

doi: 10.7690/bgzdh.2018.01.006

航天测量信息实时质量分析方法

段慧芬, 鲍俊雷, 施 斌

(中国卫星海上测控部试验技术部, 江苏 江阴 214431)

摘要: 针对航天测量信息数据分析依靠事后处理时效性差的问题, 提出一种实时数据质量分析方法。根据数据质量分析一般方法和航天测量信息的特点, 建立规范性、实时性、准确性和可靠性 4 个维度的测控信息质量分析模型和质量分析规则, 通过按需组合规则、定制分析流程, 实现了对航天测量信息的动态实时监视预警、异常数据捕获和故障判断。在测控网中测试结果表明: 该方法能够及时发现航天试验中测量信息在采集、传输、处理过程中的问题, 并对任务完成情况进行准确评价, 对其他面向实时数据流系统的数据清洗也有一定的参考价值。

关键词: 航天试验; 质量分析; 规则; 数据质量

中图分类号: TJ760.6⁺28 **文献标志码:** A

Information Real-time Quality Analysis Method in Aerospace Measurement

Duan Huifen, Bao Junlei, Shi Bin

(Experiment Technology Department, China Satellite Maritime Tracking & Control Center, Jiangyin 214431, China)

Abstract: In view of problem that bad timeliness of post processing of aerospace measurement information data analysis, put forward the method for real-time data quality analysis. According to feature data quality analysis method and aerospace measurement information, establish measurement and control information quality analysis model and quality analysis rule of standardization, reality, accuracy and reliability. Recognize rule according to requirement and customization analysis process, realize dynamic real-time monitoring pre-alarmed, abnormal data capturing and failure judgment. The result in test and control net show that the method can find out problems of measurement information in collecting, transportation and processing in aerospace test. Evaluate the completion of the mission accurately. Furthermore, it has some values for other real-time data stream systems.

Keywords: spaceflight experiment; quality analysis; rule; data quality

0 引言

随着航天试验任务的发展, 试验的内容越来越多、试验的难度也越来越大、试验的密度越来越高, 这些变化对测量过程质量提出了更高的要求。在联调、合练和任务实施中, 作为指挥决策人员, 不仅要关注目标的飞行态势, 还需要及时掌握测控信息的采集、传输、处理过程工作情况, 通过信息层面及时反映跟踪情况, 及时发现信息传输、软件和设备的故障, 提前发现产生异常的原因, 及时采取纠正和预防措施。同时, 对测量任务完成情况进行准确的评价, 以保证跟踪测量的综合效能。对每次任务中汇聚于指挥中心的大量测量信息建立一套通用、高效、规范的质量分析方法, 引入简明、灵活的分析流程和工具, 是非常必要的。

1 数据质量分析一般方法

1.1 数据质量分析的定义

数据质量是数据内在和外部特性的总和, 以此

构成其满足给定需求的能力^[1]。该定义参照了美国质量管理协会 ASQC 对“质量”的解释。基于数据质量的定义, 在不同的应用系统中, 数据不仅能反映客观事物, 还要完成信息的传递。对不同的行业, 由于其行业特点不同, 对数据质量特性要求也不相同^[2-4], 但作为数据本身应普遍具备的质量特性包括准确性、一致性、完整性、总体性、广度性、历史性、正确性、潜伏性、缺乏性、冗余性、相关完整性等^[5-7]。这些特性是进行数据质量分析的依据。对于不同的数据应用领域, 可根据领域数据的特点, 在普遍的数据质量特性基础上, 进行选择和质量特性实例化, 开展数据质量分析工作。

1.2 数据质量分析方法

根据数据质量的定义, 数据也是一种产品。每一种产品都有它固有的质量评价标准, 通常把这种数据的质量评价标准定义为数据质量元素^[8]。数据质量元素可以用若干个质量特性描述, 数据的质量属性是数据属性的扩展, 与数据属性不同的是质量

收稿日期: 2017-09-29; 修回日期: 2017-11-02

作者简介: 段慧芬(1963—), 女, 吉林人, 硕士, 高级工程师, 从事航天测控、计算机应用等研究。

属性描述了对对象在数据质量方面的约束。该约束通常定义为数据质量分析中的规则^[9-12]。

数据质量分析的一般过程可归纳为：确定数据集分析的数据元素及元素间相互关系；选择分析的质量元素，根据数据集的应用背景制定与质量元素相对应的规则集；运用规则对数据集进行质量分析，对分析结果统计计算，得出分类及综合评估结果。数据质量分析的基础是制定合理的分析规则，分析规则的确定与数据的质量特性密切相关。一般将数据质量分析规则归纳为完整性规则、一致性规则、准确性规则、深度性规则、及时性规则、关联性规则、冗余性规则^[13]等。这些规则既是对数据质量的约束条件，又是对数据质量分析的标准。在通用性规则的指导下，细化为对应相关领域业务的分析规则^[14]。分析规则本身应当是可扩展和可描述的。分析规则制定了数据质量的分析标准和判断依据，是数据质量分析的关键^[15]。

2 航天试验信息质量分析方法

2.1 测量信息质量分析数据集

在航天试验中，根据测控设备测控体制不同，分为单脉冲雷达测控设备、微波统一测控系统设备、光学测量设备及其他辅助测量设备。通过不同的测量设备，实时采集到测控目标的光测、外测、遥测和遥控信息，这些信息以不同的数据内容构成测控信息质量分析的数据集。数据质量分析的过程是对测量信息进行数据挖掘的过程。每类测量元素承载的数据内容虽然不同，但是通过统一的格式约定，使每类测量元素的质量元素可检查、可度量、可追踪。一个好的数据集视图设计，不仅可完整表示数据本身的物理属性，而且还可以表示出数据的扩展属性。不同测控信息数据集的通用表示方法如图 1。

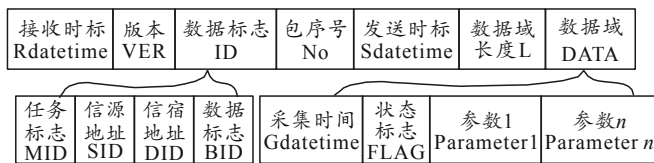


图 1 测控信息通用数据集视图

2.2 测控信息质量分析规则

根据试验信息的特点，将数据质量分析模型应用于航天测控信息质量分析中，制定出具有可操作性的质量分析规则。规则不仅确定了质量分析的判断标准，而且是开展质量分析的基础和前提。这些规则可以单独使用，也可以相互关联使用。规则的

制定首先要符合业务信息的自身规律，保证质量分析的正确性。其次要具有通用性，使规则可以适应不同的信息内容，最后还要具有可操作性，保证规则易于实现，特别是在实时信息分析中，更为重要。

测控信息质量分析对象是实时的测量数据，主要关注包括重复对象检测、缺失数据处理、异常数据检测、逻辑错误检测、不一致数据处理等。具体归纳为规范性、实时性、准确性和可靠性 4 个测控信息的质量元素。每个质量元素都包含了不同的约束规则。每个约束规则要相互独立，易于快速进行判断。具体内容见表 1。

表 1 测控信息质量元素

分类	约束规则	规则内容
规范性	信息交换格式的规范性规则	每个测源采集传输的信息类别的信息格式是固定的,通过信息类别与数据帧的长度相符合的判断,作为信息格式是否正确的通用判断条件。
	信息测源的规范性规则	信源在规定的信源数据字典中。
	信息类别的规范性规则	信息类别在规定的信息类别数据字典中。
	校验单元的规范性规则	对有校验单元的信息,判断帧同步码是否正确,状态信息是否有效。
实时性	发送时延判断规则	对同类数据每帧的数据采样时间与数据发送时间进行比较,判断数据是否被及时处理发送。
	接收时延判断规则	对同类数据每帧的数据采样时间与数据接收时间进行比较,判断数据是否被及时处理接收。
	时序判断规则	对同类数据比较相邻帧的数据采样时间,分析帧频变化是否符合约定的采样频率,可以判断数据是否被及时采集到。
	包序号判断规则	对每个测源发送的每类信息分析包序号是否符合按整数规律连续变化的规律,分析包序号是否有中断、重复、倒序等异常情况。
准确性	实测值与期望值的偏差判断规则	对有期望值的测控信息,将实测值与期望值进行比较,给出允许的误差范围,判断实测信息的准确性。
	误差分析判断规则	根据测量误差的分类,利用误差分析方法,计算出随机误差、系统误差、过失误差,根据误差的计算结果,判断实测信息的精密度和准确度。
	测量范围的判断规则	根据测源采集到有效数据的时段,与下达任务的时段比较,判断测量是否满足下达的任务要求。
可靠性	链路可靠性判断规则	分析收发链监信息的包序号,包序号按整数规律连续变化,没有中断、重复、倒序,可以判断数据收发传输链路工作基本正常。
	关键数据三判二规则	对试验中的关键数据,都规定了重复发送应遵循的规则,对每个测源发送的关键数据,检查重复的帧数、发送帧的间隔时间、数据内容是否保持一致,确保关键数据发送是可靠的。对接收的重要数据,检查接收的帧数、三判二验证通过的帧、未通过三判二验证的原因。
	双平面传输数据的比较规则	对双平面传输的数据分别进行分类统计,将相同时段内的统计结果进行比较,判断是否一致,分析出哪个路由平面数据传输的可靠性更好。

3 质量分析系统实现

3.1 应用系统结构

试验任务实时信息质量分析系统是一个以试验

数据为中心的应用系统，系统结构如图 2 所示，分为数据层、处理层、应用层。其目的是将所有可预见的变化进行隔离，从而提高系统的可扩展性。对试验任务来说，数据层包括分析对象的数据来源及形式，分析对象包括外部实时数据输入的数据、实时记录保留的数据和经过实时处理后的结果数据。处理层包括进行分析前对分析规则的维护，规则的增、删、改，规则判断中各类期望值的设置；数据接收包括实时数据分析时数据的网络接收、保留；数据处理指对与质量分析相关的数据帧中数据元素的解算；规则调度指根据不同的应用需求，调用相应的规则进行分析、统计。应用层包括规则管理，进行规则库的维护操作。实时监控预警，对经过规则判断认为异常的信息以直观的图示和声音进行提示，同时捕获记录下异常信息；数据查询，对异常信息或用户关注的信息进行追踪查询和数据对比检查；分析流程定制，将分析规则与分析对象建立关联，为规则调度提供依据；分析报告，分析结束后形成分析报告，记录下分析的结果，作为质量分析和指挥决策的依据。

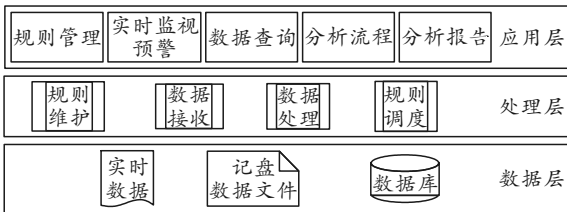


图 2 应用系统结构

3.2 实时分析流程

在实时分析系统设计中，将分析过程分为分析准备、分析实施、分析查询 3 个阶段。具体分析流程如图 3 所示。分析准备阶段可以根据经验的积累将新的分析规则增加到规则库中，将不适用的规则从规则库中删除，对不完善的规则进行修改，完成对规则库的维护；每种试验任务使用的运载工具型号不同，航天器型号不同，飞行轨道设计不同，导致分析采用的期望值不同，需要在准备阶段对有变化的期望值进行重新设置；筛选出适用于本次任务的规则集合。定制分析流程是在实时分析过程中为程序规定出规则的使用顺序，特别是在怀疑数据采集有问题时，针对各种导致问题发生的原因，调整规则调用顺序，帮助分析查找故障原因。除了进行实时分析外，对于实时记录下的数据文件还可以在事后进行有针对性的分析。

在试验中，进行实时数据质量分析工作，可以

对任务过程进行动态监视预警，及时捕获异常数据，根据异常情况进行故障判断，同时对试验任务完成情况给出定量分析结果。

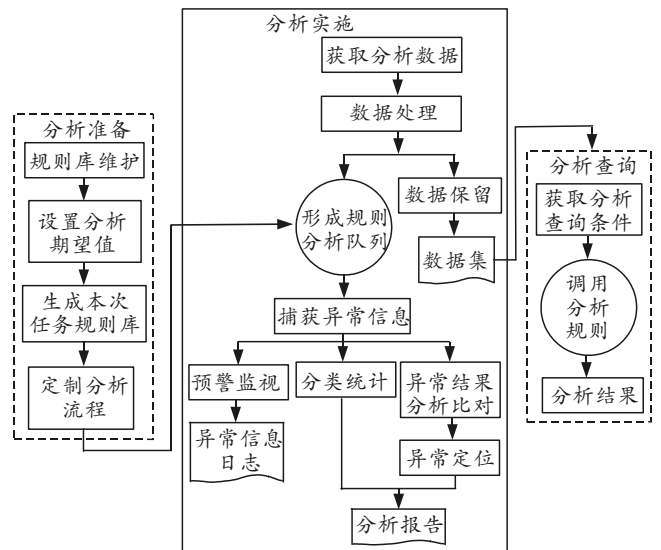


图 3 数据质量分析流程

3.3 系统应用实例

根据数据质量分析规则，笔者采用 Visual 2010.net 开发平台，用 C#语言开发出了一个试验任务实时信息质量分析系统。系统由一个主线程和若干子线程构成，数据监视模块前台主线程处于等待状态，后台线程有数据接收、保存、分析进行数据质量分析，一旦后台线程捕获异常，向前台主线程发送通知，前台线程自动发出预警信息，响应分析人员的需求，分析线程利用规则去执行推理，快速诊断确定异常原因。

系统将表 1 中的 4 类、15 个维度的数据质量约束规则，建立了任务质量分析规则库。通过规则定制信息质量分析模型，按照型号任务的航天器飞行时序、信源、信类、信息交换格式和正常值范围等，实时对异常信息进行分析诊断，支持对数据进行快速查询定位和基于源码、字段等实时分析的能力。具备数据的连续性统计、断点统计、时序统计、异常数据值统计、网络流量统计等全方位数据统计；对数据格式异常和内容异常的数据帧进行图形化和表格化分析；支持文件、表格、图形等多种视图形式输出统计结果。

在试验中，为实现链路传输质量的动态监视预警，在规则库的基础上先进行规范性检查，对符合规范性约束条件的数据包再进行实时性判断规则检查和采样时间的时序判断规则检查。图 4 通过色块颜色变化直观标识出不符合包序号连续变化规则的

数据帧，黑色表示数据有中断，提示一目了然。图 5 通过表格中数据记录背景颜色变化，用阴影标识出数据采样时间时序不连续的数据帧记录。图形化分析结果显示方式直观明了，易于快速发现问题；表格化分析结果显示方式，直接对发现异常的数据帧进行定位，易于分析问题。

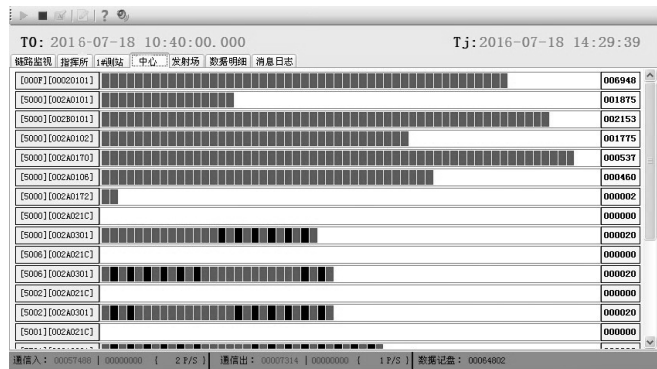


图 4 数据异常检查结果图形化显示



图 5 数据异常分析结果表格化显示

4 结束语

笔者综合分析了航天试验信息的特点，运用数据质量分析方法模型，给出了航天试验信息质量分析规则和基于规则的航天试验信息质量分析方法，通过软件系统，将信息质量分析方法应用到任务中，对任务中的实时信息进行检查和相关性分析。经过在试验任务中的应用，对快速准确发现信息传输异

常、信息错漏发挥了显著作用，为任务指挥决策提供了方便直观的分析工具，对实时信息质量分析提供了可以借鉴的方法。笔者侧重于数据质量分析方法和分析过程的工程应用研究，对分析结果的定量评价还需要继续开展研究。

参考文献：

- [1] 韩京宇, 徐立臻, 董逸生. 数据质量研究综述[J]. 计算机科学, 2008, 35(2): 1-5.
- [2] 丁海龙, 徐宏炳. 数据质量分析及应用[J]. 计算机技术与发展, 2007, 17(3): 236-238.
- [3] 吴永欢. 数据质量管理模型及应用研究[J]. 中国电机(技术版), 2013, 64(8): 46-50.
- [4] 杨青云, 赵培英, 杨冬青. 数据质量评估方法研究[J]. 计算机工程与应用, 2004, 40(9): 3-4, 15.
- [5] 史峰. 基于规则库的数据质量分析[J]. 武汉职业技术学院学报, 2010, 9(3): 79-83.
- [6] 宋金玉, 陈爽, 郭大鹏, 等. 数据质量及数据清洗方法[J]. 指挥信息系统与技术, 2013, 4(5): 63-69.
- [7] 滕东兴, 曾志荣, 杨海燕, 等. 一种面向关系型数据的可视质量分析方法[J]. 软件学报, 2013, 24(4): 810-824.
- [8] DANETTE Mc Gilvray. 数据质量工程实践[M]. 刁兴春, 曹建军, 张健美, 等. 译. 北京: 电子工业出版社, 2010: 138-196.
- [9] 刘益江. 数据仓库的数据质量分析与评价[D]. 广州: 广东工业大学, 2012.
- [10] 丛慧刚. 基于规则的数据中心数据质量研究[D]. 大庆: 东北石油大学, 2012.
- [11] 袁满, 张雪. 一种基于规则的数据质量评价模型[J]. 计算机技术与发展, 2013, 23(3): 81-84.
- [12] 曹建军, 刁兴春, 陈爽, 等. 数据清洗及其一般性系统框架[J]. 计算机科学, 2012, 39(11A): 207-211.
- [13] 安静. 基于元数据动态获取的数据质量检查技术[D]. 北京: 中国地质大学, 2015.
- [14] 田雪松. 数据质量分析评估模型的设计与实现[D]. 长春: 吉林大学, 2008.
- [15] 文峰. 数据组织过程中的数据质量评价研究[J]. 软件导刊, 2013, 12(11): 132-134.