Contribution Assessment Approach for Command and Control System Based on Force-Sparsed Stacked-Auto Encoding Neural Networks

Peng Liu 1, Zeyu Zhou 1, Ning Li 1, Bowei Zhang 1, Minglei Han 1 and Huiying Jiao 2

1 China Ship Research and Development Academy, Beijing, China
2 China Academy of Information and Communication Technology, Beijing, China

Email: delestar@126.com, zzy687096@126.com, li0702ning@163.com, bwzhang_xidian@163.com, 407517544@qq.com, 756408691@qq.com

Abstract. Aiming at the difficulty of feature extraction and related generation mechanism analysis for Command and Control (C2) System using traditional data mining method, a novel contribution assessment approach based on Force-Sparsed Stacked-Auto Encoding Neural Networks (FS-SAE) is proposed. Combined with big data and complex networks technology, the contribution assessment model to operational system of system (SoS) is built. The emergence relations between the capacity indices of C2 system are formalized. The derivation results show that formalized presentation for the emergence process of performance indices of C2 system based on the proposed model not only reflects the complexity characteristics of non-linear and uncertainty in emergence process, but also gives general-defined meaning for indices structure of C2 system. It provides a feasible method for the commanders to deeply understand, manage and control the complex operation system.

Keywords: C2 system, FS-SAE, Contribution, Emergence effect

1. Introduction

The concept of "contribution rate" first appeared in the economic field, and then it was used to describe the contribution rate of weapons to the system of operational SoS capability. The idea of evaluating the system contribution rate of weapons is to put this type of weapons in the real combat environment, and then compare and analyze its influence on the system through the "whether or not" of weapons according to the mission undertaken by the operational SoS.

Xiao-feng Hu pointed out the difference between the contribution rate of C2 system and that of weapon equipment system, and proposed the following definition of the contribution rate of C2 system [1-4]: “The contribution rate of C2 system refers to the influence degree of internal function composition on the system's operational effectiveness under the condition of approximate real combat background according to mission of operational SoS.” Therefore, it is unscientific to evaluate the contribution rate of C2 system by using the evaluation method of weapon system. The contribution of C2 system to operational SoS should be analyzed by evaluating the rise and fall of the system's internal performance indices on the operational capability.
2. Evaluation mentality

The purpose of evaluating C2 system’s contribution to operational SoS is to find the key performance indices and functions of C2 system based on the in-depth analysis of the various elements. Its essence is to establish the functional mapping relationship between characteristics of C2 system and combat effectiveness of operational SoS. The characteristics of C2 system are often described by its performance indices; and the combat effectiveness of operational SoS is characterized by mission completion. Therefore, the evaluation of C2 system’s contribution to operational SoS is mainly divided into two parts: one is to establish C2 system performance index structure; the other is to build the mapping between C2 system performance index and mission of operational SoS.

The deep neural network extracts feature extraction from original data species by constructing several hidden layers, in order to realize the formal representation and cognition of complex system law[5-7]. Taking the indices of different levels as the nodes of each layer of neural network, a set of indices structure reflecting the characteristics of C2 system can be constructed, and then the complex system can be analyzed. Its expression is as follows:

\[ E = f(H(X)) \] (1)

Where \( X \) is the initial aggregate of performance index (input set); \( H(X) \) is the network model constructed by hidden layer; \( f \) is the formal expression of neural network mapping function; \( E \) is the aggregate of mission (output set). \( H(X) \) is the main embodiment of black-box model in neural network.

The hidden layer of the network model formed by it has great randomness, which will lead to the hidden nodes have no specific reference meaning and cannot analyze the aggregation and emergence of performance indices. Therefore, it is necessary to whiten the black-box model. By using the method of big data and complex network analysis, this paper constructs the index network structure framework to describe the capability of C2 system. Taking it as a priori knowledge, Stacked Encoding Neural Networks (SAE) is used to build the contribution evaluation model to operational SoS of C2 system based on FSAE. As shown in Figure 1:

![Figure 1. Evaluation mentality of contribution rate](image)

3. Model

3.1 Input set and output set

If all the performance indices are put into the sample aggregate of performance index, its dimension will be very large. The more information extracted from the system characteristics, the more information the index aggregate can express. For an input set, the dimension is too high, which will bring great challenges to the construction of sample aggregate, training of neural network, data storage and time to process data. In order to complete the network training better and ensure the feasibility and practicability of the evaluation model, we should select the key performance indices from the C2 system according to the requirements of system assessment, and conduct unified coding processing, so as to form a standardized C2 system performance index sample aggregate.
This paper mainly considers the performance indices of C2 system which can be verified by analytical calculation or simulation, such as the types of detectable targets, the probability of recognition, the period of operational decision-making, the symbol error rate, etc. One dimensional vector \( X \) is composed of them in a fixed order, where \( k \) is the dimension of the sample aggregate of performance indices, i.e., the number of selected performance indices

\[
X = \{X_1, X_2, X_3, \ldots X_k\} \tag{2}
\]

The mission and task of combat system can be regarded as the classification problem of whether the task is achieved or not, such as whether the target is destroyed, whether the loss exchange ratio is greater than 1 or less than 1, etc. Therefore, this paper defines the output space of FS-SAE model as the vector space \( Y = \{Y_1, Y_2, Y_3, \ldots Y_l\} \), where \( l \) is the dimension of mission space.

### 3.2 Performance index network

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Maximal Information Coefficient (MIC), which belongs to Maximal Information-based Nonparametric Exploration (Mine), is used to measure the linear or nonlinear strength of two variables. For any indices \( X_{ij}, X_{jk}, (i, j \in \{1, 2, \ldots k\}) \) in the sample aggregate of performance index, their MIC is used to represent the weight of the edge connected by them. Their MIC is calculated as follows:

The data pairs formed by any two indices \( X \) in the sample aggregate of performance index are distributed in two-dimensional space. It is divided by \( m \times n \) grid, and \( \Omega \) is defined as the aggregate of \( m \times n \) grid generation methods. \( P(\alpha, \beta) \) is defined as the frequency estimation of data points in the grid \( \alpha, \beta \).

\[
P(\alpha, \beta) = \frac{\gamma(\alpha, \beta)}{\tau} \quad (\alpha \in \{1, 2, \ldots m\}, \beta \in \{1, 2, \ldots n\}) \tag{3}
\]

Where \( \gamma \) is the number of points in the grid, \( \tau \) is the number of points in the whole space.

\[
MIC = \max_{m \times n < B} \left( \frac{\max_{\alpha, \beta} P(\alpha, \beta) \log \frac{P(\alpha, \beta)}{\gamma(\alpha, \beta)}}{\min_{\alpha, \beta} P(\alpha, \beta) \log \frac{P(\alpha, \beta)}{\gamma(\alpha, \beta)}} \right) \tag{4}
\]

Where \( B \) is the resolution of grid generation, \( g \) is a method of grid generation, and \( P(\alpha) \) is the frequency estimation of all data points in the line or row. The value range of MIC is \([0, 1]\). The closer the value is to 1, the stronger the correlation between the two indices is.

### 3.3 The aggregate of functional index

According to the performance index network diagram, the symbols are defined: \( G = (V, E) \) represents the performance index network diagram, \( V \) is the aggregate of points, \( E \) is the aggregate of edges, and \( A \) is the weight matrix of edges.

FN (Fast Newman) algorithm is a community discovery algorithm based on Modularity, which is a kind of aggregation method based on greedy algorithm[8-10]. This algorithm can explore hierarchical community structure. Its optimization goal is to maximize the Modularity of the whole network. Modularity is a measure to evaluate the division of a complex network. Its essence is the difference between the number of connected edges of nodes in the community and the number of edges in the random case, and its value range is \([-1/2, 1]\).

\[
Q = \frac{1}{2m} \sum_{i,j} [a_{ij} - \frac{\delta(V_i, V_j)}{2m}] \tag{5}
\]

Where \( l_i = \sum_j A_{ij} \) is the sum of the weights of all the edges connected to \( V_i \), and \( m = \frac{1}{2} \sum_{i,j} a_{ij} \) is the sum of the weights of all the edges and \( \delta(V_i, V_j) \) is the discriminant function If \( V_i \) and \( V_j \) are in the same community, then it is 1, otherwise it is 0.

- Firstly, each node in the network is defined as a community, and the Modularity is \( Q_0 = 0 \).
- Calculate the Modularity\( Q_t \) when any two communities merge. Find \( \max_t \Delta Q_t = Q_t - Q_0 \), then let \( Q_0 = \max_t \Delta Q_t \).
• Repeat step 2 until all communities merge into one large community. Find the largest Q in the merge process. The community partition in the network corresponding to the maximum Modularity is the most suitable method of community partition.

The whole process of FN algorithm can be expressed as a dendrogram, from which the hierarchy with the largest Q is selected to get the final community structure, as shown in Figure 2. The best community division of performance index network is \( C = \{C_1, C_2, ..., C_n\} \), which is the aggregate of functional index.

![Dendrogram of FN algorithm’s process](image)

**Figure 2.** Dendrogram of FN algorithm’s process

3.4 The aggregate of community characteristic index.

In order to reflect the emergence mechanism between performance indices and their functional communities, and to exclude the coupling between performance indices, we extract the feature indices from the functional communities as the second hidden layer, so as to reflect the relationship between the aggregate of performance index and the aggregate of functional index.

Firstly, the correlation matrix \( R \) is constructed according to the maximum information coefficient between the performance indices within the community.

\[
|\lambda_I - R| = 0 \quad (6)
\]

By solving (x), the eigenvalues \( \lambda_1, \lambda_2, ..., \lambda_n \) and their corresponding eigenvectors \( \nu_1, \nu_2, ..., \nu_n \) are obtained.

\[
\nu_i = (c_{i1}, c_{i2}, ..., c_{ni})^T \quad (7)
\]

Since the eigenvectors are orthogonal, the principal component composed of eigenvectors is linearly independent. The principal components were as follows.

\[
B_j = c_{1j}X_1 + c_{2j}X_2 + ... + c_{nj}X_n, \quad j = 1, 2, ..., n \quad (8)
\]

According to the data compression principle, the cumulative contribution rate of principal components should be no less than 85%.

\[
\omega = \frac{\sum_{j=1}^{p_0} S_j}{\sum_{j=1}^{n} S_j} \geq 85\% \quad (9)
\]

Where, \( S_j \) is the variance of the jth principal component. Accordingly, select the appropriate \( p_0 \) which makes \( \omega_{p_0} \geq 85\% \) and \( \omega_{p_0-1} < 85\% \). At this time, the characteristic index set composed of the TOP \( p_0 \) previous component can effectively reflect the relationship between the performance index set and the system functional community.

4. Optimization Model

Auto Encoder Neural Network (AENN) is used between the input layer and the hidden layer of FS-SAE model. As can be seen from Fig.2(e). After being forced to sparse, the whole network model is equivalent to n SAE modules in parallel (n is the number of communities in the input index Network), and the output is the mission index through logical regression model (Softmax) indices. Therefore, the basic unit of optimization is still AENN.
4.1 AENN layer algorithm design
Sigmoid function is adopted as the activation function of AENN. For any sample aggregate \( \{x^1, x^2, x^3, \ldots, x^{(m)}\} \), the Implicit features after encoding is \( y^{(i)} = f(x^{(i)}) \).

\[
f(x^{(i)}) = \frac{1}{1 + \exp(-(Wx^{(i)} + b))}
\]

Then the decoder is used to reconstruct the implied representation \( z^{(i)} = f'(y^{(i)}) \).

\[
f'(y^{(i)}) = \frac{1}{1 + \exp(-(W'y^{(i)} + b))}
\]

Where, \( W \) and \( W' \) are the weight matrices of the input layer and the hidden layer, which usually satisfy constraints \( W = W^T \); \( b \) and \( b' \) are the bias vectors of the input layer and the hidden layer.

\[
\theta = \{W, W', b, b'\}
\]

The reconstruction error \( e^{(i)} \) is defined by Negative Log-Likelihood (NLL) function.

\[
e^{(i)} = L(x, z^{(i)}, \theta)
\]

\[
e^{(i)} = -\sum_{n=1}^{k} x_n^{(i)} \log(x_n^{(i)}) + (1 - x_n^{(i)}) \log(1 - x_n^{(i)})
\]

Where: \( k \) represents dimension of sample, \( e^{(i)} \) is the error of a single sample, and the average error is used to measure the reconstruction cost of the sample.

\[
\bar{C} = \frac{1}{M} \sum_{i=1}^{M} (e^{(i)})
\]

Where: \( M \) is the number of samples in a sample set.

4.2 Classifier
In this paper, Softmax classifier is selected as the logistic regression function of hidden layer and output layer. Softmax function has the advantages of fast convergence, high accuracy and easy cascade with AENN network. The basic principle of the function is shown in the following formula:

\[
h_{\theta}[x^{(i)}] = \frac{1}{\sum_{k=1}^{K} e^{\theta_k x^{(i)}}}
\]

The optimization algorithm of FS-SAE model still adopts the Stochastic Gradient Descent (SGD) algorithm of SAE model, while each AENN sub-module calculates update parameters separately during training optimization.

4.3 Optimization algorithm of FS-SAE
The FS-SAE model adopts the SGD algorithm. Its core idea is to randomly select a small sample subset \( B \) from the whole sample space, and estimate the expected value of the Gradient of the whole sample space by calculating the Gradient of the sample in the subset, as shown below:

\[
g = \frac{1}{M_B} \sum_{i=1}^{M_B} (\nabla_{\theta} L((x^{(i)}, y^{(i)}, \theta))
\]

The sample subset \( B \) is obtained from random sampling in the sample space. As long as the sample subset \( B \) has good distribution characteristics, it can be guaranteed \( g \) to better represent the expected value of the whole sample space gradient. Usually selected \( M_B \) is much smaller than \( M \), which greatly reduces the computational complexity of the algorithm.

5. Emergence and contribution analysis

5.1 Emergent mechanism analysis of function index
According to the definition of FS-SAE model, the emergence formula of input nodes in each community’s AENN layer for hidden layer points can be formally expressed as:

\[
h_{1j}^{(i)} = f_{1j}(x^{(i)}) = \frac{1}{1 + \exp(-(W_1x^{(i)} + b_{1j}))}
\]

\[
h_{2j}^{(i)} = f_{2j}(h_{1j}^{(i)}) = \frac{1}{1 + \exp(-(W_2h_{1j}^{(i)} + b_{2j}))}
\]
Through the mapping relationship between community indices, we can determine the emergent form of the initial index to the community characteristic index and the community characteristic index to the function index in the community structure. It should be emphasized that there are only two value states of the hidden node h, which are 0 and 1 respectively, and the calculated result of the formula is the probability that the emergent index of the hidden layer is equal to 1. Therefore, strictly speaking, the expected value of the emergent index is obtained by the formula, namely:

$$E(h_{n_j}^{(i)}) = \frac{1}{1 + \exp(-W_{nj}x^{(i)} + b_{n_j}))}$$

(20)

Where n=1,2, respectively represent the community characteristic index of the first hidden layer and the function index of the second hidden layer. This probability representation well represents the uncertainty of the emerging process of complex system capabilities and provides a method for analyzing the value range and distribution characteristics of capability indices.

5.2 Contribution rate analysis of function indices
According to the evaluation model of FS-SAE C2 system, the relationship between system functional indices and mission indices can be expressed as follows:

$$y_{h}^{[k]}[h_{2}(i)] = \frac{e^{\theta_{h}^{[k]}h_{2}(i)}}{\sum_{i=1}^{k}e^{\theta_{h}^{[k]}h_{2}(i)}}$$

(21)

Where $y_{h}^{[k]}[h_{2}(i)]$ is the expected value of the mission index $y_{h}^{[k]}[h_{2}(i)]$ when the sample is $x^{(i)}$. On this basis, we define the relative contribution degree of emergence index to mission index as follows:

$$\rho(y_{i};h_{j}) = \frac{\delta y_{i}}{y_{i}\delta h_{2j}}$$

(22)

It can be proved that:

$$\rho(y_{i};h_{j}) = \theta_{h}^{[1]}(1 - E(y_{i}))$$

(23)

Where $\theta_{h}$ is the connection weight of emergent index to mission indices. Therefore, we conclude that the relative contribution rate of emergent indices in FS-SAE evaluation model is related to the connection weight between indices and the expected value of mission indices. By measuring the response relationship between the function index and the mission index, we can monitor the mission completion ability of the C2 system, and then assist the commander to manage and control the joint operation system in a favorable direction.

6. Summary
Based on the background of C2 system operation under the condition of joint operation and the multi-layer network index system model framework obtained from big data and complex network analysis as a priori knowledge, this paper constructs a C2 system operation situation assessment model based on FS-SAE, and realizes the situation prediction and analysis. At the same time, the formal expression of emergence relationship of C2 system network index system, the analysis of emergence mechanism of system function index and the calculation of relative contribution to mission are completed. The calculation results show that FS-SAE has higher prediction accuracy than traditional sparse SAE, and it also gives relatively clear physical meaning to the hidden nodes of the model, which provides a feasible scheme for in-depth analysis of the function and influence mechanism of each function module of cognitive air defense system on mission completion. It provides an effective means to assist the commander to monitor, manage and control the C2 system.

7. Acknowledgment
In writing this paper, we have benefited from the presence our coworkers. They generously helped us collect materials we needed and made many invaluable suggestions. We hereby extend our grateful thanks to them for their kind help, without which the paper would not have been what it is.

8. References
[1] Wenfeng Wu, Xiaofeng Hu. Capacity Evaluation of Networked Operational System of Systems based on Big Data[J]. Military Operations Research and Systems Engineering.
[2] Wenfeng Wu, Xiaofeng Hu, ShengMing Guo. Analysis of Combat SoS Coordination Based on Multi-Layered Temporal Networks [J]. Complex Systems & Complexity Science, 2017, 14(2): 1-10.

[3] Xiaofeng Hu. On War System Engineering: Methodology towards to Information Age’s War [M]. Beijing: University of National Defense Press, 2013.

[4] Le Zhang, Zhong Liu, Jianqiang Zhang, et al. Reduction method of weapon system of systems assessment index system based on autoencoder [J]. Journal of Central South University, 2013, 44(10): 4130-4137.

[5] Xiaoyi Yao, Shengming Guo, Xiaofeng Hu, et al. Design evaluation indices for Air defense systems based on RF-GLR integrated machine learning algorithms [J]. Journal of command and control, 2015, 1(3): 269-277.

[6] Elaraby N M, Elmogy M, Barakat S. Deep Learning: Effective Tool for Big Data Analytics [C]. International Journal of Computer Science Engineering (IJCSE), 2016: 254-262.

[7] Bengio, Yoshua, Ian Goodfellow, Aaron Courville. Deep learning. Vol. 1. Massachusetts, USA:: MIT press, 2017.

[8] Newman M E J. Fast algorithm for detecting community structure in networks. Physical Review E [J], 2004, 69(6 Pt 2):066133.

[9] Zhao Ting, Wang Bin, Gao Qi. Congestion warning method based on the Internet of vehicles and community discovery of complex networks. The Journal of China Universities of Posts and Telecommunications [J], 2016, 23(4):37-45.

[10] Brown N, Sandholm T. Superhuman AI for heads-up no-limit poker: libratus beats top professionals [J]. Science (S0036-8075), 2017: 1733.