

doi: 10.7690/bgzd.2023.08.008

## 基于模糊聚类的工程保障力量编组方法

王东, 陈虹, 徐勇, 任利平

(西南计算机有限责任公司技术中心, 重庆 400060)

**摘要:** 针对当前工程保障指挥中兵力编组主要依赖指挥员的经验和主观决策的问题, 提出基于模糊聚类的工程保障力量编组方法。分析工程保障力量与任务需求的关联特征, 通过构筑急造军路任务示例验证工程保障力量编组方法。结果表明: 该方法能够按照任务需求划分所属工程保障力量, 为快速制定工程保障行动方案提供支撑。

**关键词:** 工程保障; 模糊聚类分析; 兵力编组

**中图分类号:** TJ07 **文献标志码:** A

## Grouping Method of Engineering Support Force Based on Fuzzy Clustering

Wang Dong, Chen Hong, Xu Yong, Ren Liping

(Technology Center, Southwest Computer Co., Ltd., Chongqing 400060, China)

**Abstract:** In view of the problem that the current engineering support force grouping mainly depends on the experience and subjective decision-making of the commander, a method of engineering support force grouping based on fuzzy clustering is proposed. The correlation characteristics between engineering support force and mission requirements are analyzed, and the method of engineering support force grouping is verified by constructing an urgent military road mission example. The results show that the method can divide the engineering support forces according to the task requirements, and provide support for the rapid formulation of engineering support action plan.

**Keywords:** engineering support; fuzzy clustering analysis; force grouping

### 0 引言

随着战争形态加速由机械化向信息化转变, 工程保障重点由具体行动转到作战体系的关键节点, 工程保障范围由平面转向立体、由要素转向体系, 工程保障行动由定点、伴随保障向作战全程和战场全域深度拓展。传统工程保障力量编组模式较单一、平战转换不够灵活、过于依赖主观决策等问题日益凸显, 难以满足工程保障力量动态组合、专业按需集成的作战运用需求, 无法有效适应工程保障全维部署、全过程使用、一体化运用的力量建设目标<sup>[1]</sup>。笔者采用聚类分析的思想, 分析现役指挥信息系统中工程保障力量编组存在的问题、过程、方法等, 适应工程兵力量体系的“身形更轻巧、结构更充实、编配更灵活、行动更敏捷”建设要求<sup>[1]</sup>, 满足基于任务运用、专业匹配、效率归并等原则, 为指挥员迅速制定工程保障力量编组方案提供辅助支持, 使工程保障力量发挥专长、保障效益最大化, 有效提升工程保障效能。

### 1 工程保障力量编组现状

工程保障力量包括具体实施工程保障的机构、

人员和装备、设备、设施、器材等, 是遂行工程保障任务的基础。工程保障力量编组是为遂行工程保障任务, 按照一定要求将编制内的部队、分队或人员进行临时编配与组合<sup>[2]</sup>, 是工程保障指挥决策的重要环节。

在工程兵指挥信息系统多代建设过程中, 工程行动指挥决策模型也经过多次迭代, 但指挥决策模型对于工程保障力量编组和配置, 没有确定关于任务特点、作业方式、兵力需求等规则或策略要求, 主要依靠指挥员的经验和主观决策完成对所属工程保障力量编组, 缺乏对编组的科学性、合理性约束, 容易造成任务与执行作业力量的专业不匹配现象, 不能科学、有效地组织所属工程保障力量, 未将专业优势、装备优势转化为行动优势和保障优势。

针对工兵部(分)队、工程装备能力特点, 分析工程保障任务对力量编组的需求, 明确任务和作业装备/分队的关联约束, 结合历史演训数据、装备保障效能数据和专家经验数据等; 根据工程保障力量的专业特征, 构建工程作业分队/装备特征矩阵, 针对工程保障任务需求进行兵力模糊聚类分析, 优选力量编组方案, 为指挥员快速制定工程保障行动方

案提供支撑。

### 2 工程保障力量编组分析

工程保障力量编组围绕工程保障任务的实施，适应工程保障任务多元化的要求，满足“平时按建制/专业训练，战时按任务科学组合”的力量组织模式要求，有利于工程保障资源的整体使用和统筹分配，便于对执行作业分队/装备的指挥和管理<sup>[1]</sup>。从工程保障任务驱动工程保障力量编组，主要以各专业分队模块化组合为手段，具备现有编制力量优势运用和针对性训练的特点，能够确保快速反应和遂行任务能力，增强完成任务的灵活性和适应性，可根据任务需要和专业特长对工程保障力量作出临机调整，任务驱动力量编组过程如图 1 所示。

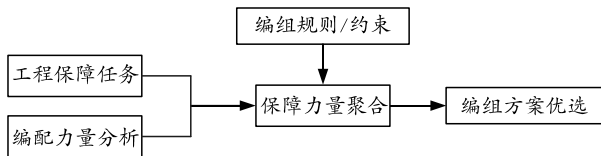


图 1 任务驱动力量编组过程

通常根据作战样式、遂行任务、配属位置、装备/分队能力以及任务间关系等进行工程保障力量编组。工程保障力量的编组要坚持“合理、适度、灵活”的原则<sup>[3]</sup>，根据任务的重要程度、时限要求等，发挥各专业分队/工程装备的技术特长，优化力量组合，编成适应不同任务需要，便于指挥、协同、群队结合的力量编组。

工程保障任务清单具备层级化的模块性、独立性和专业性特征，与工程保障力量的编配形式、作业内容、专业性质等方面相符，依据任务的特征分析，建立分队/装备与任务相对应的编组关系，完成工程保障力量编组是合理可行的，任务与力量编组关系如图 2 所示。

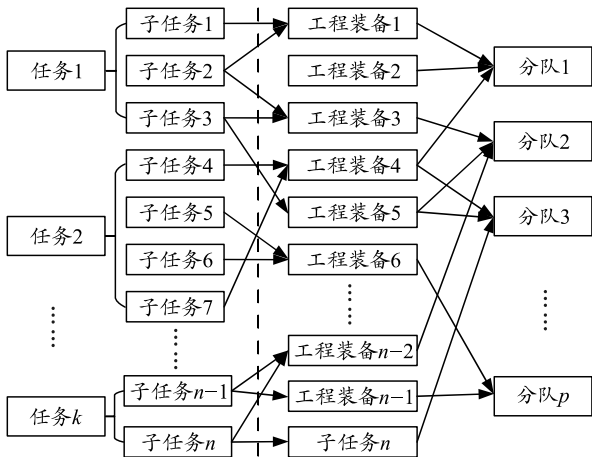


图 2 任务-力量编组关系

### 3 工程保障力量编组算法

工程保障力量编组是通过人员、装备等要素的科学组合，充分发挥工程保障力量的优势效能。如何快速、有效地将工程保障分队/装备聚合，形成各类工程保障群/队/组是各类作战行动战斗部署的重点。

#### 3.1 模糊聚类分析

模糊聚类分析是涉及事物之间的模糊界限时按一定要求对事物进行分类的数学方法，通常根据事物本身的属性来构造模糊矩阵，在此基础上根据一定的隶属度来确定聚类关系，即用模糊数学的方法定量地确定事物之间的模糊关系，从而客观且准确地进行分类。

系统聚类法是基于模糊等价关系的模糊聚类分析法。可用等价关系对数据集进行聚类，系统聚类方法的数学描述步骤如下：

1) 描述数据特征，设被聚类的数据集为  $X = \{x_1, x_2, \dots, x_n\}$ ，对于每个单元  $x_i (i=1, 2, \dots, n)$ ，均有  $m$  项特征，记作  $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ ， $x_{im}$  表示数据  $x_i$  的第  $m$  个特征的向量值。

2) 规定数据间相似系数  $R = \{r_{ij} | 0 \leq r_{ij} \leq 1; i, j = 1, 2, \dots, n\}$ ， $r_{ij}$  描述数据  $x_i$  与  $x_j$  之间的差异或相似程度。 $r_{ij}$  越接近于 1，表明  $x_i$  与  $x_j$  之间的差异越小； $r_{ij}$  越接近于 0，表明  $x_i$  与  $x_j$  之间的差异越大。 $r_{ii} = 1$  ( $x_i$  与自身没有差异)，( $x_i$  与  $x_j$  之间的差异等同于  $x_j$  与  $x_i$  之间的差异)，由  $r_{ij} (i, j = 1, 2, \dots, n)$  可得  $X$  的模糊相似关系。

3) 运用模糊运算求出最接近相似关系的模糊等价关系，生成模糊等价矩阵  $R^*$ 。

4) 根据选取适当水平  $\lambda (0 \leq \lambda \leq 1)$  对模糊等价矩阵进行分类，得到  $X$  的一种聚类。

#### 3.2 基于任务需求模糊聚类的力量编组算法

工程保障力量编组以任务需求的专业特性、作业能力需求、作业时间、时序要求等特征为基础，对编配的分队/装备进行模糊聚类分析，将相似度较高的作业分队/装备聚合到同一子集中，将具备相同专业特性的作业分队/装备按照作业力的需求集中编组，避免差异度较大的作业分队/装备在同一“群/队/组”，造成工程保障力量的浪费和闲置，充分发挥作业分队/装备的优长，实现集中优势力量完成重点任务的目的。工程保障力量编组流程如图 3 所示。

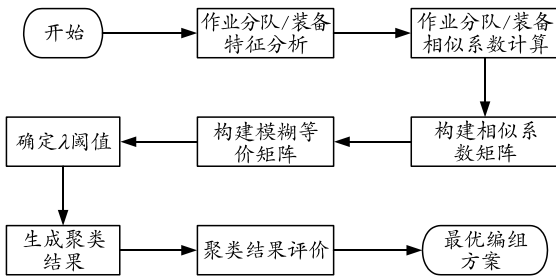


图 3 工程保障力量编组流程

首先,根据任务需求对作业分队/装备进行特征提取,进行数据标准化,建立数据矩阵;其次,根据作业分队/装备间的相似特性,采用相关系数法确定相似系数,形成作业分队/装备的模糊相似矩阵;然后,用传递闭包聚类法建立模糊等价关系,按不同的水平进行聚类;最后,对不同阈值聚类结果进行评价,选出最佳力量编组方案。

1) 原始数据矩阵进行标准化,建立特征向量矩阵。

作业分队/装备集合为  $X = \{x_1, x_2, \dots, x_n\}$ , 对于任务  $T$  每个作业分队/装备  $x_i$  均有  $m$  项特征(任务的专业特性、能力等),记作  $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ , 则原始数据矩阵为:

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}。$$

对原始数据矩阵采用式(1)进行标准差标准化,将  $x_{ij}$  转化  $x'_{ij}$  处理后数据范围为[0,1]。

$$x'_{ij} = (x_{ij} - \bar{x}_j) / S_j, (1 \leq j \leq m)。 \quad (1)$$

式中:  $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$ ;  $S_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$ 。

2) 计算相似系数,建立模糊相似矩阵。

采用式(2)相关系数法建立作业分队/装备集合对于任务  $T$  的相似关系:

$$r_{ij} = \frac{\sum_{k=1}^m |x_{ik} - \bar{x}_i| |x_{jk} - \bar{x}_j|}{\sqrt{\sum_{k=1}^m (x_{ik} - \bar{x}_i)^2} \sqrt{\sum_{k=1}^m (x_{jk} - \bar{x}_j)^2}}。 \quad (2)$$

得到作业分队/装备之间的模糊相似矩阵:

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \vdots & \vdots & & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{pmatrix}。$$

3) 构造模糊等价矩阵,选择阈值进行动态聚类。

模糊相似矩阵  $R$  具备自反性和对称性,但不具备传递性,不具有模糊等价关系,需要将模糊相似矩阵  $R$  转化为模糊等价矩阵  $R^*$ ,利用传递闭包法实现模糊相似关系转换成模糊等价关系。

从模糊相似矩阵  $R$  出发,依次求平方:

$$R \rightarrow R^2 \rightarrow R^4 \rightarrow \dots \rightarrow R^{2^k} = R^*。$$

式中:  $R^2 = R \circ R = \bigvee_{k=1}^n (r_{ik} \wedge r_{kj})$ , 其中:  $\vee$  为两数取大运算(逻辑加);  $\wedge$  为两数取小运算(逻辑乘)。当满足  $R^k \circ R^k = R^k$  时,即合成前矩阵包含合成后的矩阵,  $R^k$  就是所求的模糊等价矩阵:

$$R^* = [r^*_{ij}]_{n \times m}, 0 \leq r^*_{ij} \leq 1。$$

基于  $R^*$ ,对于不同的阈值  $\lambda \in [0,1]$  可得到不同的分类,通过对阈值控制可以实现不同形式的编组,按阈值  $\lambda$  截矩阵进行动态聚类,  $R^*$  中任意一个值  $r^*_{ij}$  均可作为阈值  $\lambda$ , 可得到:

$$R^*_\lambda = [\lambda_{ij}]_{n \times m}。$$

式中

$$\lambda_{ij} = \begin{cases} 0, & r^*_{ij} < \lambda \\ 1, & r^*_{ij} \geq \lambda \end{cases} \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m。$$

4) 评价聚类有效性阵,优选编组方案。

聚类分析属于无监督学习,内部有效性指标是最常用的聚类有效性指标,通常用来确定数据集的最佳聚类数<sup>[4]</sup>。基于数据集的几何结构信息采用 CH(calinski-harabasz) 指标对数据集的聚类划分结果进行评价。CH 指标通过计算类中各数据与中心的距离平方和来度量类内的紧密度,通过计算各类中心与数据集中心距离平方和来度量数据集的分离度,CH 指标是分离度与紧密度的比值。CH 越大表示类自身越紧密、类与类之间越分散。CH 指标定义为:

$$CH = \frac{tr(B)}{k-1} / \frac{tr(W)}{n-k}。 \quad (3)$$

式中:  $n$  表示聚类的数目;  $k$  表示当前的类;  $C_j$  表示数据集中心;  $\bar{x}$  表示类中心; 则紧密度  $tr(B) = \sum_{j=1}^k n_j \|C_j - \bar{x}\|^2$ ; 分离度  $tr(W) = \sum_{i=1}^m \|x_i - \bar{x}\|^2$ 。

将不同聚类阈值得到的系列聚类结果计算其有效性指标并比较,确定最佳指标值的聚类结果为最优编组方案。