

# 舰艇编队信息融合中模糊 C-均值算法改进研究<sup>\*</sup>

方 群<sup>1</sup> 邵 晓<sup>1</sup> 郭定荣<sup>2</sup>

(1. 海军蚌埠士官学校 蚌埠 233012)(2. 海军信息化部 北京 100036)

**摘 要** 在分析舰艇编队信息处理融合的基础上,提出舰艇编队信息融合体系结构、模型和实现框架。针对原 C-均值算法在处理数据时的不合理性,提出了改进的模糊 C-均值算法,提高分类效果,满足信息融合的实时要求和可靠性要求。

**关键词** 信息融合; C-均值算法

**中图分类号** TP301.6

## An Improved Adaptive Fuzzy C-means Algorithm for Information Fusion of Navy Armada

FANG Qun<sup>1</sup> SHAO Xiao<sup>1</sup> GUO Dingrong<sup>2</sup>

(1. Bengbu Navy Petty Officer Academy, Bengbu 233012)(2. The Navy Information Department, Beijing 100036)

**Abstract** On the base of the analysis of information transaction platform and work sequence of the integrated operation of intelligence transaction of naval information, its system framework, model, and actual frame was presented. According to the shortcomings on the C-means algorithms among information transaction, then the improved fuzzy C-means arithmetic was used in the information fusion of intelligence transaction of naval formation, which basically fulfills the need of real-time and reliability of the information fusion of intelligence transaction of naval formation.

**Key Words** information fusion, fuzzy C-means algorithm

**Class Number** TP301.6

### 1 引言

未来海战将是集电子战、C<sup>3</sup>I系统和精确制导武器的以体系对抗体系的海、陆、空、天、电磁的多维战争。毫无疑问,在现代以及未来海战中,舰艇编队的情报信息处理非常关键。情报信息的获取、传递、处理和对军事态势的评估是进行作战指挥的基本前提。因此,建立高效的舰艇编队情报信息融合系统,是所有国家军队在现代和未来海战中必须面对和解决的重大问题。

### 2 信息融合的体系结构

信息融合的层次结构主要从信息的角度来分析融合系统,信息融合的体系结构则主要是从硬件的角度来分析融合系统。

#### 1) 信息融合的层次结构

信息融合的层次结构是按照信息抽象程度来划分的。通常来说,根据信息抽象的三个层次,即数据层、特征层和决策层,可将信息融合分为三级,即数据层融合、特征层融合和决策层融合。

#### 2) 信息融合的体系结构

信息融合的硬件体系结构大致分为三类:集中式、分布式和混合式。

(1)集中式是将各传感器结点的数据都送到中央处理器进行融合处理。该方法可以实现实时融合,其数据处理的精度高、解法灵活,缺点是对处理器要求高、可靠性较低、数据量大,故难于实现。

(2)分布式是各传感器利用自己的量测单独跟踪目标,将估计结果送到总站,总站再将子站的估计合成为目标的联合估计。该方法对通信带宽要求低、计算速度快、可靠性和延续性好,但跟踪精度没有集中式高。

(3)混合式是将以上两种形式进行组合,它可以在速度、带宽、跟踪精度和可靠性等相互影响的各种制约因素之间取得平衡,是目前研究的主要形式。

### 3 舰艇编队情报处理的信息融合结构

对于一个舰艇编队,在空间上拥有分布在各个海域的舰艇、飞机、潜艇等,在时间上则有长期经营的各方面的信息记载以及将来的作战计划预算,在体制上则分成指挥机关、武器控制小组等,构成复杂的多级多职能的交叉机制。编队综合作战情报在海军综合电子信息系统的支持下完成编队空中、水面、水下警戒探测网情报的收集与管理;对收集的情报数据进行整理、分析、筛选、判定、融合,进行威胁判断,形成作战区域综合战场态势图。在舰艇编队情报处理的信息融合过程中,构成众多而且分散的多元数据源,它

<sup>\*</sup> 收稿日期:2012年2月9日,修回日期:2012年3月15日

作者简介:方群,女,硕士研究生,讲师,研究方向:指挥自动化。邵晓,男,讲师,研究方向:指挥自动化。郭定荣,男,工程师,研究方向:通信指挥。

们的结构复杂且语义多样。这些数据还具有层次性,反映决策过程中的不同级别。另外还有各级背景知识,如行为规范、法规、常识以及上级对下级的控制与限制等。利用所有这些信息进行情报处理分析和决策支持的过程将相当复杂。因此需要信息融合技术的全方位的集成,实现更高层次的智能分析,满足用户新情报的应用需要。

#### 4 模糊聚类 C-均值算法分析

在传统的模式识别中,特征选择和特征提取中关键问题就是如何评价得到的特征,使特征中更好地反映分类的信息。在编队情报信息融合中,模糊聚类 C-均值算法的流程如下:

定义: $N_i$ 是第  $i$  聚类  $\Gamma_i$  中的样本数目, $m_i$ 为样本平均值,即

$$m_i = \frac{1}{N_i} \sum_{y \in \Gamma_i} y \quad (1)$$

$J_e$ 是误差平方和的聚类准则,即

$$J_e = \sum_{i=1}^c \sum_{y \in \Gamma_i} \|y - m_i\|^2 \quad (2)$$

使  $J_e$ 最小的聚类是误差平方和准则下的最优结果。

步骤:

Step1:选择把  $N$  个样本分为  $C$  个聚类的初始划分,计算每个聚类的均值  $m_1, m_2, \dots, m_c$  和  $J_e$ 。

Step2:选择一个备选样本  $y$ ,设  $y$  现在在  $\Gamma_i$  中。

Step3:若  $N_i = 1$ ,则转 Step2,否则继续。

Step4:计算

$$\rho_j = \begin{cases} \frac{N_j}{N_j + 1} \|y - m_j\|^2 & j \neq i \\ \frac{N_j}{N_j - 1} \|y - m_j\|^2 & j = i \end{cases} \quad (3)$$

Step5:对于所有的  $j$ ,若  $\rho_k \leq \rho_j$ ,则把  $y$  从移到中去。

Step6:重新计算  $m_i$  和  $m_k$  的值,并修改  $J_e$ 。

Step7:若连续迭代  $N$  次  $J_e$  不变,则停止,否则转到 Step2。

存在问题:上述 C-均值算法是在类别数  $C$  给定的情况下进行的,并非所有的情况都能找到最优聚类数,在样本集不理想的情况下,可能导致结果不好,比如某个野值样本原理各类的聚类中心,本来它严格属于各类的隶属度都很小,但是这种野值的存在将影响迭代的最终结果,为了解决这种问题,本文提出了对原有算法的改进。

#### 5 改进的模糊聚类 C-均值算法分析

改进的模糊 C 均值算法,将问题的有关符号重新规定如下:

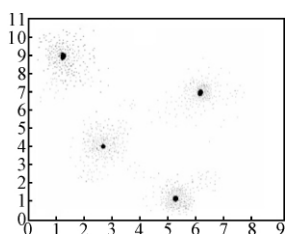


图1 正确的四个聚类中心的位置

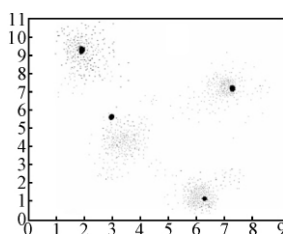


图2 模糊 C 均值算法

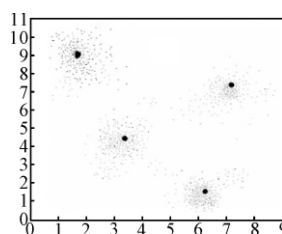


图3 改进的模糊 C 均值算法

#### 7 结语

通过分析对舰艇编队信息融合处理中原 C-均值算法在

$\{x_i, i=1, 2, \dots, n\}$  是  $n$  个样本组成的样本集合, $c$  为预定的类别数目;

$m_i, i=1, 2, \dots, c$  是每个聚类的中心;

$\mu_j(x_i)$  是第  $i$  个样本对于第  $j$  类的隶属度函数,用隶属度函数定义的聚类损失函数可以写为

$$J_f = \sum_{i=1}^c \sum_{y \in \Gamma_i} [\mu_j(x_i)]^b \|x_i - m_i\|^2 \quad (4)$$

其中, $b>1$  是一个可以控制聚类结果的模糊程度的常数,在不同隶属度定义方法下最小化该式就得到不同的模糊聚类方法,改进的模糊 C 均值法要求样本对于各个聚类的隶属度之和为  $n$ ,即:

$$\sum_{j=1}^c \sum_{i=1}^n \mu_j(x_i) = n \quad (5)$$

在这个条件下计算  $J_f$  的极小值,令  $J_f$  对  $m_i$  和  $\mu_j(x_i)$  的偏导数为 0,可得必要条件:

$$m_j = \frac{\sum_{i=1}^n [\mu_j(x_i)]^b x_i}{\sum_{i=1}^n [\mu_j(x_i)]^b}, j = 1, 2, \dots, c \quad (6)$$

$$\mu_j(x_i) = \frac{n(1/\|x_i - m_i\|^2)^{1/(b-1)}}{\sum_{k=1}^c \sum_{i=1}^n (1/\|x_i - m_k\|^2)^{1/(b-1)}} \quad (7)$$

$i = 1, 2, \dots, n; j = 1, 2, \dots, c$

具体算法步骤为:

Step1:设定聚类数目  $c$  和参数  $b$

Step2:初始化各个聚类中心  $m_i$

Step3:重复下面的运算直至各个样本隶属度值稳定:用当前的聚类中心根据式(7)计算隶属度函数;用当前的隶属度函数按式(6)更新计算各类聚类中心。

当算法收敛时,就得到了各类的聚类中心和各个样本对各类的隶属度值,从而完成了模糊聚类划分。显然,用改进的模糊 C 均值算法得到的隶属度值可能大于 1,因此并不是通常意义上的隶属度函数。必要时可以把最终得到的隶属度函数进行归一化处理,并不会影响到聚类结果,如果要求结果进行去模糊化。

#### 6 实验数据结果及比较

如图 1~3 所示,图 1 为正确的四个聚类中心的位置,把聚类数设置为三个,图 2 和图 3 分别为模糊 C 均值算法和改进的模糊 C 均值算法,可见改进的模糊 C 均值算法可以正确给出四个聚类中心中的三个,具有更好的鲁棒性,不但可以在有野值存在的情况下得到较好的聚类结果,而且放松了隶属度条件,使最终聚类结果对预先确定的聚类数目不敏感。

处理数据时的不合理,提出了新的改进的模糊 C-均值算法,旨在使该算法具备更好的鲁棒性,提高聚类效果,使其满足信息融合的实时要求和可靠性要求。 (下转第 92 页)

400、800B 为测试的传输数据包大小。

以控制与显示计算机节点到数据录取计算机节点传输延迟的性能测试为例,控制与显示计算机作为发送节点,数据录取计算机作为接收节点,双方均采用中断响应方式,使用 VC 提供的可精确到  $1\mu\text{s}$  的高精度计时器,连续测试 1000 次。当传输的数据包大小是 200B 时,传输延迟数据统计如图 4 所示。

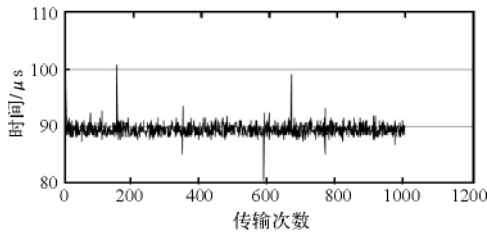


图 4 传输延迟数据统计图

当传输的数据包大小分别为 20、200、400、800B 时,传输延迟数据统计如表 1 所示。

表 1 传输延迟数据统计表

| 数据包大小/B | 最大值/ $\mu\text{s}$ | 最小值/ $\mu\text{s}$ | 平均值/ $\mu\text{s}$ |
|---------|--------------------|--------------------|--------------------|
| 20      | 40.53              | 22.67              | 28.94              |
| 200     | 161.51             | 89.28              | 100.55             |
| 400     | 838.95             | 352.01             | 545.74             |
| 800     | 1237.08            | 587.86             | 633.28             |

在数据传输正确性测试中,首先设置系统仿真步长,时统模块按设置的仿真步长发送仿真时间,控制和剧情解算模块、雷达电磁环境生成模块和接口适配模块根据仿真试验系统设计模拟相应的接收和发送数据,数据录取模块实时记录试验数据。通过对比各功能模块模拟发送的数据和数据录取模块实时记录的试验数据,验证传输数据的正确性。分别设置 1ms、5ms 和 50ms 的仿真步长进行测试,试验结果表明,系统数据传输正确。

## 7 结语

基于广播内存网的雷达侦察装备半实物仿真试验系统可以设置仿真步长,能将系统时间同步精度转换为 GPS 授时误差和广播内存网数据中断延迟,数据传输正确、确定,而且在广播内存网上扩展功能节点不影响系统整体实时性,具有良好的应用前景。

## 参考文献

- [1] 魏岳江. 复杂电磁环境下的联合训练[J]. 国防科技, 2008, 29(4): 62-67.
- [2] 高慧英, 曹宁生, 李三全. 雷达侦察装备半实物仿真试验系统设计[J]. 舰船电子工程, 2007(4): 106-108.
- [3] 廖瑛, 梁加红, 等. 实时仿真理论与支撑技术[M]. 长沙: 国防科技大学出版社, 2002.
- [4] 邓红德, 鲍鑫, 吴佳楠. 小型无人机飞控系统半实物仿真平台实时性研究[J]. 测控技术, 2012, 31(1): 121-122.
- [5] Trainor W L, G Warden G G. Shared-memory networking architectures-simplicity and elegance. AIAA, 89-3298-CPP 252-258.
- [6] 顾颖彦. 反射内存网实时通信技术的研究[J]. 计算机工程, 2002(7): 143-144.
- [7] 刘志国, 王仕成, 金光军. 基于实时网络的激光制导武器半实物仿真试验系统设计[J]. 计算机仿真, 2003(5): 20-21.
- [8] 王刚, 曹宁生. 一种实时仿真系统黑板式网络架构[J]. 舰船电子工程, 2007(4): 148-150.
- [9] 张永顺, 童宁宁, 赵国庆. 雷达电子战原理[M]. 北京: 国防工业出版社, 2006: 20-21, 60-79.

(上接第 51 页)

## 参考文献

- [1] Waltz E L. Computational Considerations for Fusion in Target Identification System. [J]. IEEE proc. of the NAECON, 1981(11): 492-497.
- [2] 薛兵. C3I 系统中数据融合理论与应用研究[D]. 石家庄: 军械工程学院, 1998.
- [3] Rawles, James W. Army IEW: New Structure For New Century [J]. Defence Electron, 1989(6): 69-80.
- [4] Hall David L. Mathematical techniques in multi-sensor data fusion[M]. New York: Artech House, 1992: 36-40.
- [5] 刘同明, 夏祖勋, 解洪成. 数据融合技术及应用[M]. 北京: 国防工业出版社, 1998: 77-80.
- [6] White F. A Model for Data Fusion[C]. SPIE Conference on Sensor Fusion. Orlando, FL, 1988: 20-23.
- [7] 王雷. 基于信息融合的虚拟现实理论研究[D]. 上海交大, 2002.
- [8] 夏学知, 涂葵, 王小龙, 等. 基于集群计算机的信息处理系统设计[J]. 计算机工程与应用, 2005(9): 32-34.
- [9] 何友, 陆大珍, 彭应宁. 多传感器数据融合算法综述[J]. 火力与指挥控制, 1995, 21(1): 12-21.
- [10] 李淑玉, 楼树美, 孔李军. 基于强跟踪滤波器的多传感器信息融合应用研究[J]. 计算机与数字工程, 2009, 37(8).
- [11] 郭敏. 多源信息融合在故障诊断技术中的应用研究[J]. 计算机与数字工程, 2011, 39(3).
- [12] 董志荣. 再论信息融合[J]. 情报指挥控制系统与仿真技术, 2002(11): 30-34.