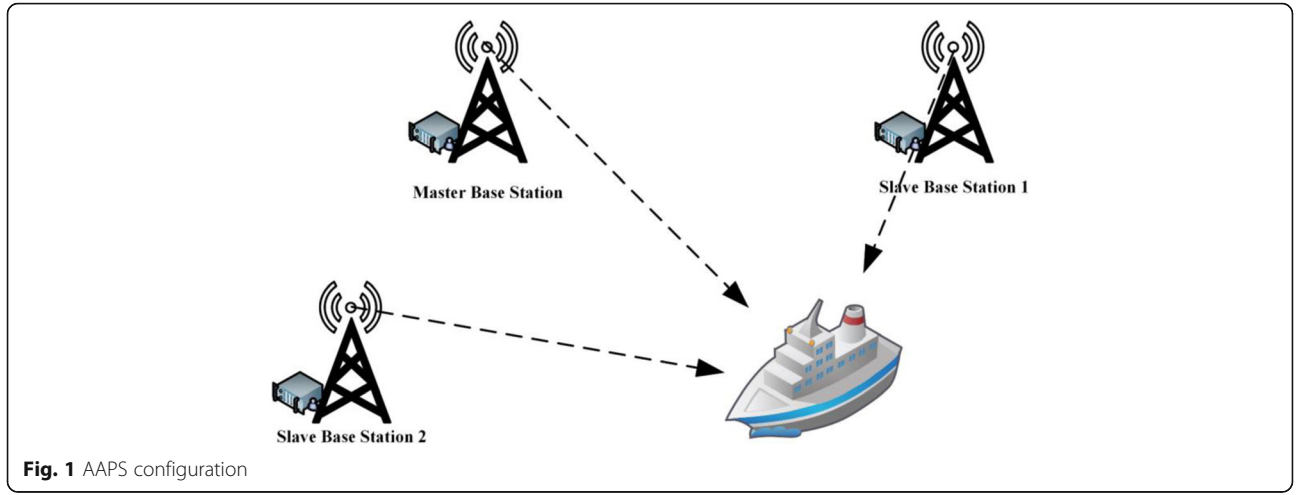
### 2.1 AIS signal model

The AAPS is composed of a master AIS base station, some slave base stations and shipborne AIS equipment. The system configuration is shown in Fig. 1. All the base stations take turns in sending signals, and the shipborne AIS equipment receives these signals. An additional secondary factor correction system is present, which is used to improve the positioning accuracy.

The existing AIS operates in the VHF band of 161.975 and 162.025 MHz with a bandwidth of 25 kHz. The signal modulation method is GMSK, and the transmission rate is 9.6 kb/s. The AIS signal model can be viewed as a GMSK signal model and expressed as

$$\begin{aligned} s(t) &= A \cos[\omega_c t + \theta(t)] \\ &= A [\cos\theta(t) \cos\omega_c t - \sin\theta(t) \sin\omega_c t] \end{aligned} \quad (1)$$

The phase  $\theta(t)$  is expressed as



**Fig. 1** AAPS configuration

$$\begin{aligned}\theta(t) &= \int_{-\infty}^t \omega d\tau = 2\pi h \int_{-\infty}^t \sum_{n=-\infty}^{\infty} a_n g(\tau - nT_b) d\tau \\ &= \sum_{n=-\infty}^{\infty} a_n q(t - nT_b)\end{aligned}\quad (2)$$

where  $g(t)$  is the rectangular impulse response of the Gaussian filter,  $q(t)$  is the integral of  $g(t)$  and the modulation index  $h$  equals 0.5, i.e.

$$q(t) = 2\pi h \int_{-\infty}^t g(\tau) d\tau = \pi \int_{-\infty}^t g(\tau) d\tau \quad (3)$$

The transfer function of the Gaussian filter is expressed as

$$H(f) = \exp(-\alpha^2 f^2) \quad (4)$$

where  $\alpha$  is a parameter related to the Gaussian filter with 3-dB bandwidth  $B$ . The definition of  $B$  is given as

$$H^2(B) = \frac{1}{2} \quad (5)$$

$$\exp(-2\alpha^2 B^2) = \frac{1}{2} \quad (6)$$

$$B = \sqrt{\ln 2 / 2\alpha^2} \quad (7)$$

The impulse response of the Gaussian filter is expressed as

$$h(t) = \frac{\sqrt{\pi}}{\alpha} \exp\left(-\frac{\pi^2}{\alpha^2} t^2\right) \quad (8)$$

GMSK is the MSK modulation of a baseband digital signal processed by a Gaussian low-pass filter. The signal characteristics such as the mutual influence degree between adjacent symbols are related to the parameters of the Gaussian filter. The value of the product with 3-dB bandwidth  $B$  of the Gaussian filter and input symbol

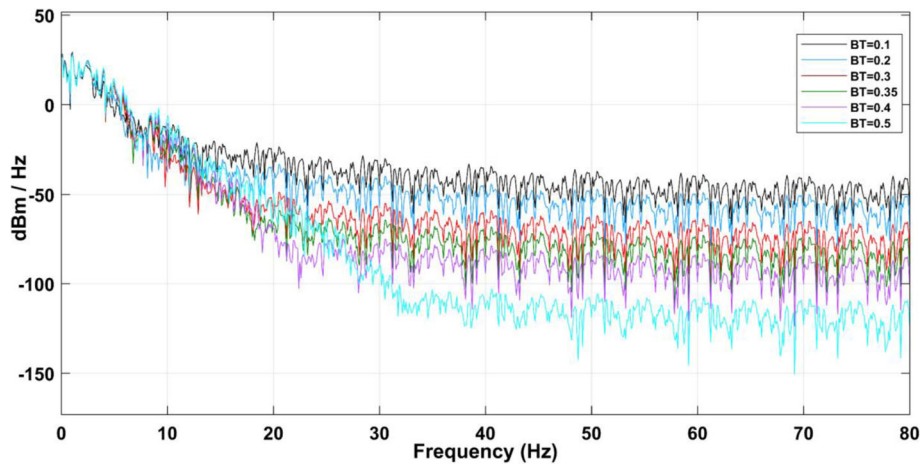
width  $T$  is usually considered as among the main parameters for designing a Gaussian filter. According to the AIS standard specification, the maximum  $BT$  product of the GMSK modulator used for transmitting data in the AIS should be 0.4, and the maximum figure should be 0.5 when receiving data [14].

Figure 2 shows that the smaller the  $BT$  product is, the slower is the decay of the GMSK signal-power spectral density with the frequency and the bigger is the main lobe. In this paper, for convenience of calculation, the  $BT$  product is set to 0.3. According to Recommendation ITU-R M.1371-4, the AIS data transmission should begin with a 24-bit demodulator training sequence (preamble) consisting of one segment synchronisation, as shown in Fig. 3. This segment should consist of alternating zeros and ones using the non-return-to-zero-inverted encoding. This sequence always starts with a zero [14].

## 2.2 Sparse representation theory

The recent years have witnessed a rapid growth in interest in the research of signal sparse representation. From the Fourier transform to the wavelet transform to the multi-scale geometric analysis, the purpose of the researchers is to provide a more concise and direct analysis of signals in different function spaces. All these transforms are aimed at discovering the signal features and obtaining the sparse representation of the signal. The signal sparse representation is derived from the nonlinear approximation theory. Given dictionary  $D = \{d_k, k = 1, 2, \dots, K\}$ , its atoms are the unit vectors that span the whole Hilbert space  $H = R^N (K \geq N)$ . For any signal  $x \in H$ , we adaptively select  $m$  atoms for  $m$  approximation of signal  $x$  in  $D$ .

$$x_m = F_m x = \sum_{k \in I_m} \alpha(k) d_k = \sum_{k \in I_m} \frac{\langle x, d_k \rangle}{\langle d_k, d_k \rangle} d_k \quad (9)$$



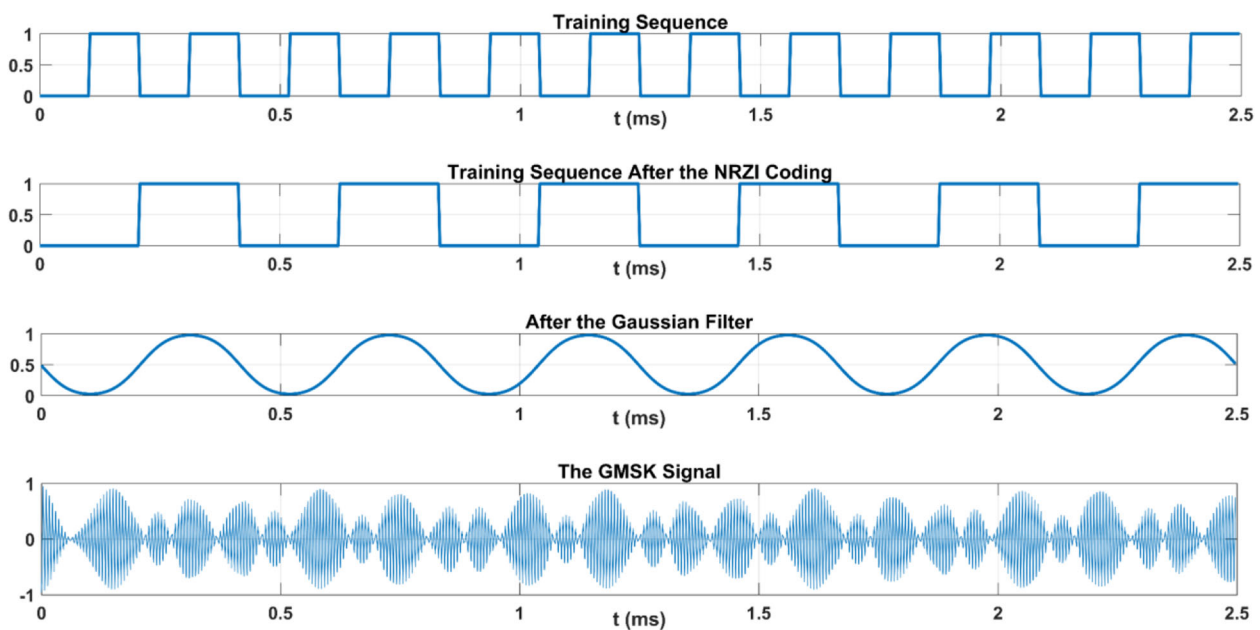
**Fig. 2** Power spectral density: the power spectral density of GMSK signal. Lines of different colours represent different BT values

in which  $\text{Card}(I_m) = m$  and  $\alpha = [\alpha(1), \dots, \alpha(K)]^T$  is the vector of the decomposition factor. The approximation absolute error is  $\varepsilon_m = \|x - F_m \alpha\|_2$ . When  $D$  is an orthogonal base, we preserve the largest  $m$  atoms of  $|\alpha(k)|$ , and the optimal  $m$  approximation of signal  $x$  is obtained. When  $D$  is a redundant dictionary, multiple solutions of Eq. 9 are available. The objective of the signal sparse representation is to choose the one with the sparsest coefficient or that with the smallest  $m$  value [15]. For a random redundant dictionary, this is a nondeterministic polynomial (NP) problem.

### 3 The sparse representation method

#### 3.1 Adaptive redundant dictionary training algorithm based on K-SVD

The singular value of a matrix is a mathematical concept, usually obtained by singular value decomposition (SVD). SVD has a wide range of applications in data compression, such as image compression and noise reduction. Singular values often correspond to the important information hidden in the matrix, and the importance is positively correlated with the size of the singular value. The K-SVD algorithm is based on this concept. In the AAPS, we need



**Fig. 3** Training sequence: the training sequence of AIS signal at different stages

to obtain and save the AIS signal within a certain time interval and we cannot lose its main features. In this regard, we construct the dictionary using the K-SVD algorithm for the AIS signal sparse representation.

For a given set of training signals, it adaptively obtains a redundant dictionary for sparse representation according to the training characteristics of the signals and sparse constraint conditions. K-SVD is characterised by the alternating application of a norm sparse constraint and SVD algorithms as well as the synchronous update of sparse dictionary and coefficients, which have strong sparse representation ability. The algorithm is flexible and works with any pursuit algorithm.

We assume that matrix  $Y = [y_i]_{i=1}^N$  represents the training signals,  $y_i$  represents one of these signals,  $D = [d_k]_{k=1}^K$  represents the dictionary to be trained and  $X = [x_i]_{i=1}^N$  is the sparse coefficient. The essence of dictionary training is the optimization problem shown by the following equation:

$$\min_{D, X} \{ \|Y - DX\|_F^2 \} \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leq T_0 \quad (10)$$

in which  $T_0$  represents the sparseness.

In order to implement the K-SVD algorithm, we summarise the specific steps, as follows:

1. Initialisation

We assign an initial value to dictionary matrix  $D$ . In general, we simply set the training signals in the matrix columns, and the optimisation problem is transformed into a problem of finding sparse representation  $X$  of sample set  $Y$ .

2. Sparse coding

We use OMP algorithm to compute the sparse coefficient, i.e.

$$\min_{x_i} \{ \|y_i - Dx_i\|_2^2 \} \quad \text{s.t.} \quad \|x_i\|_0 \leq T_0 \quad i = 1, 2, \dots, N \quad (11)$$

3. Dictionary update

We update the dictionary column by column. We suppose that the column that needs to be updated is the  $k$ th column of the dictionary matrix, which is expressed as  $d_k$  and  $x_k$  represents the  $k$ th row of matrix  $X$ . Equation 10 is thus rewritten as

$$\begin{aligned} \|Y - DX\|_F^2 &= \left\| Y - \sum_{j=1}^K d_j x_j \right\|_F^2 \\ &= \left\| \left( Y - \sum_{j \neq k} d_j x_j \right) - d_k x_k \right\|_F^2 \\ &= \|E_k - d_k x_k\|_F^2 \end{aligned} \quad (12)$$

in which  $E_k$  is the error in the  $k$ th column of the dictionary and can be expressed as

$$E_k = Y - \sum_{j \neq k} d_j x_j \quad (13)$$

Thus, we decompose the multiplication  $DX$  to the sum of  $K$  rank 1 matrices.  $K-1$  terms are assumed fixed, and one remains in question.

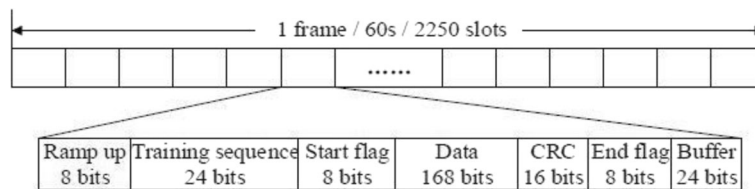
Applying SVD on  $E_k$ , we can obtain

$$E_k = U \Delta V^T \quad (14)$$

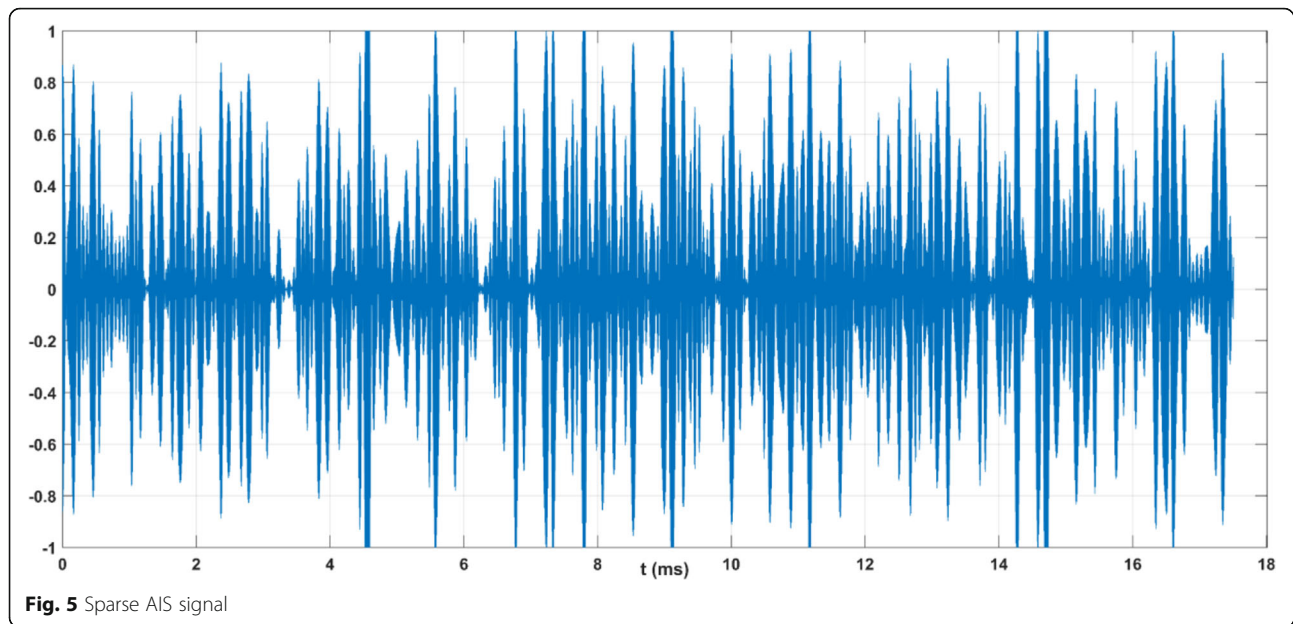
The resulting first column of  $U$  is an update of the  $k$ th column of the dictionary, and the first column of  $V$  multiplied by  $\Delta(1, 1)$  is an update of the  $k$ th coefficient vector. When all the columns of the dictionary are updated, we repeat step 2 until the iteration is completed.

Two conditions are available for stopping the iteration. In detail, the first one is to set the number of iterations. The second one is to set a fixed sparse error rate so that the iteration stops when the target value is reached.

An important issue of the K-SVD algorithm is its convergence [16, 17]. In the dictionary update stage of the K-SVD algorithm, SVD can ensure that the mean square error decreases or remains unchanged without affecting the sparse constraint conditions. The mean square error after iteration monotonically decreases, which ensures that the K-SVD converges to a local minimum. However, this does not mean that the convergence of the K-SVD algorithm is established. Instead, it depends on the convergence of the pursuit algorithm. Fortunately, the classical matching pursuit (MP), OMP and basis pursuit (BP) algorithms all exhibit good performance when  $T_0$  is sufficiently small; thus, convergence of the K-SVD can be guaranteed.



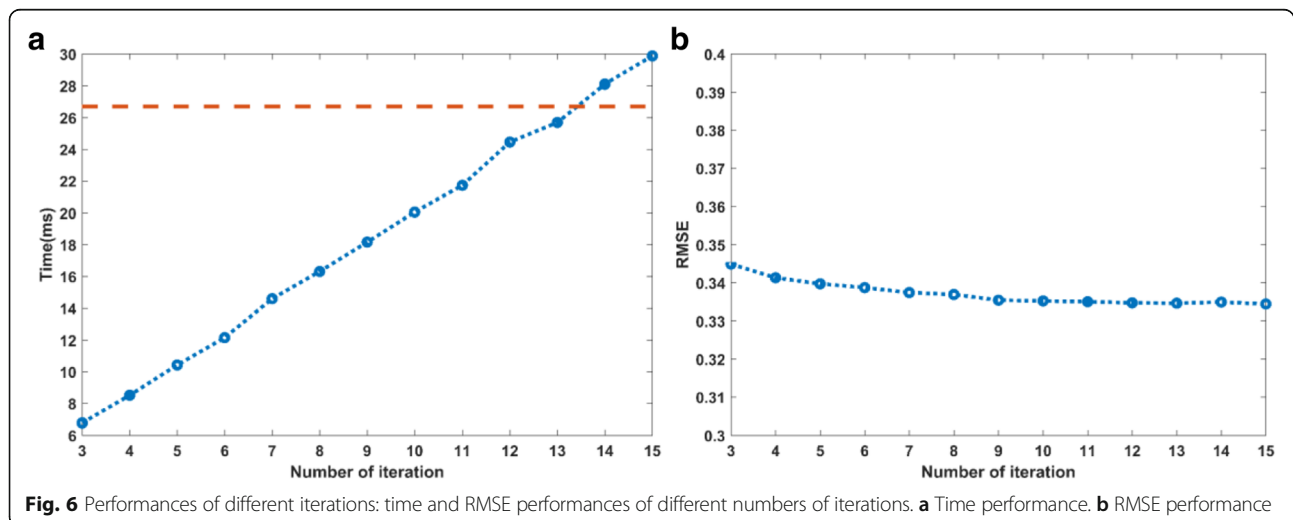
**Fig. 4** Structure of AIS signal

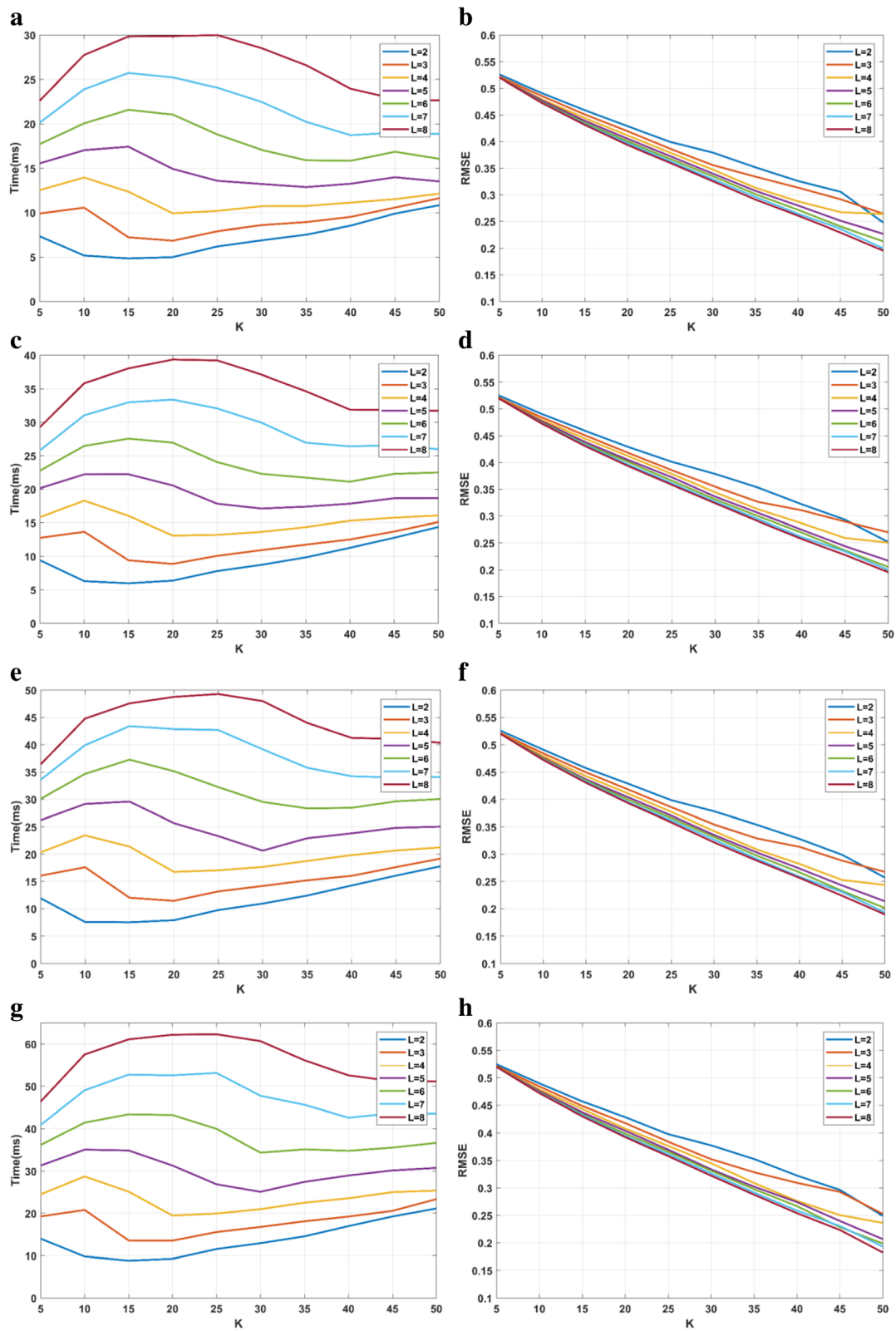


### 3.2 Orthogonal matching pursuit

As discussed in Section 2.2, Eq. 9 is an NP problem. On the surface, solving it appears to be hopeless. However, according to the premise that the signal is sparse, this problem can be solved. Researchers have conducted extensive and in-depth research on this issue and proposed many algorithms for the signal sparse approximation, such as the convex relaxation and greedy algorithms [18, 19]. BP is one type of convex relaxation algorithms. It has the advantages of global optimisation, but its computational complexity is very high. MP is a typical greedy algorithm. It converges faster than the BP, but it does not have global optimality. The basic idea of the MP algorithm is to select an atom (one column) that best matches the signal from

the dictionary matrix to construct a sparse approximation and calculate the signal residual. Then, it iteratively continues to select the atom that best matches the signal residual. The signal can be represented by the linear combination of these atoms and the last residual value. Obviously, if the residual value is negligible, the signal is a linear combination of these atoms. However, if the signal is non-orthogonal in the vertical projection of the selected atom, this will make the result of each iteration suboptimal, thus requiring more iterations. In this circumstance, the idea of OMP was proposed [20, 21]. The main improvement of OMP is to orthogonalise all the selected atoms at every step of the decomposition, which leads to faster convergence than the MP under the same accuracy





**Fig. 7** (See legend on next page.)

(See figure on previous page.)

**Fig. 7** Performances of different  $K$  and  $L$ : time and RMSE performances of different values of  $K$  and  $L$ . **a** Time performances when iteration is 6. **b** RMSE performances when iteration is 6. **c** Time performances when iteration is 8. **d** RMSE performances when iteration is 8. **e** Time performances when iteration is 10. **f** RMSE performances when iteration is 10. **g** Time performances when iteration is 12. **h** RMSE performances when iteration is 12

requirements. The OMP procedures can be described as given the matrix  $A$ , the vector  $b$  and the error threshold  $\varepsilon_0$ , approximate the solution of

$$\min_x \|x\|_0 \quad \text{s.t. } Ax = b \quad (15)$$

The specific steps summarised are as follows:

1. Initialisation

We initialise  $k = 0$  and set the initial solution  $x_0 = 0$ , the initial residual  $r_0 = b - Ax_0 = b$  and the initial solution support  $S_0 = \text{supp}\{x_0\} = \emptyset$ .

2. Iteration

Compute the errors

$$\varepsilon(j) = \min_{z_j} \|a_j z_j - r_{k-1}\|_2^2 \quad (16)$$

for all  $j$  using the optimal choice

$$z_j^* = a_j^T r_{k-1} / \|a_j\|_2^2 \quad (17)$$

Find a minimizer  $j_0$  of  $\varepsilon(j)$ ,  $\forall j \notin S_{k-1}$  and  $\varepsilon(j_0) \leq \varepsilon(j)$ .

Update  $S_k = S_{k-1} \cup \{j_0\}$ .

Compute  $x_k$ , find a minimizer of  $\|Ax - b\|_2^2$  subject to  $\text{supp}\{x\} = S_k$ .

Compute  $r_k = b - Ax_k$ .

The iteration stops when  $\|r_k\|_2 < \varepsilon_0$ . Otherwise, do another iteration.

## 4 Results and discussion

In this section, we assess the performance of the proposed sparse representation algorithm in AIS via numerical simulations. Figure 4 shows that in AIS, the data part of the default transmission packet is 168 bits; considering this as an example, we produce 64 data signals with dimension of 168 to train the dictionary. Since the AIS is a real-time system, the signal processing time must satisfy the AIS requirements. Figure 4 shows that one frame in AIS is 60 s and divided into 2250 slots; thus, one slot is 26.7 ms. The time of the sparse representation of the AIS signal must be shorter than a slot, that is 26.7 ms [14]. We explore the effects of three parameters on the final results: the number of coefficients in each linear combination, number of iterations and size of the dictionary. The choice of these three parameters has a decisive influence on the time. The sparse representation accuracy is also calculated, and we use the root mean square error (RMSE) as representative

of the accuracy. Considering the impact of noise, the SNR is set to 10 dB [22]. Figure 5 is the sparse AIS signal obtained. The hardware specifications used in this study are Intel Core i7-6700 CPU at 3.40 GHz with 16-GB memory.

All the three parameters have an effect on the time for sparse representation. The computation time must be increased with the number of iteration. Thus, we tentatively set the number of coefficients in each linear combination to 5 and the size of dictionary to 30 to find a proper number of iteration (the number of iteration is preset). Figure 6a shows that as the number of iterations increases, the processing time essentially displays a rapid linear increase. However, in Fig. 6b, the reduction rate of the RMSE does not obviously sufficiently drop from 0.345 to 0.334 which indicate that increasing the number of iterations in the algorithm does not significantly improve the accuracy. Instead, the amount of data that needs to be calculated greatly increases, resulting in a substantial increase in computation time. In Fig. 6a, the orange dashed line represents 26.7 ms, to satisfy the AIS requirements, and the number of iteration should not be more than 13.

Use  $L$  to represent the number of elements in each linear combination and  $K$  to represent the size of the dictionary. To study the effects of these two parameters on the time and accuracy of the sparse representation, use  $L$  equals 5 as the median, taking the values from 2 to 8, and  $K$  equals 30 as the median, taking the values from 5 to 50. Figure 7 shows the effect of different values of  $K$  and  $L$  on the time and accuracy when the number of iteration equals 6, 8, 10 and 12.

As shown in Fig. 7a, c, e, g, as the value of  $L$  increases, the processing time shows a significant increase. With the increase of  $K$ , the time curves basically have a process of rising then descending and then tending to be flat. The larger  $L$  is, the more obvious the process is. When  $K$  reaches a certain value, the effect of  $K$  on time tends to be flat. However, considering the requirements of AIS real-time signal processing, the value of  $L$  should not be too large. Figure 7b, d, f, g shows that as  $K$  increases, the RMSE essentially shows a significant linear decline. For example, in Fig. 7b, when  $L$  equals 2, the RMSE decreases from approximately 0.53 to around 0.25. It means the larger the size of the dictionary, the more precise the sparse representation is. When the value of  $K$  is small, the differences in RMSE of different  $L$  are also small. The differences gradually increase as  $K$  increases.

Subsequently, demodulation of the sparse signals is carried out, and we compare the results with the original AIS binary messages to calculate the error rate. In Table 1, the

**Table 1** Error rate of the different parameters

K	L						
	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)
5	36.2	35.6	35.6	35.2	35.1	35.1	34.9
10	31.1	30.3	29.3	29.0	28.8	28.5	28.4
15	27.1	25.7	25.2	24.6	24.3	23.8	23.6
20	25.2	22.0	21.4	21.1	20.2	20.1	19.8
25	28.3	19.9	18.2	17.9	17.5	16.9	16.1
30	32.6	17.3	14.7	14.4	14.0	13.5	13.2
35	36.6	19.4	12.2	11.8	11.6	11.1	10.9
40	42.4	25.6	10.7	9.4	9.0	8.8	8.4
45	43.2	37.4	17.3	9.1	8.4	7.1	6.5
50	49.3	46.5	33.6	19.5	7.0	5.5	4.8

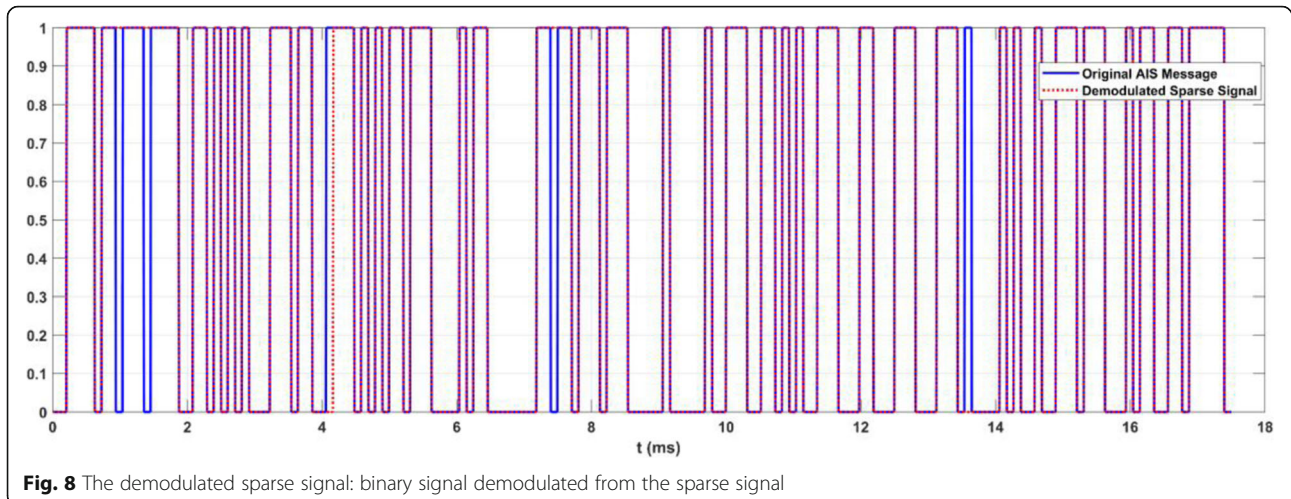
number of iterations is 6; we can see that as  $L$  increases, the error rate declines, but it is not remarkable. The error rate decreases tremendously as  $K$  increases at first. However, when  $K$  reaches a certain value (depending on different values of  $L$ ), the error rate rapidly increases. Taking  $L$  equals 3 as an example, when  $K$  equals to 50, the error rate is more than two times the value when  $K$  is 30. This is because of multiple repeated matching processes that occur on the same feature of the signal, which leads to the increase in the signal sparse representation accuracy. However, it also results in erroneous feature responses that increase the error rate. In the AIS, the frame check sequence uses the cyclic redundancy check, and an error rate of approximately 10% is acceptable [23]. According to the experimental results, considering the time of sparse representation, the accuracy of sparse representation and the demodulation error rate, a suitable combination is the size of the dictionary is set to 45 and the number of coefficients in each linear combination is set to 5.

Figure 8 is a binary signal demodulated from the sparse signal when the iteration is 6,  $K$  is 45 and  $L$  is 5.

The blue solid line represents the original AIS binary message, and the red dashed line represents the demodulated sparse signal. We can see that only five code differences exist between the two different lines; for most of the time, the two lines are coincident, which means that the error rate is quite small.

## 5 Conclusions

In this paper, a novel signal sparse representation algorithm has been proposed and applied on AIS signals. The time and accuracy performances of the sparse representation are studied with different values of iterations, number of elements in each linear combination and size of the dictionary to test whether the proposed algorithm meets the requirements of AIS real-time signal processing. For the accuracy of the AIS signal sparse representation, the number of iterations did not show much influence. However, both the number of elements in each linear combination and size of the dictionary displayed a significant effect. The adaptive redundant dictionary construction algorithm guaranteed accuracy. With regard to the signal processing time, each of the three parameters has an effect on it. The accuracy of the sparse representation increased as the size of the dictionary increases. However, a larger size cannot always guarantee a better performance by considering the error rate of the demodulated sparse signal. According to the experimental results, by considering 64 AIS signals with a dimension of 168 as an example, the best match dictionary matrix had a size of  $168 \times 45$ , the number of iterations was set to 6 and each signal was created by a linear combination of only 5 different dictionary atoms. Both the error rate and processing time satisfied the AIS real-time signal processing requirements, which implied that the proposed sparse representation algorithm can be well used in AIS.



The current work is an important part of realising the autonomous positioning function of the AIS. For further work such as the AIS signal holographic correlation detection, the proposed algorithm is applied to obtain the data of AIS signals within a certain time interval, which can significantly reduce the amount of data storage as well as the computation time. In other fields, such as the remote sensing, the detection and tracking of real-time signals with different frequencies are also necessary. The theory proposed in this paper provides a reference for real-time signal processing system.

#### Abbreviations

AAPS: AIS autonomous positioning system; AIS: Automatic Identification System; BP: Basis pursuit; GMSK: Gaussian-filtered minimum shift keying; IMO: International Maritime Organization; K-SVD:  $K$  singular value decomposition; NP: Nondeterministic polynomial; OMP: Orthogonal matching pursuit; RMSE: Root mean square error; VHF: Very high frequency

#### Acknowledgements

This work was supported by the Chinese National Science Foundation (nos. 61231006, 61501079 and 61601078).

#### Authors' contributions

SH provided the idea of algorithm, carried out the simulations, arranged the architecture and drafted the manuscript. SZ supervised the work and revised the manuscript. Both authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 11 July 2018 Accepted: 10 September 2018

Published online: 18 September 2018

#### References

- B. Hu, R.W. Liu, K. Wang, Y. Li, M. Liang, Statistical analysis of massive AIS trajectories using Gaussian mixture models. Paper presented at the Second Int. Conf. on Multimedia and Image Process, Wuhan, China 2017, 113–117
- L. Liu, X. Liu, X. Chu, J. Li, Z. Jiang, Coverage effectiveness analysis of AIS base station: a case study in Yangtze River. Paper presented at the Fourth Int. Conf. Transportation Information and Safety, Banff, AB, Canada 2017, 178–183
- Q. Hu, Y. Jiang, S.F. Zhang, Development of an automatic identification system autonomous positioning system. *Sensors* **15**, 28574–28591 (2015)
- Y. Jiang, S.F. Zhang, D.K. Yang, A novel position estimation method based on displacement correction in AIS. *Sensors* **14**, 17376–17389 (2014)
- J.B. Zhang, S.F. Zhang, J.P. Wang, Pseudorange measurement method based on AIS signals. *Sensors* **17**, 1183 (2017)
- X.Y. Wang, S.F. Zhang, X.W. Sun, The additional secondary phase correction system for AIS signals. *Sensors* **17**, 736 (2017)
- X.Y. Wang, S.F. Zhang, Evaluation of multipath signal loss for AIS signals transmitted on the sea surface. *Ocean Eng.* **146**, 9–20 (2017)
- S. Mallat, Z. Zhang, Matching pursuit with time–frequency dictionaries. *IEEE Trans. Signal Process.* **41**(12), 3397–3415 (1993)
- E. Candes, Compressive sampling. Paper presented at the Proc. Int. Cong. of Mathematicians, Madrid, Spain 2006, 3, 1433–1452
- E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Info. Theory* **52**(2), 489–509 (2006)
- D.L. Donoho, Compressed sensing. *IEEE Trans. Info. Theory* **52**(4), 1289–1306 (2006)
- T.J. Wang, B.Y. Zheng, Z. Yang, A speech signal sparse representation algorithm based on adaptive overcomplete dictionary. *Electron. Info. Technol.* **33**(10), 2372–2377 (2011)
- S. Chen, D.L. Donoho, M. Saunders, Atomic decomposition by basis pursuit. *SIAM J. Scientific Comput.* **20**(1), 33–61 (1999)
- Recommendation ITU-R M. 1371-4, technical characteristics for an automatic identification system using time-division multiple access in the VHF maritime mobile band. 2010
- A. Cohen, J. J. Dales, Nonlinear approximation of random functions. *SIAM J. Appl. Math.* **57**(2), 518–540 (1997)
- M. Aharon, M. Elad, A. Bruckstein, K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Image Process.* **15**(11), 4311–4322 (2006)
- R. Rubinstein, T. Peleg, M. Elad, Analysis K-SVD: a dictionary-learning algorithm for the analysis sparse model. *IEEE Trans. Signal Process.* **61**(3), 661–677 (2013)
- E.J. Candes, T. Tao, Near optimal signal recovery from random projections: universal encoding strategies. *IEEE Trans. Info. Theory* **52**(12), 5406–5425 (2006)
- R. Baraniuk, A lecture on compressive sensing. *IEEE Signal Process. Mag.* **24**(4), 118–121 (2007)
- Y.C. Pati, R. Rezaifar, P.S. Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. Paper presented at the Proc. 27th Asilomar Conf. Signals, Systems and Computers, Pacific Grove, CA, USA 1993, 1, 40–44
- A.M. Bruckstein, D.L. Donoho, M. Elad, From sparse solutions of systems of equations to sparse modelling of signals and images. *SIAM Rev.* **51**(1), 34–81 (2009)
- M. Zhou, A.J.V.D. Veen, R.V. Leuken, Multi-user LEO-satellite receiver for robust space detection of AIS messages. *IEEE International Conference on Acoustics* **22**(10), 2529–2532 (2012)
- T.S. Q. Sun, H.X. Zhang, H.X. Li, A satellite-borne AIS receiver based on CRC error correction and iterative interference cancellation. *Telecommunication Eng.* **56**(10), 1140–1146 (2016)

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)