Design and Implementation of Blind Source Separation Based on BP Neural Network in Space-Based AIS

Chengjie Li, Lidong Zhu, Zhongqiang Luo, Zhen Zhang, Yilun Liu and Ying Yang

In space-based AIS (Automatic Identification System), due to the high orbit and wide coverage of the satellite, there are many self-organizing communities within the observation range of the satellite, and the signals will inevitably conflict, which reduces the probability of ship detection. In this paper, to improve system processing power and security, according to the characteristics of neural network that can efficiently find the optimal solution of a problem, proposes a method that combines the problem of blind source separation with BP neural network, using the generated suitable data set to train the neural network, thereby automatically generating a traditional blind signal separation algorithm with a more stable separation effect. At last, through the simulation results of combining the blind source separation problem with BP neural network, the performance and stability of the space-based AIS can be effectively improved.

Keywords: space-based AIS, blind source separation, BP neural network, parallel processing, error function

INTRODUCTION

In space-based AIS (Automatic Identification System), satellite communications are one of important parts, due to the satellite orbit altitude is relatively high and the number of satellites increasing, the openness of satellite orbits makes satellite communications vulnerable to be jammed from external signals. To solve that problem, improving the efficiency of information processing, new technologies must be developed. The framework of space-based AIS is shown in Figure 1. In Figure 1, due to the openness of satellite channels, satellite communication is easily disturbed by external signals, therefore, it is great significance to study the security of satellite communication in space-based AIS (Kourogiorgas et al., 2015; Chen et al., 2020).

Blind separation technology can improve the efficiency of signal processing without occupying excess bandwidth. With the increasing processing power of computers, neural network technology has also been introduced into blind separation technology (Lee et al., 2008; Lahat and Christian, 2016; Fu et al., 2014). In 1985, C. Jutten and J. Herault proposed a neural network recursively connecting two initial signals to two mixed signals passing through the channel, gradient descent method is used to change the network weights, the residual error of the output signal is minimized, the separation of the signal under unknown initial conditions is accomplished. In 1991, the variability of the solution of blind separation problem is completed by L. Tong. In recent years, a robust neural network blind separation algorithm is developed by A. Cichockic and R. Unbehauen (Inan and Erdogan, 2015; Li et al., 2020; Kim et al., 2007). In (Inan and Erdogan, 2015), introduce a family of convolutive Bounded component analysis (BCA) criteria and corresponding algorithms, and prove that the global optima of the proposed criteria, under generic BCA assumptions, are equivalent to a set of perfect separators. In (Li et al., 2020), separate the mixed source signals by Particle swarm optimization.
Knowledge Background

Blind Source Separation Model

N original signals pass through the mixed channel, the signals are statistically independent of each other, and M sensors receive the mixed signal, the mixed model of the observed signal is

\[ x_j(t) = \sum_{i=1}^{N} a_{ji} s_i(t) + n_j(t), \quad i \in [1, 2, \cdots N], \quad j \in [1, 2, \cdots M] \quad (1) \]

where \( a_{ji} \) are the unknown mixing coefficients, \( s_i(t) \) are the original signals, \( x_j(t) \) are the observed signals, \( n_j(t) \) are the noise signals. The matrix form of formula (1) is

\[
\begin{bmatrix}
  x_1(t) \\
  x_2(t) \\
  \vdots \\
  x_M(t)
\end{bmatrix} =
\begin{bmatrix}
  a_{11}(t) & a_{12}(t) & \cdots & a_{1N}(t) \\
  a_{21}(t) & a_{22}(t) & \cdots & a_{2N}(t) \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{M1}(t) & a_{M2}(t) & \cdots & a_{MN}(t)
\end{bmatrix}
\begin{bmatrix}
  s_1(t) \\
  s_2(t) \\
  \vdots \\
  s_M(t)
\end{bmatrix} +
\begin{bmatrix}
  n_1(t) \\
  n_2(t) \\
  \vdots \\
  n_M(t)
\end{bmatrix} \quad (2)
\]

Among the numbers of original signal (M) and the numbers of receiving antennas (N), blind source signals can be classified into overdetermined blind separation (\( M < N \)), determined blind separation (\( M = N \)) and underdetermined blind separation (\( M > N \)) (Robert et al., 1995).

The blind source separation model in space-based AIS is slightly different from common blind separation. In this paper, consider only the signal processing methods between satellites, the probability distribution of the channel model obeys Rayleigh distribution, the probability density function is expressed as follows,

\[ f(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x \geq 0 \quad (3) \]

Back Propagation Neural Network

BP (Back Propagation) neural network was proposed by Rumelhart and McClelland in 1986, which is a multi-layer feedforward neural network trained according to error back propagation algorithm. In the BP neural network system, hidden layer connection weight learning of multi-layer neural network is solved, which has four basic features (Wu et al., 2005).

Nonlinear

Each connection unit in BP neural network has a different form, BP neural network can deal with nonlinear problems well. BP neural network contains a large number of specific connection units, and the connection mode of BP neural network is not a specific connection unit, so the BP neural network is nonlinear. The relationship between the initial data and the expected data is not linear, and a neural network is designed to train and learn the initial data and the expected data. Then the trained BP neural network system can simulate the relationship between the initial data and the expected data.

Parallel Processing

BP neural network is composed of many parallel concrete join units, and the order in which these join units are used in BP
BP neural network can change the weight of each specific connection unit according to the initial data type. In this case, change itself structure is called self-organization. The self-organization of BP neural network means that when the trained network inputs different types of initial data again, it can adjust the weight according to the change of the initial data type. In this case, change itself structure is called adaptive ability.

BP neural network is a kind of multi-layer feedforward network trained by using the method of transmitting data error from back to forward. It is one of the most commonly used models, which includes both the process of data propagating backwards and the process of error propagating forwards. Classic BP neural network diagram contains a data input layer, an intermediate layer, and a result output layer. The task of data input layer is to receive and detect incoming data and then send it to the neurons at the next layer, the task of intermediate layer is to process and transform the information transmitted from the upper layer accordingly and sends the information to the final layer. These three parts constitute the BP neural network, which transmit data backwards. The diagram of BP neural network is as follows in Figure 2 (Jin et al., 2000).

**Frequency-Hop Signal Model**

In space-based AIS, considering to the frequency-hop signal model (Loof and Pratt, 2019a) (Loof and Pratt, 2019b),

\[
f(t) = \sqrt{2S} \sum_{k} \text{rect}_{\tau_{H}}(t - kT_{H} - aT_{M}) \cdot e^{i2\pi f_{k}(t - kT_{H} - aT_{M}) + j\theta} + n(t), 0 < t \leq L
\]  

(4)

In formula (4), \(L\) is the length of signal, \(T_{H}\) is the hop duration, \(\text{rect}_{\tau_{H}}\) stands for is the rectangular window, \(f_{k}\) is the carrier frequency of the \(k\)th hop, \(aT_{M}\) is the frequency hopping time, \(\theta\) is the phase of a frequency hopping signal, \(S\) is the signal power; \(n(t)\) is the Gaussian white noise. The frequency hopping signal distribution in time-frequency domain is described in Figure 3. In Figure 3, the time-frequency domain image of frequency hopping signal is determined by the following three factors: the position of the \(k\)th hop in the time domain, the position of the \(k\)th hop in the frequency domain and the length of the time-frequency domain.

**ALGORITHM EXECUTION PROCESS**

In the algorithm execution process, BP neural network is the main body of algorithm execution.

**Forward Propagation**

Sampling data input formula is as follows (Porter et al., 2014; Balouchestani et al., 2014; Logeswari et al., 2014),

\[
n_{i}^{m} = \sum_{j=1}^{m-1} \omega_{i}^{m} j + b_{i}^{m}, m = 1, 2, \cdots M (M \geq 2)
\]  

(5)

In formula (5), \(n_{i}^{m}\) stands for the connection value of neurons in \(m\) row and \(i\) column of BP neural network, \(M\) stands for the number of layers of the neural network. The output of \(m\) layer is \(y_{m} = f^{m}(\mathbf{w}_{m}), m\) is the number of layers of the neural network, \(f^{m}\) is the transfer function at the \(m\)th layer. Then, the output of the first layer is \(y_{1} = f^{1}(\mathbf{w}_{1}^{T}x + b_{1})\), the output of the last layer is \(y = y_{M}\).

**Back Propagation**

Back propagation is the main characteristic of BP neural network, in this paper, BP neural network is mainly composed of four parts.

**Error Function**

In this paper, MSE (Mean Squared Error) function is chosen as the error function of BP network, the expression of MSE is as follows,

\[
F(z) = E(e^{2}) = E[(y_{i} - \hat{y}_{i})^{2}]
\]  

(6)

here, \(y_{i}\) is the output of the BP neural network true value, \(\hat{y}_{i}\) is the output of the BP neural network estimated value.
Weight Correction Method

In the above error function (6), gradient descent method as the weight correction method, the iterative formula is as follows,

\[
\begin{align*}
\omega_{i,j}^{m}(k+1) &= \omega_{i,j}^{m}(k) - \eta \frac{\partial F}{\partial \omega_{i,j}^{m}} \\
b_{i}^{m}(k+1) &= b_{i}^{m}(k) - \eta \frac{\partial F}{\partial b_{i}^{m}}
\end{align*}
\]  

(7)

(8)

here, \( \eta \) stands for training rate.

Chain Rule

To minimize MSE, it’s necessary to take the partial derivative of the objective function, the chain rule is the rule that you have to follow when you take partial derivatives, suppose \( F \) is the objective function, then the chain rule is as follows,

\[
\frac{dF(n(w))}{dw} = \frac{dF(n)}{dn} \cdot \frac{dn}{dw}
\]  

(9)

According to formula (9), formulas (7) and (8) can be expressed as,

\[
\begin{align*}
\frac{dF}{dw_{i,j}^{m}} &= \frac{dF}{dn_{i}^{m}} \cdot \frac{dn_{i}^{m}}{dw_{i,j}^{m}} \\
\frac{dF}{db_{i}^{m}} &= \frac{dF}{dn_{i}^{m}} \cdot \frac{dn_{i}^{m}}{db_{i}^{m}}
\end{align*}
\]  

(10)

(11)

In BP neural network, \( s_{i}^{m} = \frac{\partial F}{\partial n_{i}^{m}} \) represents the sensitivity of the iteration, then formulas (10) and (11) can be expressed as

\[
\begin{align*}
\frac{\partial F}{dw_{i,j}^{m}} &= s_{i}^{m} s_{j}^{m-1} \\
\frac{\partial F}{db_{i}^{m}} &= s_{i}^{m}
\end{align*}
\]  

(12)

(13)

Weights Updated

In the above process, according to gradient descent method, the connection value of the BP neural network is as follows,

\[
\begin{align*}
\omega_{i,j}^{m}(k+1) &= \omega_{i,j}^{m}(k) - \eta \omega_{i,j}^{m} y_{j}^{m-1} \\
b_{i}^{m}(k+1) &= b_{i}^{m}(k) - \eta \omega_{i}^{m}
\end{align*}
\]  

(14)

(15)

BP Neural Network Algorithm

According to the 3.2, the steps of BP neural network in this paper are as follows.

1) Calculate the forward propagation value of BP neural network

\[
\begin{align*}
n_{m} &= W_{m} y_{m-1} + b_{m}, m = 1, 2, \ldots, M (M \geq 2) \\
y^{m} &= f^{m}(n^{m}) \\
y &= y^{m}
\end{align*}
\]  

(16)

(17)

(18)

2) Calculate the back propagation value of BP neural network

\[
\begin{align*}
s^{M} &= -2(f^{M})'(n^{M})(t - y) \\
s^{m-1} &= (f^{m-1})'(n^{m-1}) (W_{m})^{T} s^{m}, m = M, \ldots, 2, 1 (M \geq 2)
\end{align*}
\]  

(19)

(20)

3) Change connection value

\[
\begin{align*}
s^{m-1} &= (f^{m-1})'(n^{m-1}) (W_{m})^{T} s^{m}, m = M, \ldots, 2, 1 (M \geq 2) \\
b^{m}(k+1) &= b^{m}(k) - \eta s^{m}
\end{align*}
\]  

(21)

(22)
SIMULATION RESULTS

Parameter Setting
In this paper, determined blind separation \( (M = N) \) is considered, the number of original signals is equal to the number of received antennas, \( M = N = 3 \), the parameters are set as follows.

Sparse Representation
Since the target signal is processed as sparse signal, the sparse processing should be carried out (Meganem et al., 2014).

Definition
Dictionary: It is a collection of waveforms, that is, an analytic function library or a signal sample library, which may not satisfy orthogonality.

Atom: Basic waveforms in dictionaries, that is the basic carrier and representation unit of information and can represent any signal by superposition.

Theorem
Atomic Decomposition Theorem: \( x \) is a random signal, and the dictionary \( \Phi \) consists of many column vectors, each of which is an atom \( \phi_y \), then \( x \) can be expressed as a weighted sum of some optimal atoms, \( x = \sum_{y \in \Gamma} \alpha_y \phi_y = A \alpha \), here \( \alpha_y \) is a decomposition coefficient, \( y \) is the atomic number, \( \Gamma \) is the set of atomic serial numbers.

Residual Signal: In fact, it’s very difficult to match \( \alpha_y \) and \( \phi_y \), in general, random signals \( x \) is obtained by means of signal approximation, \( x = \sum_{m=1}^{m} \alpha_y \phi_y + r_m \), here \( r_m \) is the residual signal, the smaller the value of \( r_m \), the higher the matching accuracy.

Sparse Representation: Sparse representation is actually an optimization problem, that is, the signal has as few atoms as possible, but it can’t be distorted, suppose \( \| \alpha \|_0 \) is the objective function of the optimization problem, then the mathematical model of sparse representation is as follows,

\[
\min \| \alpha \|_0, \quad \text{s.t.} \quad x = \sum_{y \in \Gamma} \alpha_y \phi_y
\]

Separation Simulation Results
After the signal is sparse, the signal separation performance will be discussed in this subsection. In this paper, the number of transmitting antennas and receiving antennas is equal to three, the original signal are sparsely processed according to 4.2 and the waveforms are shown in Figure 4, in Figure 4, the X-axis represents sparse time and the Y-axis represents \( f \), the expression for \( f \) is described in formula (24),

\[
f = A \cdot m(N) \ast \cos \left(2\pi \cdot \frac{f_c}{f_s} \cdot N\right)
\]

In the next experiment, we aim to separate each signal from the received mixed signals. Considering to the realistic signal transmission, we assume that the transmission channel is a Gaussian channel, the initial signal goes through the Gaussian channel and are sparsely processed according to 4.2, the mixed signal’s waveforms are shown in Figure 5.
We use the proposed algorithm in this article, in BP neural network algorithm, the number of hidden layers is 100, the processing flow chart is given in Figure 6. In the data training process, 70% of the data set is used as the training set, 15% of the data set as the verification set and the remaining 15% of the data set as the test set. After the proposed algorithm in this article, the original signals are recovered efficiently, the separated signal waveforms are displayed in Figure 7. From Figure 7, it is clear that the signal waveform after separation is very similar to the original signal waveform.

To measure the separation effect, we compare the signals in Figure 4 and Figure 7 with correlation coefficient value as the evaluation standard. The correlation coefficient expression is given in formula (25) (Li et al., 2016),

\[ r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \sum_{i=1}^{n} (y_i - \bar{y})^2}} \]  

(25)

To measure the advantages of the algorithm in this paper, we further compare the separation performance with Classical FastICA Algorithm. The algorithm performance comparison results are shown in Figure 8. In Figure 8, the algorithm has better separation performance than Classical FastICA Algorithm (Reju et al., 2010).

**Network Security Analysis**

Blind separation belongs to signal processing, it can not only improve the information processing ability of the system, but also improve the security of space-based AIS network. To measure the security of the system, invulnerability analysis of space-based AIS network is used as the safety standard, the parameter settings are shown in Table 1. A cost function is constructed to measure the invulnerability analysis of the space-based AIS network in formula (26),
FIGURE 7 | Waveforms of separated signals.

FIGURE 8 | Separation performance comparison with classical fastICA algorithm.
where \( l_{ij} \) is the distance of node \( i \) and node \( j \).

The results of the analysis are shown in Figure 9. The results show that: 1) The invulnerability performance of space-based AIS network under random attack has significant advantages than the DE (Differential Evolution algorithm) network optimization and PSO (Particle Swarm Optimization) network optimization; 2) As the number of nodes increases, the invulnerability of the system decreases; 3) The invulnerability performance of space-based AIS network with blind source separation was significantly better than that without blind source separation (Yang et al., 2019) (Tai et al., 2017).

**CONCLUSION**

In this paper, the blind separation technique is used to deal with the conflict between signal receivers of space-based AIS network. This method not only improves the signal processing capability of the system, but also improves the security of the space-based AIS network. That fundamentally solves the security of the system and belongs to the category of Endogenous Safety and Security (ESS). Finally, the paper discusses the signal processing ability and the security of the system through the simulation experiment. The simulation results show that the algorithm in this paper has better advantages than conventional information processing methods.
DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CL completed the overall idea of the paper, LZ completed the structure of the paper, ZZ, YL, and YY completed the simulation experiments.

REFERENCES

Balouchestani, M., Sugavaneswaran, L., and Krishnan, S. (2014). Advanced K-means clustering algorithm for large ECG data sets based on K-SVD approach[C]. 2014 9th International Symposium on Communication Systems, Networks and Digital Signal. Tokyo: CSNDSP, 177–182.

Chen, S., Sun, S., and Kang, S. (2020). System Integration of Terrestrial Mobile Communication and Satellite Communication——The Trends, Challenges and Key Technologies in BSG and 6G. China Commun., 156–171.

Fu, G.-S., Phlypo, R., Li, M., Li, X.-L., and Adali, T. (2014). Blind Source Separation by Entropy Rate Minimization. IEEE Trans. Signal. Process. 62 (16), 4245–4255. doi:10.1109/tsp.2014.2333563

Inan, H. A., and Erdogan, A. T. (2015). Convolutive Bounded Component Analysis Algorithms for Independent and Dependent Source Separation. IEEE Trans. Neural Netw. Learn. Syst. 26 (4), 697–708. doi:10.1109/tnnls.2014.2320817

Jin, Wen., Zhao, Wei., Luo, Jia-Li., Wei, Si., and Zhen, Han. The Improvements of BP Neural Network Learning Algorithm. Proc. ICSP2000, 1647–1649.

Kim, T., Attias, H. T., Lee, S.-Y., and Lee, T.-W. (2007). Blind Source Separation Exploiting Higher-Order Frequency Dependencies. IEEE Trans. Audio Speech Lang. Process. 15 (1), 79–79. doi:10.1109/taslp.2006.872618

Kourougiorgas, C. I., Arapoglou, P., and Panagopoulos, A. D. (2015). Statistical Characterization of Adjacent Satellite Interference for Earth Stations on Mobile Platforms Operating at Ku and Ka Bands. IEEE Wireless Commun. Lett. 4 (No. 1), 82–85. doi:10.1109/lwc.2014.2374177

Lahat, D., and Christian, J. (2016). An alternative proof for the identifiability of independent vector analysis using second order statistics[C]. ICASSP, Shanghai, 4363–4367.

Lee, J.-H., Lee, T.-W., Iolesz, F. A., and Yoo, S.-S. (2008). Independent vector analysis (IVA): Multivariate approach for fMRI group study. Neuroimage 40 (1), 86–109. doi:10.1016/j.neuroimage.2007.11.019

Li, C., Zhu, L., Luo, Z., and Zhang, Z. (20202020). “Solutions to Data Reception with Improve Blind Source Separation in Satellite Communications,” in IEEE International Symposium on Networks, Computers and Communications (ISNCC’20) (Montreal, Canada: IEEE), 1–5. doi:10.1109/isncc49221.2020.9297206

Li, C., Zhu, L. D., and Zhang, Z. (2016). Non-orthogonal frequency hopping signal underdetermined blind source separation in time-frequency domain[J]. Infocommunications J. 8 (3), 1–7.

Logeswari, G., Sangeretha, D., and Vaidehi, V. (2014). “A cost effective clustering based anonymization approach for storing phr in cloud[C],” in International Conference on Recent Trends (Sydney: Information Technology), 1–5.

Loof, J. L., and Pratt, T. G. (2019). Frequency-Hopped Signal Source Identification in Frequency-Selective Channels. IEEE Trans. Aerospace Electron. Syst. 55 (No. 6), 3316–3329.

Meganem, I., Deville, Y., Hosseini, S., Deliop, P., and Briottet, X. (2014). Linear-Quadratic Blind Source Separation Using NMF to Unmix Urban Hyperspectral Images. IEEE Trans. Signal. Process. 62 (7), 1822–1833. doi:10.1109/tpsp.2014.2360618

Porter, R., Tadic, V., and Achim, A. (2014). Blind separation of sources with finite rate of innovation[C]. London: Signal Processing Conference (EUSIPCO), 136–140.

Reju, V. G., Soo Nqee Koh, S. N., and Ing Yann Soon, I. Y. (2010). Underdetermined Convolutive Blind Source Separation via Time-Frequency Masking. IEEE Trans. Audio Speech Lang. Process. 18 (1), 101–116. doi:10.1109/tasl.2009.2024380

Robert, H., Zaragoza, M., and Lin, S. (1995). QPSK block-modulation codes for unequal error protection[J]. IEEE Trans. Inf. Theor. 41 (2), 576–581.

Tai, W.-L., Chang, Y.-F., and Li, W.-H. (2017). An IoT notion-based authentication and key agreement scheme ensuring user anonymity for heterogeneous ad hoc wireless sensor networks. J. Inf. Security Appl. 34, 133–141. doi:10.1016/j.isa.2017.04.002

Wu, W., Feng, G., Li, Z., and Yuesheng, X. (2005). Deterministic Convergence of an Online Gradient Method for BP Neural, IEEE Transactions on Neural Networks. 16, 533–540. doi:10.1109/tnn.2005.844903

Yang, J., He, S., Xu, Y., Chen, L., and Ren, J. (2019). A trusted routing scheme using blockchain and reinforcement learning for wireless sensor networks. Sensors (Switzerland) 19, 4. doi:10.3390/s19040970

FUNDING

This work is fully supported by Natural Science Foundation of China Project (61871422), Science and Technology Program of Sichuan Province (2020YFH0071), National Natural Science Foundation of China under Grant (61801319), in part by Sichuan Science and Technology Program under Grant (2020DJQ0061), (2021YFG0099), in part by the Sichuan University of Science and Engineering Talent Introduction Project under Grant (2020RC33), Innovation Fund of Chinese Universities under Grant (2020HYA04001).

Copyright © 2021 Li, Zhu, Luo, Zhang, Liu and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.