

存在证据冲突的临近空间飞行器遥测数据 相关性评价与融合方法

张少杰^{1,2,3}, 荣海军^{1,2,3}, 杨朝旭^{1,2,3}, 侯凯振^{1,2,3}

(1. 西安交通大学机械结构强度与振动国家重点实验室, 710049, 西安;

2. 西安交通大学陕西省先进飞行器服役环境与控制重点实验室, 710049, 西安;

3. 西安交通大学航天航空学院, 710049, 西安)

摘要: 为了解决单一的相关性分析方法在分析临近空间飞行器遥测数据时存在局限性以及证据冲突问题,在相关系数分析评价与优势组合的基础上,提出了基于支持因子的证据理论融合算法。首先分别利用 Pearson 相关系数、Spearman 相关系数与距离相关系数对遥测数据进行相关性评价分析,表明 3 种相关系数可优势互补。其次,建立了基于支持因子的证据理论融合算法,实现证据的冲突基本概率赋值函数分配,避免 Dempster-Shafer 证据理论(D-S 证据理论)的一票否决和合成规则失效问题。最后,利用 3 种相关系数构造相关性证据并开展遥测数据相关性分析实验。结果表明:基于支持因子的证据理论融合方法能使 Pearson 相关系数、Spearman 相关系数与距离相关系数的证据融合更加合理;在证据冲突大的情况下,相比相容系数的证据理论融合方法,相关概率计算结果准确度提高约 6.55%,能更有效地处理证据冲突问题。

关键词: 遥测数据;相关系数;证据冲突;支持因子;D-S 证据理论

中图分类号: TK311.5 **文献标志码:** A

DOI: 10.7652/xjtuxb202311010 **文章编号:** 0253-987X(2023)11-0100-10

Correlation Evaluation and Fusion Method for Telemetry Data of Near Space Vehicles with Conflicting Evidence

ZHANG Shaojie^{1,2,3}, RONG Haijun^{1,2,3}, YANG Zhaoxu^{1,2,3}, HOU Kaizhen^{1,2,3}

(1. State Key Laboratory for Strength and Vibration of Mechanical Structures, Xi'an Jiaotong University, Xi'an 710049, China;

2. Shaanxi Key Laboratory of Environment and Control for Flight Vehicle, Xi'an Jiaotong University, Xi'an 710049, China;

3. School of Aerospace Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: This paper proposes a support factor-based evidence theory fusion algorithm to address the limitations and evidence conflicts of a single correlation analysis method in analyzing telemetry data from nearby space vehicles. Firstly, the Pearson correlation coefficient, Spearman correlation coefficient, and distance correlation coefficient are used to evaluate and analyze the correlation of telemetry data. The results show that the three correlation coefficients complement each other's advantages. Secondly, an evidence theory fusion algorithm based on support factors is established to achieve the assignment of basic probability assignment functions for evidence conflicts, thus avoiding the one-vote veto and composite rule invalidation issues of Dempster-Shafer

收稿日期: 2022-12-23。 作者简介: 张少杰(1999—),男,硕士生;杨朝旭(通信作者),男,副教授,硕士生导师。 基金项目: 国家自然科学基金资助项目(61976172,12002254);陕西省重点研发计划资助项目(2023-YBGY-235)。

网络出版时间: 2023-05-19

网络出版地址: <https://kns.cnki.net/kcms/detail/61.1069.T.20230518.1721.010.html>

evidence theory (D-S evidence theory). Finally, the Pearson correlation coefficient, Spearman correlation coefficient, and distance correlation coefficient are used to make evidence and an experiment is conducted to analyze the correlation of telemetry data. The results show that the proposed algorithm can ensure the reasonability of the fusion from the Pearson coefficient, Spearman coefficient, and distance coefficient. In case of high evidence conflicts, the accuracy of correlation probability calculation is improved by about 6.55% compared to the evidence theory fusion method based on compatibility ratio. These findings indicate that the proposed algorithm is more effective in handling evidence conflict problems.

Keywords: telemetry data; correlation coefficient; conflicting evidence; support factor; D-S evidence theory

临近空间的物质构成、能量输运以及相互作用极其复杂,对飞行器的长期驻空飞行带来了挑战。利用遥测数据实时监测飞行器飞行状态,是确保高可靠长航时稳定飞行的有效手段^[1]。飞行器的能源、姿控、载荷等分系统之间接口交互,决定了遥测数据之间存在相关性。全参数的数据分析不仅导致信息冗余,增加了飞行器状态监测的工作量,而且难以检测出遥测数据之间相关关系的多元异常,使异常检测存在盲目性^[2]。遥测数据之间关联复杂,任何一种飞行参数可能与多种参数之间存在相关关系,并且相关关系随时间而变化。人工通过物理知识获取相关性模型耗时且费力,并且一旦飞行器结构和环境发生变化,通常要重新分析相关性。借助相关性分析算法能实时分析遥测数据相关性,获取数据的潜在关系,便于监测飞行器状态^[3]。欧洲太空局正是利用遥测数据之间的统计相关性进行异常检测,挖掘出与异常事件相关的遥测参数^[4]。因此,分析飞行器遥测数据的相关性有助于监测飞行器在轨运行状态,保障飞行器正常运行。

围绕遥测数据相关性,国内外学者提出了多种分析方法,包括 Pearson 相关系数^[5]、互信息^[6]、最大信息系数(MIC)^[7]和距离相关系数^[8]等方法。Zhe 等^[9]利用 Pearson 相关系数,分析了飞行器遥测数据相关性,能够快速检测出与舱门泄漏率相关的参数。Pearson 系数计算复杂度低,对随机噪声的鲁棒性强,但 Pearson 系数由于正态分布假设导致计算结果易受异常数据的影响。互信息不要求变量分布,但是难以计算概率密度函数,在互信息基础上学者提出了 MIC 方法^[10]。崔树银等利用 MIC 对多元负荷数据进行相关性分析,筛选出相关变量。MIC 能刻画数据间非线性关系,但是受网格划分数和方式的影响,计算时间复杂度高^[11]。孙宇豪等采用距离相关系数对遥测数据进行自变量选择,提高了模型的预测性

能^[12]。距离相关系数通过特征函数距离衡量变量之间的相关关系,相比 MIC 时间复杂度低,但计算结果易受噪声数据的影响。此外,上述方法在遥测数据相关性分析的实践中存在结果冲突的问题。

由于现有的单一相关性方法存在一定的局限性,如果将不同优势的相关性分析方法融合,则能够提高分析结果的可靠性,因此考虑将优势互补的相关性分析方法融合。D-S 证据理论具有坚实的理论基础,能在先验概率和条件概率都未知的情况下处理信息,提高判决结论的正确度和可信度^[13]。考虑到 D-S 证据理论具有上述优势,本文采用 D-S 证据理论进行融合。但是当存在高度冲突证据时,D-S 证据理论融合结果往往与常理相悖。Zhao 等将证据冲突部分作为未知信息概率以处理证据冲突,但该方法会导致未知信息概率增加,无法获得有效融合结果^[14]。孙全等使用证据的可信度处理证据之间的冲突部分,但由于可信度直接定义,导致结果存在主观性^[15]。Shang 等计算出 n 个证据的平均值,并使用 D-S 证据理论组合规则融合平均值^[16]。该方法能够处理冲突,然而未考虑到证据之间的权重。利用证据权重修正证据源,能够降低冲突证据的影响。赵静等利用 Jousselme 距离计算证据权重,并对证据加权平均后再采用 D-S 证据理论组合规则融合^[17]。该方法通过 Jousselme 距离考虑了证据间的关联性,利用焦元的交集与并集比值反映证据之间的相似性,但是证据焦元彼此独立,焦元的比值并不能充分利用各证据之间的信息。Lin 等采用相容系数计算证据之间的权重并修正证据,考虑了证据之间的相互支持度,但是无法有效处理高度冲突证据下的一票否决和合成规则失效问题^[18]。

为了解决临近空间飞行器的状态监测与相关性分析问题,考虑到现有单一的相关性分析方法的局限性,以及相关性分析实践中的证据冲突问题,本文

在相关系数分析评价与优势组合的基础上提出了基于支持因子的证据理论融合算法。首先,选取优势互补的 Pearson 相关系数、Spearman 相关系数与距离相关系数,分析遥测数据的相关性,并根据分析结果,构建相关性证据的初始基本概率赋值函数。其次,利用支持因子计算出各证据之间的权重并对证据进行修正,使修正后的证据具有相同的重要程度,计算证据命题的分布权重,分配修正证据的冲突基本概率赋值函数,从而获得融合结果。最后,开展临近空间飞行器遥测数据相关性分析实验,结果表明所提方法提高了遥测数据相关性分析的可信度。

1 遥测数据相关性分析

存在证据冲突的遥测数据相关性分析方法包括相关性评价和证据融合两部分内容,如图 1 所示。考虑到现有单一的相关性分析方法的局限性,选取能够优势互补的 Pearson 相关系数、Spearman 相关系数与距离相关系数,对临近空间飞行器遥测数据进行相关性计算,根据计算结果构造各相关性证据的初始基本概率赋值函数。考虑到相关性分析实践中存在冲突问题,在相关系数分析评价与优势组合的基础上,提出了基于支持因子的证据理论融合算法。利用支持因子计算出各初始证据之间的支持度,获得证据的权值并对初始证据进行修正,使修正后的证据具有相同的重要程度;根据证据支持度计算证据可信度,并获得证据命题的分布权重;最后分配修正证据的冲突基本概率赋值函数,获得相关性证据融合结果。

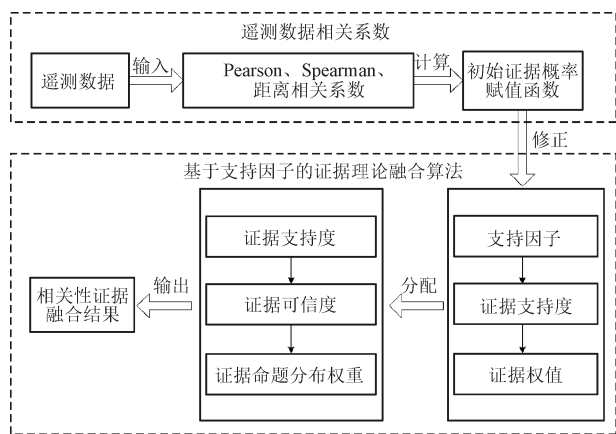


图 1 遥测数据相关性分析方法框图

Fig. 1 Correlation analysis method for telemetry data

2 相关性评价

本节考虑到临近空间飞行器遥测数据之间关联

复杂、数据量大、伴有随机噪声和扰动等特征,因此从随机噪声鲁棒性、异常数据抗干扰和数据非单调性角度,分别选取了 Pearson 相关系数、Spearman 相关系数与距离相关系数进行分析对比。

2.1 相关系数选取

2.1.1 Pearson 相关系数

Pearson 相关系数 $\rho_{x,y}$ 用于衡量数据间线性相关程度^[19],表示为

$$\rho_{x,y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

式中: (X_i, Y_i) 是任意两个集合 X 与 Y 的第 i 个值; $(\bar{X}, \bar{Y})$ 为两个集合的均值; n 为集合元素个数。 $\rho_{x,y}$ 取值在 $[-1, 1]$ 之间,能够体现出遥测数据的相关关系强度和方向。 $\rho_{x,y} = -1$ 表示 Y 随着 X 的增大而减小, $\rho_{x,y} = 1$ 表示 Y 随着 X 的增大而增大。 $\rho_{x,y}$ 计算复杂度低,在遥测数据的相关性分析中获得了普遍应用。

2.1.2 Spearman 相关系数

Spearman 相关系数 r_s 本质在于根据原始数据的排序位置进行相关性分析^[20]。对两个集合 (X, Y) 进行升序或降序排序,并获得排行集合 X', Y' , 元素 x_i, y_i 分别为 X_i 在 X 中的秩以及 Y_i 在 Y 中的秩。将个数为 n 的集合 X', Y' 中的元素 x_i, y_i ($1 \leq i \leq n$) 对应相减,获得一个排行差分值 $d_i (= x_i - y_i)$, r_s 表示为

$$r_s = 1 - 6 \sum_{i=1}^n d_i^2 / n(n^2 - 1) \quad (2)$$

r_s 采用单调变换描述集合之间的相关性,不会因个别异常数据影响整体排序,故抗干扰能力强。

2.1.3 距离相关系数

距离相关系数通过 X 与 Y 间联合特征函数与各自边际特征函数的差量化相关程度^[21]。距离相关系数 $R(X, Y)$ 表示为

$$R^2(X, Y) = \begin{cases} \frac{V_n^2(X, Y)}{\sqrt{V_n^2(X) V_n^2(Y)}}, & V_n^2(X) V_n^2(Y) > 0 \\ 0, & V_n^2(X) V_n^2(Y) = 0 \end{cases} \quad (3)$$

式中: $V_n^2(X, Y)$ 为样本距离协方差,表示为^[8]

$$V_n^2(X, Y) = \frac{1}{n^2} \sum_{i,j=1}^n A_{i,j} B_{i,j} \quad (4)$$

$$A_{i,j} = \|x_i - x_j\|_2 - \frac{1}{n} \sum_{k=1}^n \|x_k - x_j\|_2 -$$

$$\frac{1}{n} \sum_{l=1}^n \|x_i - x_l\|_2 + \frac{1}{n^2} \sum_{k,l=1}^n \|x_k - x_l\|_2 \quad (5)$$
$$B_{i,j} = \|y_i - y_j\|_2 - \frac{1}{n} \sum_{k=1}^n \|y_k - y_j\|_2 - \frac{1}{n} \sum_{l=1}^n \|y_i - y_l\|_2 + \frac{1}{n^2} \sum_{k,l=1}^n \|y_k - y_l\|_2 \quad (6)$$

式中: $k, l=1, 2, \cdots, n$ 。同理, 获得遥测数据距离方差 $V_n^2(X)$ 和 $V_n^2(Y)$, 表达式为

$$V_n^2(X) = \frac{1}{n^2} \sum_{i,j=1}^n A_{i,j}^2 \quad (7)$$

$$V_n^2(Y) = \frac{1}{n^2} \sum_{i,j=1}^n B_{i,j}^2 \quad (8)$$

当且仅当 $V_n^2(X, Y)=0$ 时, 表明 X 与 Y 独立。距离相关系数与 $\rho_{x,y}, r_s$ 相比, 通过特征函数距离能够更全面描述遥测数据的非单调关系。

2.2 相关性评价实验

本节开展临近空间飞行器的遥测数据相关性评价实验。受传输环境影响, 遥测数据下发时不可避免出现噪声, 以连续分布的高斯噪声和不同时刻下的脉冲噪声为例分析噪声鲁棒性。同时, 遥测数据之间关联复杂, 本节也从非单调性角度进行评价。

2.2.1 高斯噪声鲁棒性评价

数据噪声模型可表示为

$$y(i) = x(i) + \alpha n(i) \quad (9)$$

式中: $x(i)$ 是不加任何噪声的遥测数据; $y(i)$ 是加入噪声后的混合数据; α 为噪声强度系数, $\alpha \in [0, 1]$, 通过调节 α 去控制信噪比。随着 α 增大, 信噪比减小, 数据的噪声程度也随之增大。图 2 表示不含高斯噪声的原始遥测数据和含高斯噪声的遥测数据(信噪比为 40 dB)。

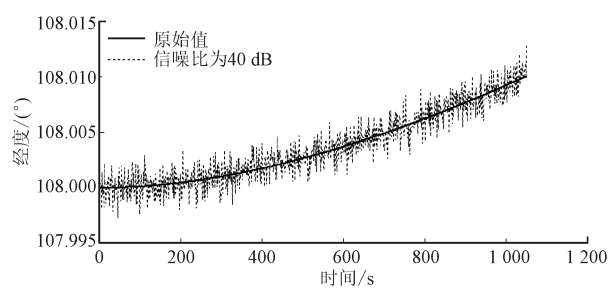


图 2 含或不含高斯噪声的遥测数据示意图
Fig. 2 Telemetry data with Gaussian noise

随着信噪比降低, 数据之间的关联性减小, 意味着相关系数的值随着噪声程度的增加而下降。因此, 增加噪声程度, 并比较相关系数的下降程度, 若下降程度最小, 则说明噪声鲁棒性最强。图 3 反映

了不同高斯噪声下的相关系数, 可见随着噪声程度的增加, 相关系数在逐渐下降, 但是下降程度却不同。图 3 中 Pearson 相关系数下降程度最低, 表明其高斯噪声鲁棒性最强。因为 Pearson 相关系数适合描述随机分布的数据相关性, 虽然加入了高斯噪声, 但是数据整体依然服从随机分布, 所以 Pearson 相关系数相比其他两种方法受高斯噪声影响更小。

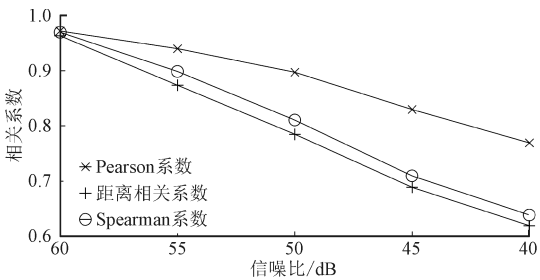


图 3 不同高斯噪声下的相关系数
Fig. 3 Correlation coefficients under different Gaussian noises

本节通过下降率衡量各相关系数的高斯噪声鲁棒性, 下降率为不加噪声相关系数和加入噪声后相关系数之间的差值与不加噪声的相关系数的比值, 计算结果如表 1 所示。Pearson、Spearman 以及距离相关系数在不同高斯噪声下的下降率之和分别为 0.59, 0.939 1 和 1.050 0。因为下降率越小, 噪声鲁棒性越强。因此, 上述 3 种相关系数的高斯噪声鲁棒性排序为 Pearson 相关系数最高, Spearman 相关系数次之, 距离相关系数最低。

表 1 不同高斯噪声下的相关系数下降率

系数	相关系数下降率				
	60 dB	55 dB	50 dB	45 dB	40 dB
Pearson 相关系数	0.013 5	0.055 4	0.108 3	0.176 1	0.236 7
Spearman 相关系数	0.029 6	0.095 6	0.177 8	0.283 1	0.353 0
距离相关系数	0.047 5	0.117 8	0.208 4	0.304 5	0.371 8

2.2.2 异常数据抗干扰评价

脉冲噪声的鲁棒性体现了对异常数据的抗干扰能力。图 4 是加入不同时刻脉冲噪声下的经度遥测数据。图 5 反映了不同脉冲噪声下的相关系数, 图中 Spearman 相关系数下降程度最低, 表明其抗干扰能力最强。因为 Spearman 相关系数通过数据的秩进行相关性分析, 个别扰动数据并不会改变整体数据的秩, 所以 Spearman 相关系数相比其他系数抗干扰能力更强。

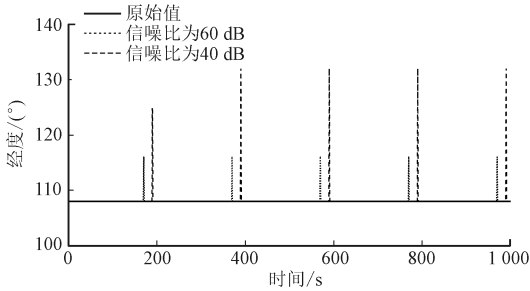


图 4 含脉冲噪声的遥测数据示意图

Fig. 4 Telemetry data with impulse noise

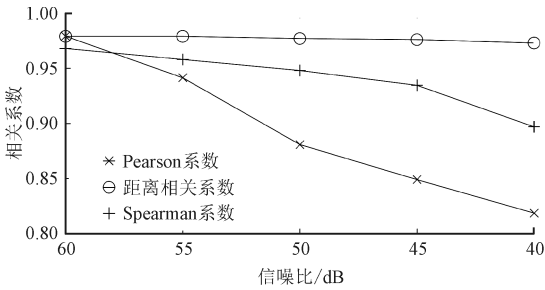


图 5 不同脉冲噪声下的相关系数

Fig. 5 Correlation coefficients under different impulse noises

本节依然通过下降率对比各相关系数的异常数据抗干扰能力,计算结果如表 2 所示。Pearson、Spearman 以及距离相关系数在不同脉冲噪声下的下降率之和分别为 0.488、0.042 5 和 0.220 6。因此,上述 3 种相关系数的异常数据抗干扰能力排序为 Spearman 相关系数最高,距离相关系数次之,Pearson 相关系数最低。

表 2 不同脉冲噪声下的相关系数下降率

Table 2 Decreasing rate of correlation coefficients under different impulse noises

系数	相关系数下降率				
	60 dB	55 dB	50 dB	45 dB	40 dB
Pearson 相关系数	0.018 1	0.055 9	0.106 6	0.138 5	0.168 9
Spearman 相关系数	0.005 3	0.006 2	0.007 6	0.010 7	0.012 7
距离相关系数	0.019 1	0.029 2	0.039 3	0.052 6	0.080 4

2.2.3 复合噪声鲁棒性评价

向原始遥测数据中加入含脉冲噪声和高斯噪声的复合噪声,比较各相关系数的复合噪声鲁棒性。图 6 为不同信噪比下的含复合噪声的遥测数据。图 7 反映了不同复合噪声下的相关系数,图中 Spearman 相关系数下降程度最低,表明其对复合噪声的鲁棒性最强。Pearson 相关系数受到其中脉冲噪声影响,在

三者中的鲁棒性最差,距离相关系数位居中间。

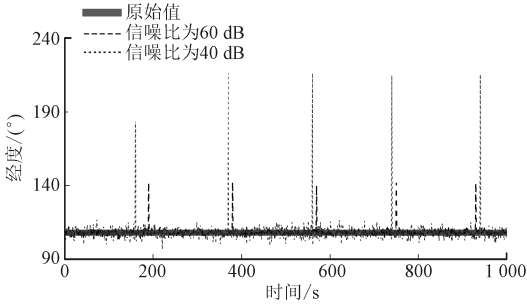


图 6 含复合噪声的遥测数据示意图

Fig. 6 Telemetry data with composite noises

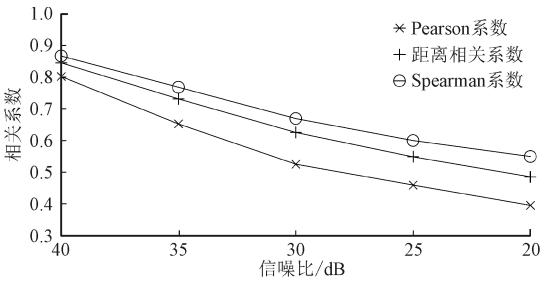


图 7 不同复合噪声下的相关系数

Fig. 7 Correlation coefficients under different composite noises

表 3 反映了相关系数对复合噪声的鲁棒性。Pearson、Spearman 以及距离相关系数在不同复合噪声下的下降率之和分别为 2.141 8、1.544 8 和 1.755。因此,上述 3 种相关系数的复合噪声鲁棒性排序为 Spearman 相关系数最高,距离相关系数次之,Pearson 相关系数最低。

表 3 不同复合噪声下的相关系数下降率

Table 3 Decreasing rate of correlation coefficients under different composite noises

系数	相关系数下降率				
	40 dB	35 dB	30 dB	25 dB	20 dB
Pearson 相关系数	0.191 1	0.342 2	0.470 3	0.536 8	0.601 4
Spearman 相关系数	0.132 8	0.231 5	0.330 4	0.399 8	0.450 3
距离相关系数	0.152 3	0.266 6	0.372 8	0.450 1	0.513 2

2.2.4 非单调性评价

遥测数据之间关联复杂,多种复杂关系并存。数据之间也存在非单调关系。本节通过距离相关系数,与 Spearman 相关系数描述遥测数据之间的非单调关系。图 8 反映了平台北速(飞行器平台沿正北方向的运动速度)与平台俯仰角之间的非单调关系。分别用距离相关系数和 Spearman 相关系数计

算两者相关性,距离相关系数的计算结果是 0.975 3, Spearman 相关系数的计算结果是 0.180 4。由于 Spearman 相关系数需要通过数据的单调变化获得数据的秩,所以当数据非单调变化时,会影响 Spearman 相关系数的分析结果,而距离相关系数不需要对数据进行单调变化,因此距离相关系数相比 Spearman 相关系数更能够准确描述数据之间的非单调关系。

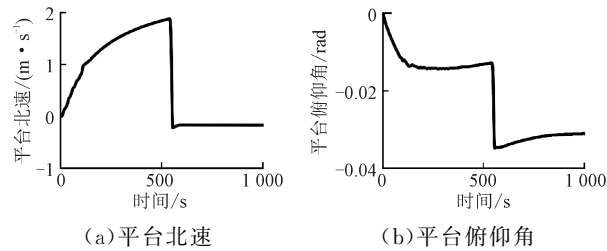


图 8 平台北速与平台俯仰角变化过程示意图

Fig. 8 Variation process of platform north speed versus platform pitch angle

评价分析结果表明:Pearson 相关系数计算复杂度低,对随机噪声的鲁棒性强,但异常数据抗干扰能力差;Spearman 相关系数对异常数据抗干扰能力强,但不能准确识别数据之间的非单调关系;距离相关系数能够更准确地描述数据之间非单调关系,但是对随机分布的高斯噪声鲁棒性差。因此,Pearson 相关系数、Spearman 相关系数与距离相关系数相互互补,三者融合结果能够更全面体现遥测数据相关性。

2.3 存在证据冲突的相关性分析

Pearson 相关系数、Spearman 相关系数与距离相关系数相互互补,本节分别采用上述 3 种相关系数作为相关性证据对遥测数据进行相关性分析。表 4 是 Pearson、Spearman 和距离相关系数的遥测参数相关性分析结果。分析表 4 可知,3 种相关系数在分析悬浮高度与大气压、悬浮高度与表面温度时,分析结果都趋于 1,表明悬浮高度与大气压以及表面温度相关。在飞行器运动速度与氦气囊内的氦气温度相关性分析结果中,Pearson、Spearman 相关系数结果表明二者不相关,距离相关系数的分析结果与 Pearson、Spearman 相关系数分析结果差异不大,但认为二者存在一定相关性。在空气温度与氦气体积的相关性分析结果中,Pearson 相关系数结果表明二者不相关,而 Spearman 和距离相关系数结果表明二者强相关,此时 3 种相关系数的分析结果不一致,产生了冲突。上述方法在遥测数据相关性分析实践中存在结果冲突的问题,需要融合算法有效

抑制证据冲突对分析结果的影响。

表 4 遥测数据相关性结果

Table 4 Correlation results of telemetry data

遥测参数	Pearson 相关系数	Spearman 相关系数	距离相关 系数
悬浮高度与大气压	0.999 538	1.000 000	0.999 705
悬浮高度与表面温度	0.999 458	0.999 824	0.999 663
速度与氦气温度	0.340 00	0.310 000	0.670 000
空气温度与氦气体积	0.000 02	0.964 692	0.958 082

3 证据理论融合算法

在相关系数分析评价与优势组合的基础上,提出了基于支持因子的证据理论融合算法。首先,由于 Pearson、Spearman 和距离相关性证据的分析角度不同,各相关性证据重要程度不一定相同。所以,采用支持因子计算出各证据之间的权重并对证据进行修正,使修正后的证据重要程度相同。其次,为了解决 D-S 证据理论存在的一票否决和合成规则失效问题,通过计算证据命题的分布权重,分配修正证据冲突基本概率赋值函数,从而获得融合结果。

辨识框架 $\Theta = \{A_1, A_2\}$,命题 A_1 表示相关, A_2 表示不相关。对于任意两个证据 m_i 与 $m_j (1 \leq i, j \leq 3)$ 之间支持因子 S_{ij} 为

$$S_{ij} = \frac{3 \sum_{t=1}^2 m_i(A_t) m_j(A_t)}{\sum_{t=1}^2 (m_i(A_t)^2 + m_j(A_t)^2 + m_i(A_t) m_j(A_t))} \quad (10)$$

式中: $A_t \subseteq \Theta, t = 1$ 或 2,代表上述相关性证据的结果,即相关和不相关。支持因子既能够表征证据之间的支持程度和一致性,也从侧面反映了证据之间的冲突程度。式(10)所计算的支持因子保留了各证据之间的初始信息,同时综合考虑到证据的所有命题 A_t ,能直接体现出证据整体之间的相互支持程度。 S_{ij} 取值范围为 $[0, 1]$, S_{ij} 若接近 0,表明证据之间支持程度低,冲突性强; S_{ij} 若接近 1,表明证据之间支持程度高,冲突性弱。

在证据理论中, Θ 包含所有命题构成的集合 2^Θ 。若存在函数 m 满足 $2^\Theta \rightarrow [0, 1]$, 且 $m(\emptyset) = 0$ 和 $\sum_{A \subseteq \Theta} m(A) = 1$, 则称 m 为 Θ 上的基本概率赋值函数(BPA)^[22-23]。Pearson、Spearman 和距离相关系数是支持遥测数据之间彼此相关的概率统计量,计算结果的绝对值属于 $[0, 1]$ 。所以,上述 3 种相关

系数结果能够直接构造各证据 BPA。Pearson、Spearman 和距离相关性证据的 BPA 分别为 m_1 、 m_2 、 m_3 。每种 BPA 都包含 2 个命题,例如 Pearson 相关性证据的 BPA $m_1: |m_{pe}|, (1 - |m_{pe}|)$, $|m_{pe}|$ 为 Pearson 相关性证据支持数据相关命题的概率, $(1 - |m_{pe}|)$ 为 Pearson 相关性证据支持数据不相关命题的概率。在获得 3 种相关性证据的 BPA 后,通过以下步骤进行融合:

步骤 1 通过式(10)计算证据支持因子 S_{ij} ,并依据 S_{ij} 计算证据的支持度 $p(m_i)$ 。在计算出相关性证据 m_i 与 m_j 之间的支持因子 S_{ij} 后,构造出一个 3×3 的支持因子矩阵 S 如下

$$S = \begin{bmatrix} 1 & S_{12} & S_{13} \\ S_{21} & 1 & S_{23} \\ S_{31} & S_{32} & 1 \end{bmatrix} \tag{11}$$

矩阵 S 中的元素体现出相关性证据之间相互支持的程度,如 S_{12} 表示 Pearson 相关性证据与 Spearman 相关性证据之间相互支持的程度。当一条证据被其他证据支持的程度越高,说明该证据的可信度越高;反之,可信度越低。系统中其他 s 相关性证据对证据 m_i 的支持度为 $p(m_i) = \sum_{j=1, j \neq i}^3 S_{ij}$ 。

步骤 2 计算相关性证据权值 w_i 并对相关性证据修正,使修正后的证据重要程度相同。支持度最大即 $\max_{1 \leq i \leq 3} p(m_i)$ 对应的证据为关键证据,各相关性证据相对关键证据的权重为 $w_i = \frac{p(m_i)}{\max_{1 \leq i \leq 3} p(m_i)}$ 。利用证据权重修正证据^[24]

$$m'_i(A) = w_i m_i(A), A \neq \Theta \tag{12}$$

$$m'_i(\Theta) = m_i(\Theta) + 1 - w_i \tag{13}$$

式(12)、(13)获得了修正后的基本概率赋值函数 $m'_i(1 \leq i \leq 3)$ 。式(12)使证据权重小的相关性证据对命题 A 所提供的确定性信息减小;式(13)使证据权重小的相关性证据所提供的不确定性信息增加。因此,减小了证据权重小的证据对整个融合系统的影响。经过式(12)、(13)修正后,各相关性证据的重要程度相同。

步骤 3 计算证据的命题分布权重并改进组合规则。通过命题分布权重分配修正证据的冲突基本概率赋值函数,获得融合结果^[25],表示为

$$m(A) = \sum_{\cap A_i = A} \prod_{1 \leq j \leq 3} m'_j(A_i) + K \sum_{i=1}^3 C_i m'_i(A) \tag{14}$$

式中: $K = \sum_{\cap A_i = \emptyset} \prod_{1 \leq j \leq 3} m'_j(A_i)$ 为修正后的相关性证据的总冲突值; $m(A)$ 为相关性证据的融合结果, $A \subseteq \Theta$; $m'_j(A_i)$ 为修正后的第 j 条相关性证据对命题 A_i 的概率; C_i 为第 i 条证据的可信度, $C_i = p(m_i) / \sum_{j=1}^3 p(m_j)$ 。根据证据可信度 C_i ,计算每种相关性证据对命题 A 的分布权重 $\alpha(A, m'_i) = \sum_{i=1}^3 C_i m'_i(A)$,它决定了每个命题所占的权重且满足 $\sum_{A \subseteq \Theta} \alpha(A, m') = 1$ 。

4 遥测数据分析结果

本节基于表 4 的相关性结果,构造相关性证据并开展证据融合,从证据冲突大和证据冲突小两种情况分别进行相关性分析。

4.1 证据冲突小情况下的相关性分析结果

飞行器悬浮高度与其表面温度的相关性证据及融合结果如表 5 所示。上述基本概率赋值函数相关概率都接近 1,表明悬浮高度与表面温度相关,证据分析结果一致。图 9 是悬浮高度与表面温度的变化过程曲线,表面温度随悬浮高度的增加而增加。图 9 与表 5 都充分体现出二者的相关性。飞行器在上升过程中表面温度升高,两者整体上呈现强相关关系,但上升阶段初期悬浮高度出现振荡,导致不是严格线性相关关系。

表 5 飞行器悬浮高度与表面温度相关性证据及融合结果

Table 5 Correlation evidence and fusion results between aircraft suspension height and surface temperature

基本概率 赋值函数	相关概率	不相关概率	未知信息 概率
m_1	0.999 458	0.000 542	0
m_2	0.999 824	0.000 176	0
m_3	0.999 663	0.000 337	0
融合结果	1	0	0

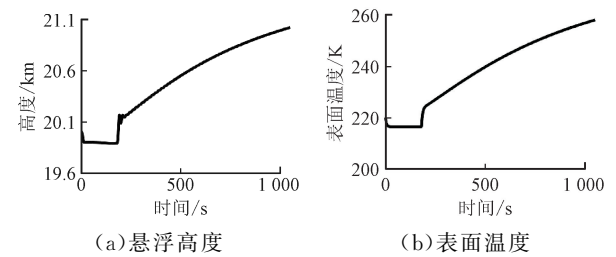


图 9 飞行器悬浮高度与表面温度变化过程
Fig. 9 Variation process of aircraft suspension height versus surface temperature

飞行器运动速度与氦气温度的相关性证据如表 6 所示。Pearson、Spearman 相关性证据结果表明速度与氦气温度不相关,但是距离相关性证据结果表明二者存在一定相关性。此时存在证据冲突,采用式(10)计算支持因子,可得 $S_{12}=0.998\ 9$, $S_{13}=0.859\ 9$, $S_{23}=0.834\ 4$,由此可得支持度分别为 1.858 8、1.833 4 和 1.694 3。分析表 4 中速度与氦气温度的 Pearson 与 Spearman 相关系数,可得出两者支持度接近。选择支持度最高的证据进行修正, m_1 的支持度最高,故为关键证据。

表 6 飞行器运动速度与氦气温度的相关性证据

Table 6 Correlation evidence between aircraft motion speed and helium temperature			
基本概率 赋值函数	相关概率	不相关概率	未知信息 概率
m_1	0.340 000	0.660 000	0
m_2	0.310 000	0.690 000	0
m_3	0.670 000	0.330 000	0

根据 $w_i = \frac{p(m_i)}{\max_{1 \leq i \leq 3} p(m_i)}$,可得各证据权重分别为 1、0.986 3 和 0.911 5。 m_1 权重最大,所以证据融合结果应更接近 m_1 。证据修正及融合后的结果如表 7 所示。融合结果接近 m_1 ,判断二者不相关,没有受到距离相关系数的干扰,本文所提方法在证据冲突小的情况下依然有效。

表 7 修正的飞行器运动速度与氦气温度相关性证据及融合结果

Table 7 Corrected correlation evidence and fusion results between aircraft motion speed and helium temperature

修正基本概率 率赋值函数	相关概率	不相关概率	未知信息 概率
m'_1	0.340 0	0.660 0	0
m'_2	0.305 8	0.680 5	0.013 7
m'_3	0.610 7	0.300 8	0.088 5
融合结果	0.349 8	0.579 1	0.026 0

4.2 证据冲突大情况下的相关性分析结果

空气温度与氦气体积的相关性证据如表 8 所示。分析表 8 可知,Pearson 相关性证据表明二者不相关,而 Spearman 和距离相关性证据表明二者相关。此时,Pearson 相关性证据为高度冲突的证据。采用式(10)可得 $S_{12}=0.053\ 9$, $S_{13}=0.064\ 1$, $S_{23}=1$ 。由此可获得支持度分别为 0.118、1.053 8

和 1.064 1。选择支持度最高的关键证据 m_3 进行修正,根据 $w_i = \frac{p(m_i)}{\max_{1 \leq i \leq 3} p(m_i)}$ 可得,各证据相对于关键证据的权重分别为 0.110 9、0.990 4 和 1。 m_3 权重最大,所以融合结果应该更接近 m_3 。根据式(12)、(13)对证据进行修正,表 9 是证据的修正及融合结果,融合结果接近 m_3 ,判断二者相关。图 10 是空气温度与氦气体积的变化过程曲线,图中空气温度随着氦气体积的增加而减小,但是数据中出现了较大的脉冲扰动,从而影响了 Pearson 相关系数的分析结果,因此分析结果应表明二者相关。

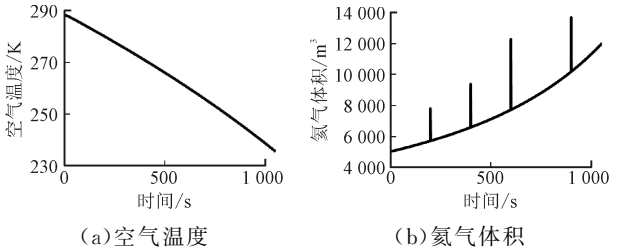


图 10 空气温度与氦气体积变化过程

Fig. 10 Variation process of air temperature versus helium volume

表 8 空气温度与氦气体积相关性证据

Table 8 Correlation evidence between air temperature and helium volume

基本概率 赋值函数	相关概率	不相关概率	未知信息 概率
m_1	0.000 019	0.999 981	0
m_2	0.964 692	0.035 308	0
m_3	0.958 082	0.041 918	0

表 9 修正的空气温度与氦气体积的相关性证据及融合结果

Table 9 Corrected correlation evidence and fusion results between air temperature and helium volume

修正基本概率 率赋值函数	相关概率	不相关概率	未知信息 概率
m'_1	0.000 002	0.110 862	0.889 136
m'_2	0.955 410	0.034 960	0.009 630
m'_3	0.958 082	0.041 918	0
融合结果	0.965 453	0.023 751	0.010 796

为验证本文所提方法的有效性,与其他文献方法进行对比,结果如表 10 所示。经典 D-S 证据理论融合结果极易受到冲突证据的影响;文献[14]方法在处理冲突时,将冲突部分全部赋给未知信息,导致

未知信息过大,分析结果造成偏差;文献[16]方法对平均证据进行 D-S 证据理论组合,但忽略了证据之间的支持度,结果并不理想;文献[18]方法采用相容系数计算证据权重,考虑了证据之间的相互支持度,融合结果相对较好。本文方法的相关概率计算结果相比文献[18]计算结果提高了约 6.55%,比文献[16]计算结果提高了约 26.84%,故本文利用基于支持因子的证据理论融合算法使相关性分析结果的可信度进一步提高,有效解决了相关性分析中的证据冲突。

表 10 空气温度与氦气体积的相关性证据融合结果对比

Table 10 Correlation evidence fusion results comparison between air temperature and helium volume			
合成方法	相关概率	不相关概率	未知信息概率
经典 D-S 方法	0.011 731	0.988 269	0
文献[14]方法	0.000 018	0.001 479	0.998 503
文献[16]方法	0.761 131	0.238 869	0
文献[18]方法	0.906 135	0.042 423	0.051 442
本文方法	0.965 453	0.023 751	0.010 796

5 结 论

临近空间飞行器长时间滞空飞行,状态发生异常甚至系统故障的频率大大增加。遥测数据相关性分析结果有助于监测飞行器的状态及检测故障。

(1)针对实际工程中遥测数据关联复杂、数据量大,实时下发伴有噪声等特点,考虑现有的单一相关性分析方法无法同时解决上述问题,采用融合优势互补的相关性分析方法,以提高分析结果的可靠性。评价实验证明,Pearson 相关系数、Spearman 相关系数与距离相关系数性能互补,3 种相关系数融合有助于更全面地分析遥测数据的内在关联。

(2)由于 Pearson 相关系数、Spearman 相关系数与距离相关系数在遥测数据相关性分析的实践中存在结果冲突的问题,因此在相关系数分析评价与优势组合的基础上,提出了基于支持因子的证据理论融合算法,通过支持因子计算权重及修正相关性证据,并分配修正证据的冲突基本概率赋值函数,解决了证据冲突,取得了更客观的融合结果。

(3)实验结果表明,本文所提方法能够有效解决证据冲突问题,提高了临近空间飞行器遥测数据相关性分析结果的可信度。

参考文献:

[1] 康旭. 时序遥测数据异常检测方法研究 [D]. 南京: 南京航空航天大学, 2018.

[2] 谭春林, 胡太彬, 王大鹏, 等. 国外航天器在轨故障统计与分析 [J]. 航天器工程, 2011, 20(4): 130-136. TAN Chunlin, HU Taibin, WANG Dapeng, et al. Analysis on foreign spacecraft in-orbit failures [J]. Spacecraft Engineering, 2011, 20(4): 130-136.

[3] 杨甲森, 孟新, 陈托, 等. 基于遥测数据相关性的航天器异常检测 [J]. 仪器仪表学报, 2018, 39(8): 24-33. YANG Jiasen, MENG Xin, CHEN Tuo, et al. Anomaly detection of spacecraft based on the telemetry data correlation [J]. Chinese Journal of Scientific Instrument, 2018, 39(8): 24-33.

[4] MARTÍNEZ-HERAS J A, DONATI A, SOUSA B, et al. DrMUST-a data mining approach for anomaly investigation [C]//SpaceOps 2012 Conference. Reston, VA, USA: AIAA, 2012: AIAA 2012-1275109.

[5] 纪德洋, 金锋, 冬雷, 等. 基于皮尔逊相关系数的光伏电站数据修复 [J]. 中国电机工程学报, 2022, 42(4): 1514-1522. JI Deyang, JIN Feng, DONG Lei, et al. Data repairing of photovoltaic power plant based on Pearson correlation coefficient [J]. Proceedings of the CSEE, 2022, 42(4): 1514-1522.

[6] JI Chen, WANG Jue, ZHANG Guoan. Approximate expression for the mutual information of dense PAM [J]. IEEE Communications Letters, 2018, 22(11): 2182-2185.

[7] RESHEF D N, RESHEF Y A, FINUCANE H K, et al. Detecting novel associations in large data sets [J]. Science, 2011, 334(6062): 1518-1524.

[8] SZÉKELY G J, RIZZO M L, BAKIROV N K. Measuring and testing dependence by correlation of distances [J]. The Annals of Statistics, 2007, 35(6): 2769-2794.

[9] LI Zhe, AN Meizhen, BAI Linhou, et al. Correlation analysis on telemetry data of manned spacecraft [C]//2018 Chinese Control and Decision Conference (CCDC). Piscataway, NJ, USA: IEEE, 2018: 377-380.

[10] WEN Tao, DONG Deyi, CHEN Qianyu, et al. Maximal information coefficient-based two-stage feature selection method for railway condition monitoring [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(7): 2681-2690.

[11] 崔树银, 汪昕杰. 基于最大信息系数和多目标 Stacking 集成学习的综合能源系统多元负荷预测 [J]. 电力自动化设备, 2022, 42(5): 32-39.

- CUI Shuyin, WANG Xinjie. Multivariate load forecasting in integrated energy system based on maximal information coefficient and multi-objective Stacking ensemble learning [J]. *Electric Power Automation Equipment*, 2022, 42(5): 32-39.
- [12] 孙宇豪, 李国通, 张鸽. 距离相关系数融合 GPR 模型的卫星异常检测方法 [J]. *北京航空航天大学学报*, 2021, 47(4): 844-852.
- SUN Yuhao, LI Guotong, ZHANG Ge. A satellite anomaly detection method based on distance correlation coefficient and GPR model [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(4): 844-852.
- [13] 周莉, 张歆茗, 郭伟震, 等. 基于改进冲突度量的多证据直接融合算法 [J]. *电子与信息学报*, 2019, 41(5): 1145-1151.
- ZHOU Li, ZHANG Xinming, GUO Weizhen, et al. A direct fusion algorithm for multiple pieces of evidence based on improved conflict measure [J]. *Journal of Electronics & Information Technology*, 2019, 41(5): 1145-1151.
- [14] ZHAO Kaiyi, CHEN Zeqiu, SUN Shulin, et al. A novel evidence combination rule based on compromise conflict indicator and conflict focal element [J]. *Knowledge-Based Systems*, 2022, 257: 109898.
- [15] 孙全, 叶秀清, 顾伟康. 一种新的基于证据理论的合成公式 [J]. *电子学报*, 2000, 28(8): 117-119.
- SUN Quan, YE Xiuqing, GU Weikang. A new combination rules of evidence theory [J]. *Acta Electronica Sinica*, 2000, 28(8): 117-119.
- [16] SHANG Qiuyan, LI Hanwen, DENG Yong, et al. Compound credibility for conflicting evidence combination: an autoencoder-K-means approach [J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, 52(9): 5602-5610.
- [17] 赵静, 关欣, 刘海桥. 冲突证据决策新方法及应用 [J]. *北京航空航天大学学报*, 2019, 45(9): 1838-1847.
- ZHAO Jing, GUAN Xin, LIU Haiqiao. A new conflict evidence decision method and its application [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2019, 45(9): 1838-1847.
- [18] LIN Zhen, XIE Jinye. Research on improved evidence theory based on multi-sensor information fusion [J]. *Scientific Reports*, 2021, 11(1): 9267.
- [19] KONG Liang, NIAN Heng. Fault detection and location method for mesh-type DC microgrid using Pearson correlation coefficient [J]. *IEEE Transactions on Power Delivery*, 2021, 36(3): 1428-1439.
- [20] JIA Ke, YANG Zhe, ZHENG Liming, et al. Spearman correlation-based pilot protection for transmission line connected to PMSGs and DFIGs [J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(7): 4532-4544.
- [21] FANG Jian, XU Chao, ZILLE P, et al. Fast and accurate detection of complex imaging genetics associations based on greedy projected distance correlation [J]. *IEEE Transactions on Medical Imaging*, 2018, 37(4): 860-870.
- [22] FU Bo, FANG Jinwei, ZHAO Xilin, et al. A belief coulomb force in D-S evidence theory [J]. *IEEE Access*, 2021, 9: 82979-82988.
- [23] JIAO Zaibin, GONG Heteng, WANG Yifei. A D-S evidence theory-based relay protection system hidden failures detection method in smart grid [J]. *IEEE Transactions on Smart Grid*, 2018, 9(3): 2118-2126.
- [24] 刘海燕, 赵宗贵, 刘熹. D-S 证据理论中冲突证据的合成方法 [J]. *电子科技大学学报*, 2008, 37(5): 701-704.
- LIU Haiyan, ZHAO Zonggui, LIU Xi. Combination of conflict evidences in D-S theory [J]. *Journal of University of Electronic Science and Technology of China*, 2008, 37(5): 701-704.
- [25] BI Wenhao, ZHANG An, YUAN Yuan. Combination method of conflict evidences based on evidence similarity [J]. *Journal of Systems Engineering and Electronics*, 2017, 28(3): 503-513.

(编辑 杜秀杰)