Visualization analysis method based on multidimensional scale transformation for indicators on simulation data of weapon system

Lizhen Zheng¹, Lu Gao and Tonggang Yu

The Department of Equipment Command and Management, Army Engineering University, Shijiazhuang, China

¹E-mail: zhlzh1981@126.com

Abstract. Visual analytic methods are a growing area of research that targets the effective description of weapon system, which capitalizing on displaying of various indicators information of weapon system simulation. We provide a method which can display all kinds of indicators and architectural features of simulation results in the same screen. First of all, the simulation data of multidimensional information index system and the corresponding dimension is mapped into the two-dimensional matrix. Secondly, the index system of two-dimensional graphs and planar map simulation data are fused by computation. Finally, distinguish the architectural features by drawing the influence range of the indicators to the systems. The method can analyze the correlation between indicators, similarity comparisons system features, and describe some kind of index and the system characteristics of the correlation degree. Because of the above characteristics of the method, we can adjust the index system, using visualization method to display the comprehensive characteristic of the weapon equipment system, and select the weapon equipment system in accordance with the specific conditions of the system.

1. Introduction

Advances in use of information technologies in military such as visualize the indicators of weapon equipment system are providing new opportunities for description of the characteristics of the weapon equipment system. Typically, the process of data analysis is often inseparable from the interaction of machines and people and complementary advantages. From this standpoint, the research on the theory and method of large data analysis can be carried out from two dimensions: one is from the perspective of machine or computer, emphasizing the computing power and artificial intelligence of machine, with various high-performance processing algorithms, intelligent search and mining algorithms as the main content, such as large data processing methods based on Hadoop and MapReduce [1] framework and various large-scale oriented methods. Machine learning and data mining are also the mainstream of research in the field of large data analysis at present. Another dimension emphasizes the analysis method based on human-computer interaction and conforming to human's cognitive law from the perspective of human being as the main body of analysis and demand. It intends to integrate human's cognitive ability, which machine is not good at, into the analysis process. Visual analysis of big data is the main representative. [2, 3] More than 80% of the information that human beings get from the outside world comes from the visual system. [4, 5] When large data is presented in the form of visual
graphics in proper form, the analyst can often get a glimpse of the hidden information behind the data and transform it into knowledge and wisdom.

Weapon equipment system is a higher-level system composed of various weapon equipment systems which are functionally interconnected and interacting under certain conditions of strategic guidance, operational command and support conditions, in order to accomplish certain combat tasks. With the advent of the information age, the composition of the combat capability of the weapon equipment system is no longer determined only by the quantity and type of the equipment, but by the network distribution, functional complimentarily, battlefield environment, combat objects, performance parameters and other comprehensive factors. These influencing factors are usually expressed by the method of comprehensive index system, of which the key indexes reflect the different characteristics of various types of weapon equipment system. The construction method of the index system mainly has the supervisor experience method, the mathematical analysis method and the machine learning and et al. These methods tend to focus only on the subjective experience judgment or the objective data analysis, which has obvious flaw in the practical application. The data visualization analysis is the best way to solve this difficult problem.

In the field of capability evaluation of weapon equipment system based on simulation data, the research on data visualization is still only the preliminary exploration stage, among which there are two kinds of visual analysis oriented to indicators: one is the analysis of single index, aiming at the numerical analysis and time series analysis of individual index of the system, the main method is the traditional way of line chart or graph, [6] and the second is the analysis of multi-dimensional index, which is aimed at the comprehensive analysis of multiple indexes of the system, and achieves the characteristics of system characteristics or ability from different sides, including the analysis of the deep relationship between indexes and systems, and its performance method involves the visualization analysis of multidimensional data.

In this paper, a visual analysis method for multidimensional indicators based on multidimensional scale transformation is proposed. The algorithm takes the data in the simulation experiment of weapon equipment system as the form of data matrix, using the idea of multidimensional scale transformation. It calculates the similarity matrix of multiple simulation system represented in the data matrix, and then fuses and gets the comprehensive similarity matrix. On this basis, it uses the kernel density function to draw the influence range density graph of single index, finally draw the interaction effect visualization effect of multidimensional index. Experiments show that this method can find the hidden system feature information and the clustering information between the systems in various weapons and equipment system schemes accurately and quickly, the correlation information between multidimensional indexes, and according to the specific task target, through the construction of differentiated index system and threshold range, which is possible to select the weapon equipment system that meets the expectations from a large number of simulation data.

2. Relative work
Multidimensional visualization technology is widely used as an effective abstract information presentation tool to assist knowledge workers to understand and analyze massive high-dimensional datasets. At present, considerable number of multidimensional visualization methods are proposed.

Parallel coordinates [7] is one of the earliest and most classical visualization techniques for representing multidimensional data in two-dimensional form. The dimension and coordinate axis are mapped, and multidimensional information is represented by straight lines or curves between multiple parallel axes. The disadvantage of the parallel coordinate is that the expression dimension is limited by the horizontal width of the screen. As the dimensionality increases, the vertical axis gets closer, causing difficulties in identification. As the data set increases, the line density increases since the identification of a large amount of overlap polyline is difficult, and its manifestations are too professional, not suitable for general style promotion is the most difficulty.

Scatter plot matrix [8] take each dimensional variable combination of two multidimensional data as an element (referred to as a panel) of the matrix. Each panel plotted scatter plot corresponding variable
multidimensional data, obtained from information implicit in each comparison between pairwise dimensional variables. Its main drawback is that when the dimension increases, the matrix will be limited by screen size, and it can only be found the relationship between the two-dimensional, hard to find relationships between multiple data dimensions.

The radioactive coordinate system, including Radial Coordinate visualization [9] and Star Coordinates [10], is a multidimensional visualization method based on the idea of circular parallel coordinate system. They use the circular K radius to represent K dimensional space, so that one point in the coordinate system represents multi-dimensional information objects. The disadvantage is that when the dimension increases or the number of data points increases, the visualization effect of the polyline group will be significantly reduced. The more important problem is that the ray coordinate system does not have the ability to express the distance between points and points, so it can not show the correlation between indexes.

The dimension reduction mapping technology is to consider the multidimensional data as points in the same dimension space, and its coordinates are determined by the corresponding dimension values. Then the data points are mapped to the low dimensional space. At the same time, the best method is used to maintain certain relations between the data points as far as possible. The most representative ones are principal component method (PCA) [11] and multi-dimensional scale change (MDS) [12]. The drawback of PCA is that the mapping data no longer keep the information of the original attribute space, while MDS is unable to map attribute information and sample information into the same plane, which it can not observe the relationship between the index and the system characteristics at a plane scale. Our method effectively overcomes this problem, and designs a comprehensive distance matrix to display the index information and body at the same time.

3. Multidimensional scaling analysis

Multidimensional Scaling Analysis (multidimensional scaling, MDS is a classic method for dimensionality reduction, but also a mean of data visualization. Rebuild the problem originated in geographic coordinates. If we can only get the matrix of the distance between any two cities, how to reconstruct their Euclidean coordinates. For example, we can get the distribution of high-dimensional data points, their absolute positions not meaningful to us, we are concerned with the distance relationship between point to point. So we can map high dimensional points to two-dimensional space through MDS method, and maintain their distance relationship well. In our method, the composition produced by this method is extended to the analysis of the characteristic similarity of the weapon and equipment system.

3.1. Basic terminology

An $n \times n$ order matrix of $D = (d_{ij})_{n \times n}$, if the condition is satisfied:

$$
D = D^T \quad \{d_{ij} \geq 0, d_{ii} = 0, i, j = 1, ..., n\}
$$

(1)

is called the generalized distance matrix, it referred to the distance between the point $i$ and the $j$-th point.

$S = (s_{ij})_{n \times n}$ is the square of generalized distance matrix, where $s_{ij} = d_{ij}^2$.

For an $n \times n$ order matrix $D = (d_{ij})_{n \times n}$, if there is a positive integer,

$$
d_{ij}^2 = \|x_i - x_j\|^2, \forall i, j = 1, ..., n.
$$

(2)

then $D$ is called the Euclidean distance matrix.

3.2. Algorithm principle

In the MDS algorithm, we do not need the specific coordinates of the data points. Assuming that the data points mean at the origin of the coordinates, we define the inner product matrix $B$. If we can get the value of the inner product matrix $B$, the corresponding data set coordinates can be obtained. From
the distance matrix $D$, we can get the similarity value of the centre distance matrix $B$, and then we can get the composition of the distance matrix $D$ from the singular value decomposition, that is, the distribution of multidimensional data in two-dimensional space.

4. Visualization analysis method based on multidimensional scale transformation
The aim of our method is to get a comprehensive map, which can retain three relationships simulation data weapon system: the relationship between systems, the correlation between the index system and relationships between indicators. Here, the relationship can be abstracted as a distance (e.g., Euclidean distance), or may be of a similar nature (e.g., Pearson correlation). For example, the correlation between index systems may be correlation index measurements, and the similarity system may be measured by Euclidean distance.

4.1. Data matrix description
Through simulation experiments, we can get the original simulation result data and transform it into the form of matrix of $m$ row and $n$ column, as shown in Table 1.

| Indicator 1 | Indicator 2 | ... | Indicator 5 | Indicator 6 |
|-------------|-------------|-----|-------------|-------------|
| scenario-1  | 0.46        | 7.3 | ...         | 25.60       | 1602.00     |
| scenario-2  | 0.53        | 10.2| ...         | 18.7        | 1871.00     |
| scenario-3  | 0.51        | 7.5 | ...         | 26.70       | 1602.00     |
| scenario-4  | 0.50        | 8.3 | ...         | 25.30       | 1804.00     |
| scenario-5  | 0.44        | 5.1 | ...         | 26.70       | 2524.00     |
| scenario-6  | 0.52        | 7.0 | ...         | 23.70       | 1909.00     |
| scenario-7  | 0.59        | 14.5| ...         | 22.90       | 1810.00     |
| scenario-8  | 0.45        | 7.4 | ...         | 25.60       | 1852.00     |
| ...         | ...         | ... | ...         | ...         | ...         |
| scenario-100| 0.48        | 8.3 | ...         | 18.20       | 1909.00     |

Here, rows represent the various weapon and equipment systems corresponding to each experimental scheme, lists the various indicators that need to be evaluated and analyzed. We use $T_i$ represents row $i$, $Z_j$ represents column $j$, and $X_{ij}$ represents the value of indicators $j$ of weapon and equipment system $i$.

$$DM = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}$$ \hspace{1cm} (3)

System data obtained by converting the data space indicator $T$ and $Z$. Data T space contains all the $m$ systems $T_i$:

$$T_i = [x_{i1}, x_{i2}, ..., x_{in}] (i = 1, 2, ..., m)$$

$$DM = [T_1 T_2 ... T_m]$$ \hspace{1cm} (4)

Data Z space contains all the $n$ systems $Z_j$:

$$Z_j = [x_{1j}, x_{2j}, ..., x_{mj}] (j = 1, 2, ..., n)$$

$$DM = [Z_1 Z_2 ... Z_n]$$ \hspace{1cm} (5)

4.2. Mapping matrix and fusion
Using Euclidean distance formula, the similarity matrix $DT$ of $m$ systems in $n$-dimensional space is:

$$DT = (d_{ij})_{m \times n}$$

$$d_{ij} = \|T_i - T_j\|$$ \hspace{1cm} (6)

Using the cosine correlation formula, the similarity matrix $DZ$ of $n$ indicators in $m$-dimensional space is obtained by (7)
\[ DZ = (\cos(i,j))_{n \times n} \]
\[ \cos(i,j) = \frac{Z_i \cdot Z_j}{||Z_i|| ||Z_j||} \]  

(7)

In TZ space, the distance of \( T_i \) and \( Z_j \) can be transformed into data item \( X_{ij} \) of matrix \( DM \), which represents the distance matrix between system and index.

\[ DTZ = DM \]  

(8)

According to the four distance matrices obtained above and our focus on image layout, we can choose different matrix combinations to complete the comprehensive mapping diagram. We chose Glimmer MDS [13] multilayer tables to perform distance retention optimization algorithms and control map layouts, as opposed to PCA or dual-plotted linear projection, it gives us more freedom to choose constraints that control layouts, such as the function of mixing distances, layout arrangements, and cartographic standards.

![Figure 1. Mapping and fusing distribution.](image)

Through the above analysis, the correlation between the index and the system can be obtained by the data matrix, where we can use a visualization method to make this association manifest in the form of density graph. In order to achieve this goal, we adopt the adaptive kernel Density estimation algorithm (AKDE) [14]. It first estimates the local density of each sample data point, and then shrinks or expands the bandwidth of the sample data point.

4.3. Association area drawing

Through the above analysis, the correlation between the index and the system can be obtained by the data matrix, where we can use a visualization method to make this association manifest in the form of density graph. In order to achieve this goal, we adopt the adaptive kernel Density estimation algorithm (AKDE) [14]. It first estimates the local density of each sample data point, and then shrinks or expands the bandwidth of the sample data point.

4.3.1. Compute the local density \( f \) of sample data point \( P \), the formula is:

\[ f(P) = \frac{1}{N} \sum_{l=1}^{N} K_H(||P - P_l||) \]  

(9)

where \( N \) represents the total number of sample data points, and \( ||P - P_l|| \) represents the distance between data points, where each point \( P_l \) has a fixed broadband \( H \).

4.3.2. Calculate the local smooth parameter \( \lambda_l \) and local bandwidth \( H_l \).

\[ \lambda_l = \left( \frac{g}{f(P_l)} \right)^2 \]  

\[ H_l = H \times \lambda_l \]  

(10)  \hspace{1cm} (11)

Here, \( G \) represents the geometric average of the local density of all sample data points.
4.3.3. Calculate the density value of the sample.

\[ x = \sum_{i=1}^{N} \frac{K_H(\|P - P_i\|)}{\sum_{j=1}^{N} K_H(\|P - P_j\|)} x_i \]  

(12)

where \( X_i \) represents the total number of sample data points.

4.3.4. Set the display range for the density cloud:

\[ \sum_{j=1}^{N} K_H(\|P - P_j\|) \geq \varepsilon \]  

(13)

\( \varepsilon \) represents the range of thresholds.

5. Experiments

We rely on the integrated weapon simulation test platform based on information system to carry out the area air defence anti-missile countermeasure experiment. By adjusting the parameter information such as equipment type, quantity, networked connection mode and deployment position, 100 groups of different weapon equipment systems were formed as \( A_i (i=1,\ldots, 100) \), and 6 evaluation indexes of the corresponding system were measured in the course of the experiment, which were recorded as \( Z_j (j=1,\ldots, 6) \), as shown in Figure 2. Through our method, we finally get a visual analysis diagram of 6 indexes of 100 sets of systems.

The black dots represent the spatial distribution of the 100 group systems, the red dots represent 6 indicator spaces, and the green area represents the index value of the system \( Z_3 \) greater than 8, the brown area within the region represents the index value of the system \( Z_6 \) is greater than 10. The results are shown in Figure 3.

![Figure 2. Fusing mapping graph.](image)

![Figure 3. Relative area graph.](image)
The following conclusions can be obtained through visual analysis: The correlation between Z1 and Z2 is strong, and the correlation between Z5 and Z6 is stronger.

According to the comprehensive judgment of six indicators, 100 groups of weapon systems can be divided into about three groups. Among them, the value of the system measurement indicator Z3 in the green area is greater than 8, and the value of the system measurement index Z6 in the Brown region is greater than 10. In the cross area, its measured values meet Z3 values greater than 8 and Z6 values greater than 10, indicating that the region's system indicators are strongly correlated by Z3 and Z6.

6. Conclusions
We have designed a set of visualization methods to display multi-dimensional indicators and systems in the same plane at the same time. Through the fusion mapping of this visualization method, users can quickly judge the correlation between indicators and the similarity between systems. On this basis, the region of correlation density is drawn. The degree of correlation between indicators and systems is displayed by visualization method. Users can quickly judge the correlation between indicators and systems. Current methods can satisfy the comprehensive analysis of multi-dimensional static data, but do not consider the time series expression of dynamic data [15]. Our future direction can also consider data visualisation of contributions to evidence-based decision-making [16]. Those are our research direction in the future. These studies could make important contributions to understanding the pathways to research uptake and the role that data visualisation can play in promoting greater application of evidence to the design and practice of weapon system.

References
[1] Zikopoulos P and Eaton C 2011 Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data New York: McGraw-Hill Osborne Media vol. 1
[2] Keim D, Qu H and Ma K L 2013 Big-Data visualization IEEE Computer Graphics and Applications 33 vol. (4)
[3] Yuan X R 2013 Big data visualization and visual analysis http://www.chinacloud.cn/upload/2012-12/13122814565
[4] Ren L 2009 Research on interaction techniques in information visualization[Ph. D Thesis]. Beijing: The Chinese Academy of Sciences (in Chinese with English abstract)
[5] Card S K, Mackinlay J D and Shneiderman B 1999 Readings in Information Visualization: Using Vision to Think San Francisco: Morgan-Kaufmann Publishers
[6] Ahlberg C and Shneiderman B 1994 Visual information seeking: Tight coupling of dynamic query filters with starfield displays In: Beth A, Susan D, Judith O, eds. Proc. Of the CHI New York: ACM Press pp 313-317
[7] Inselberg A and Dimsdale B 1990 Parallel coordinates: A tool for visualizing multi-dimensional geometry. In: Kaufman A, ed. Proc. Of the Visualization San Francisco: IEEE Press pp 361-378
[8] Elmqvist N, Dragicevic P and Fekete J D 2008 Rolling the dice: Multidimensional visual exploration using Scatterplot matrix navigation IEEE Trans. On Visualization and Computer Graphics 14(6) pp 1539-1548
[9] Stephan Diehl, Fabian Beck and Michael Burch 2010 Uncovering Strengths and Weaknesses of Radial Visualizations-an Empirical Approach IEEE Transactions on visualization and computer Graphics vol. 16 pp 985-941
[10] Tran Van Long Lars Linsen 2011 Visualizing high density clusters in multidimensional data using optimized star coordinates Comput Stat 26 pp 655-678
[11] I Jolliffe 2002 Pricipal Component Analysis Series: Springer Series in Statics 2nd ed., Springer, NT, XXIX, pp 28-32
[12] E Kendogan 2000 Star Coordinates: A Multi-Dimensional Visualization Techniques with Uniform Treatment O Dimensions Proc. IEEE Information Visualization, Late breaking Topics pp 9-12
[13] S Ingram, T Munzner and M Olano 2009 Glimmer: Multilevel MDS on the GPU IEEE Trans. Visualization and Computer Graphics vol. 15(2) pp 249-261

[14] P Kerm 2002 Adaptive Kernel Density Estimation The Stata Journal 2 pp 148-156

[15] Svetlana Mareva 2010 Integration and Visualization of Time Series Expression Data of Gene Regulatory Networks

[16] Imogen Robinson 2016 Data visualisation: Contributions to evidence-based decision-making A SciDev.Net Learning Report