

# 多流多科目实时数据处理算法的设计与实现

王 灏, 郝 朝, 刘升护

(中国飞行试验研究院 测试所, 西安 710089)

**摘要:** 随着飞行试验测试技术的快速发展, 单架次需要处理的试飞科目越来越多, 而多科目处理参数分布在不同的网络流中, 这对机载实时数据处理的效率带来严峻的挑战; 利用“合并—处理—分流”的方法, 设计一种多流多科目网络数据实时处理算法; 该算法将多科目处理参数按照参数名、采样率等信息进行整合形成合并参数列表; 在实时处理时根据合并参数信息实时处理多网络流中的数据, 并将处理结果存储在数据缓存中; 最后将多科目数据进行分流即可得到处理结果; 测试结果表明, 该算法可实际应用于多流多科目数据处理中, 大幅提高实时处理的效率, 同时可以得到准确的数据处理结果。

**关键词:** 多数据流; 多科目; 参数合并; 实时处理; 数据分流

## Design and Implementation of Multi—stream and Multi—subject Real—time Data Processing Algorithm

Wang Hao, Hao Zhao, Liu ShengHu

(Chinese Flight Test Establishment, Xi'an 710089, China)

**Abstract:** Abstract: With the rapid development of flight test and test technology, more and more test flight subjects need to be processed in a single operation, and multi—subject processing parameters are distributed in different network flows, which brings severe challenges to the efficiency of on—board real—time data processing. A multi—flow multi—subject network data real—time processing algorithm is designed by using the method of “merge—process—split”. The algorithm integrates the multi—subject processing parameters according to the parameter name and sampling rate to form a combined parameter list; in real—time processing, the data in the multi—network stream is processed in real time according to the merged parameter information, and the processing result is stored in the data cache; The multi—subject data is shunted by the location retrieval method to obtain the processing result. The test results show that the algorithm can be applied to multi—stream and multi—subject data processing, which can greatly improve the efficiency of real—time processing and obtain accurate data processing results.

**Keywords:** multi—stream; multi—subject; parameter merging; real—time processing; data shunting

## 0 引言

随着航空技术的飞速发展, 现代飞行器的集成度和复杂度越来越高<sup>[1-2]</sup>, 传统的单架次单科目试飞远不能满足飞行试验的需求, 多科目试飞已经成为飞行试验发展的必然选择。多科目试飞需要的参数量大、采样率高<sup>[3-4]</sup>, 为解决这一问题, 当前飞行试验已广泛采用网络化测试架构, 在网络化测试架构下采集器将采集的数据以网络包的形式发送至测试系统, 多科目需要处理的参数就分布在这些网络包中<sup>[5]</sup>。由于数据以网络包的形式在测试系统上传输, 所以可以使用机载实时数据处理系统来实时处理数据, 这会极大地缩短获取结果数据的时间。但是实时处理设备是机载嵌入式设备, 对数据处理的性能有限, 如果没有高效的多流多科目数据处理算法, 实时处理时数据就会发生错误从而得到不到正确的处理结果<sup>[3]</sup>。对于当前多数据流多科目数据处理的需求, 如何在多数据流的情况下实时处理多个科目, 在机载实时处理设备性能允许的范围内, 以最大的效率处理数据, 是飞行试验实时数据处理中必须解决的

一个关键问题。

## 1 多流多科目实时数据处理

### 1.1 多流多科目实时数据处理特点

飞行试验是武器装备设计、生产、试验的三大环节之一<sup>[6-7]</sup>, 现代飞行器在飞行试验阶段需要处理性能、品质、发动机、雷达、航电等众多试飞科目的数据, 传统单架次单科目的数据处理方式不能满足众多试飞科目的需求, 为了提高试飞数据处理的效率和及时性, 多科目试飞数据处理成为飞行试验的必然选择<sup>[8]</sup>。另外由于武器装备的试飞周期比较紧, 多科目数据处理有助于加快飞行试验的进度, 缩短试飞周期。

多科目试飞数据处理需求的测试参数多、采样率大, 某型飞机在进行飞行试验时需要采集一百多条数据流, 处理三十多个科目, 共计四万多个参数、科目的采样率从 8 到 2048 不等<sup>[9-10]</sup>。为了解决参数个数多, 采样率高的问题, 现在飞行试验中广泛采用网络化测试系统架构。网络化测试系统架构下, 采集器将采集的数据以网络包的形式发送至测试系统, 多台采集器会发送多条网络数据流, 多科目需要处理的参数就分布在这些网络数据流中<sup>[11-12]</sup>。

为提高数据处理的效率, 众多型号在飞行试验过程中使用了机载实时数据处理系统, 该系统接收机载测试系统

收稿日期: 2019-12-12; 修回日期: 2020-01-13。

作者简介: 王 灏(1992-), 男, 陕西渭南人, 硕士研究生, 助理工程师, 主要从事飞行试验实时监控与数据处理技术方向的研究。

的网络包数据, 然后根据多科目处理参数进行数据处理并将结果写入结果文件<sup>[13]</sup>。机载实时数据处理极大地缩短了数据获取的时间, 但是由于多科目处理时参数数量多且采样率大, 而机载实时数据处理设备的性能有限, 所以需要设计高效的多流多科目数据处理算法, 使得数据处理设备能够满足多流多科目数据处理的需求。多流多科目数据实时处理如图 1 所示。

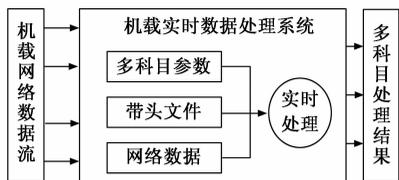


图 1 多流多科目数据实时处理

根据实际试飞数据处理的情况, 总结出多流多科目数据处理有如下几个特点:

- 1) 机载测试系统有多条网络流, 每条数据流由连续的网络包组成, 每个网络包有各自独立的 KEY 值, 该值是该网络流的唯一标识;
- 2) 网络包的带宽比较高, 例如某型机可达 50~60 Mbps, 在实时处理时需要快速处理网络包数据, 否则就会发生错误;
- 3) 每个科目都有各自的参数组, 实时处理时需要根据每个科目的参数组处理数据, 处理完成后将科目需要的结果发给课题;
- 4) 每个科目需求的采样率是一固定值, 但科目之间的采样率不完全一致;
- 5) 同一参数在不同科目中需要处理的采样率不完全一致, 但参数的采样率和课题需求的采样率一致;
- 6) 多科目需要的参数可能出现重复, 例如: 气压高度、马赫数等飞机基本参数几乎所有的科目都需要进行处理。

### 1.2 多流多科目实时数据处理方法

为解决多流多科目实时处理时的效率问题, 设计多流多科目实时数据处理算法, 最简单的办法是将多科目数据按照单科目数据处理的方法逐个处理, 具体做法是: 读取网络数据包数据, 处理第一个科目的参数, 第一个科目的参数处理完后处理第二个科目的参数, 以此类推, 直至处理完所有的科目, 然后读取下一个网络包数据进行同样的处理。按照单科目处理方法循环处理存在如下几个方面的问题:

- 1) 算法的处理效率比较低, 对于多科目中的重复参数处理多次, 浪费了宝贵的计算资源。
- 2) 单科目处理完后需要写入结果文件, 频繁地进行 I/O 操作, CPU 的利用率不高;
- 3) 数据处理时大量时间花费在对参数的轮询查找上了, 浪费了时间资源, 难以保证数据处理的实时性。

因为多科目参数中有部分参数会重复出现, 所以一个高效的办法是: 将多科目参数合并成一个参数列表, 合并参数列表中包含所有的多科目处理参数且参数没有重复, 在处理时按照合并后的多科目参数进行实时处理, 处理完

成后将数据进行分流得到各科目的处理结果。采用多科目数据合并处理的有如下几个优点:

- 1) 多科目参数中重复的参数只用处理一次, 极大地提高了实时数据处理的效率;
- 2) 可以通过合理设置数据缓存的办法, 减少 I/O 操作, 提高 CPU 的利用率, 在同等设备性能的情况下可以处理更多的参数;
- 3) 数据处理时只需要对参数进行一遍轮询查找, 节省了数据处理的时间, 提高了实时性。

由此可知: 采用“合并—处理—分流”方法处理多流多科目数据, 能够有效提高数据处理的效率、保证实时性, 所以将其作为机载实时数据处理时的多流多科目处理算法。

## 2 多流多科目实时数据处理算法的实现

为解决多数据流情况下试飞科目的快速增长与实时数据处理效率之间的矛盾, 设计并实现多流多科目实时数据处理算法, 在保证数据处理实时性的前提下实现多科目数据的高效快速处理。按照“合并—处理—分流”的方法, 首先将多科目处理参数按照参数基本信息进行合并, 形成合并参数列表, 在合并参数列表中参数不会重复且没有缺失, 同时该列表是实时处理时需要处理的参数列表; 然后在实时处理时, 读取并分析每一个网络包数据, 对合并参数列表中的参数进行码值提取和物理量转化, 通过合理设置缓冲区, 将处理的结果数据存入缓冲区中; 最后在处理完成后, 按照数据分流的方法, 将多科目中的每个参数从缓冲区中读取出来写入多科目的结果文件中, 至此完成多流多科目数据实时处理。该算法的整体结构如图 2 所示。

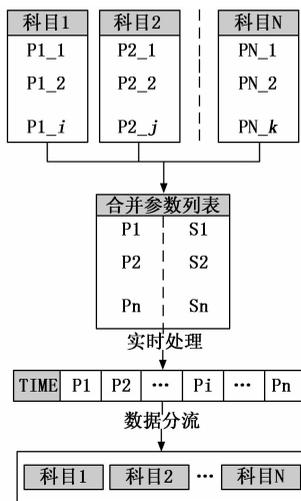


图 2 多流多科目数据处理算法结构

### 2.1 合并多科目处理参数

合并多科目处理参数有助于减少实时处理时对于多科目参数中重复参数的重复处理, 所以数据合并时以减少参数重复为主要合并原则, 多科目参数合并的方法是:

- 1) 剔除多科目参数中重复的参数, 保证合并参数列表中每个参数只出现一次, 并且没有遗漏;
- 2) 对于同一参数的不同采样率, 取最大的采样率作为

合并参数列表的采样率;

最终形成的合并参数列表包含的参数为 P1、P2、...、Pn, 采样率分别为 S1、S2、...Sn。此外为了在数据处理时准确定位某个数据包里要处理的参数, 还需要分析合并参数列表中每个参数所属的数据包。具体做法是:

1) 分析机载信息配置文件, 读取网络包的参数信息, 将参数名和数据包一一对应, 找出机载测试系统的网络包中都包含哪些参数;

2) 按照合并参数列表的参数名查找每个参数所属的数据包, 形成待处理数据包列表, 并详细列出每个待处理数据包中包含的待处理参数。

### 2.2 合并参数实时处理

在实时数据处理时, 如果对于每个参数进行解算完成之后直接写入结果文件, 这样会频繁的进行 I/O 操作, 降低了 CPU 的利用率, 所以需要设置缓存来存储数据处理的中间结果, 处理完成后再进行存储。为了方便处理及实现, 以一秒的结果数据为单位进行缓存, 设计图 3 所示的缓冲区, 该缓冲区为一个数组, 数组前三个元素存放时、分、秒信息, 后面的元素存放合并参数列表的处理结果。在实时数据处理时, 将一秒钟数据处理的结果放入该缓存中, 当缓冲区存满后, 对数据进行分流写入多科目的结果文件中, 然后清空缓存区继续存储下一秒的数据。



图 3 实时处理缓冲区结构

合并参数列表中的参数处理结果在缓冲区的存放的原则是:

1) 每个参数按照采样率的大小占据缓存中的连续元素, 参数  $P_i$  的采样率为  $S_i$ , 那么该参数在缓存中占据  $S_i$  个连续元素;

2) 每个参数在缓存中的起始位置  $POS_i$  由该参数在参数列表总的位置及之前参数的采样率共同决定,  $POS_i = 3 + \sum_{j=1}^{i-1} S_j$ , 即参数  $P_1$  开始位置为 3,  $P_2$  参数位置为  $3 + S_1$ , 以此类推;

3) 每个参数的缓冲区的元素按序存放采样时刻的参数值, 例如参数  $P_i$  的采样率为 8, 那么这 8 个元素依次存放 000、125、250、375、500、625、750、875 毫秒时刻的数据。

缓冲区设置完成之后对数据进行实时处理, 处理过程主要是在网络流中找出待处理数据包和待处理参数, 然后进行码值提取及物理量转化, 最后将结果写入缓存中, 具体过程如图 4 所示。

1) 读取网络包数据, 判读网络包的 KEY 值, 分析其是否属于需要处理的网络包, 若不需要处理, 则跳过该数据包读取下一个网络包;

2) 分析数据包处理列表, 找出需要处理的网络包都包含哪些待处理参数, 详细列出处理参数的信息, 包括参数名、采样率、校准方式等

3) 对于每一个需要处理的参数, 在网络包中进行码值

提取及物理量转化。码值提取可通过直接取值的方式获取, 一般包括单字提取、双字提取、三字提取和四字提取; 物理量转化是将码值按照参数校准的方式转化成物理量, 校准方式一般包括点对校准、直线校准、双曲线校准、多项式校准等;

4) 判断该网络包的时间, 分析待处理的参数属于缓存中的位置, 将处理参数的结果按照参数在缓存中的存放方法放在缓存中;

5) 读取下一个网络包, 重复进行上述操作, 完成数据实时处理。

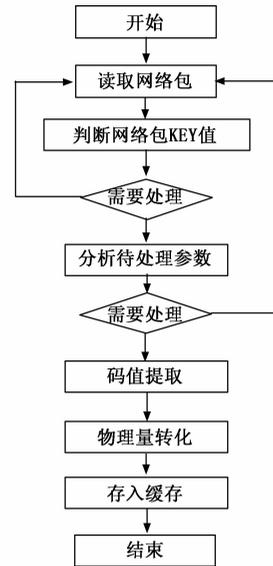


图 4 实时数据处理流程

### 2.3 多科目参数数据分流

数据处理时将处理结果放入了缓存中, 但课题需要的是单个科目的结果文件, 所以需要将缓存中的数据写入结果文件中。缓存中存储的是合并参数列表中的数据, 而要得到多课题结果文件就需要对数据进行分流。在数据缓冲区的数据存满后进行分流并将结果写入课题需求的结果文件中, 数据分流的方法是:

1) 分析每一个科目的参数在合并参数列表中的位置, 比较每一个科目参数需求的采样率与综合参数列表中该参数的采样率, 用综合参数采样率除以需求采样率得到缓存中的取值步长。

2) 分析第一个科目的参数, 按照该科目待处理参数的顺序, 依次在缓冲区中进行取值, 先取第一个时刻的数据, 然后根据步长依次取下一时刻的数据, 把该秒的参数值都写入文件中;

3) 按照步骤 2) 的方法处理其他科目的参数, 将所有科目数据分流完成并将处理的物理量写入结果文件中;

4) 改变缓冲区前三个元素的时间, 将秒数加一, 然后该缓冲区存储下一秒的处理数据, 当缓冲区存满后, 重复进行上述操作, 完成数据分流。

经过“合并—处理—分流”的过程, 所有科目中需求

参数都已完成处理, 每个科目都能得到正确的处理结果, 并且多科目中的参数都不会被重复计算, 这样节省了时间和空间资源, 保证了数据处理的实时性。

### 3 实验测试与验证

利用 C 语言实现了多流多科目数据实时处理算法, 应用该算法处理多流多科目数据, 在实验中使用某型号的机载实时数据处理设备为系统运行环境, 以某机载测试系统发送的网络流为处理对象, 该系统共有 128 条网络流。数据处理时共需要处理 9 个试飞科目, 共计 1 814 个处理参数, 多科目的名称、采样率、参数个数信息如表 1 所示。

表 1 多科目参数信息

序号	课题名称	采样率	参数个数
1	性能	32	234
2	品质	32	272
3	操稳	64	189
4	振动	256	78
5	雷达	16	127
6	航电	32	257
7	惯导	32	312
8	发动机	16	186
9	燃油	1	159

在多科目处理参数中, 气压高度是所有科目都关注的参数, 以该参数作为实验结果准确与否的判断依据。按照“合并—处理—分流”的方法处理多流多科目参数, 合并后的参数列表中共有 1 258 个处理参数, 每个参数的采样率不等。实时数据处理完成后, 在每个课题的结果文件中分析气压高度的数值, 发现所有课题的气压高度数值相同。气压高度处理结果如图 5 所示。

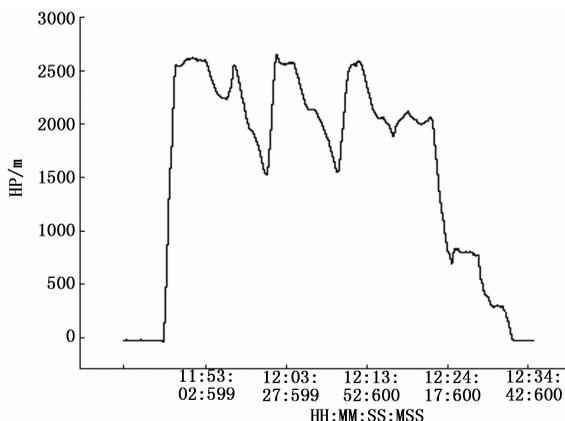


图 5 气压高度处理结果

(上接第 234 页)

[6] Parvathi P. Comparative Analysis of CBRP, AODV, DSDV Routing Protocols in Mobile Ad-hoc Networks [A]. 2012 International Conference on Computing, Communication and Applications [C]. 2012: 22-24.

[7] 郭鹏程. Ad Hoc 网络 AODV 按需路由协议的研究 [J]. 科学技术与工程, 2013, 13 (18): 5207-5213.

[8] 王军袖. Ad Hoc 网络中基于 AODV 路由协议能量优化算法的

由数据处理的结果可知, 多流多科目数据处理算法的处理过程正确, 对参数合并后能有效减少数据处理时的参数个数, 并且能够得到准确的处理结果。基于该算法设计的实时数据处理系统可应用于多个型号中, 有助于提高数据处理的效率, 加快试飞进度。

### 4 结束语

针对多流多科目数据处理的需求, 提出了一种基于“合并—处理—分流”方法的多流多科目实时数据处理算法, 首先将多科目数据处理参数参数名、采样率等信息进行合并, 形成一个合并处理参数列表, 该参数列表包含了处理参数的基本信息; 然后在实时处理时按照合并参数列表对网络数据流进行实时处理, 并将数据处理的结果存储在缓存中; 处理结束后通过书分流的办法获取处理结果, 并将结果写入课题需求的结果文件中。在试验总测试了该算法的有效性, 实验结果表明, 该方法能有效解决多流多科目数据处理的问题, 在得到准确数据处理结果的同时大幅提高数据处理的效率。

### 参考文献:

[1] 张卫方, 何晶靖, 阳劲松. 面向飞行器结构的健康监控技术研究现状 [J]. 航空制造技术, 2017, 60 (19): 38-47.

[2] 孙 伟. 直升机总体优化设计技术研究 [D]. 南京: 南京航空航天大学, 2012.

[3] 祁 春, 聂瑶佳, 周训强. 嵌入式新型机载实时数据处理系统的设计与实现 [J]. 计算机测量与控制, 2018, 26 (3): 208-211.

[4] 霍朝晖, 魏建新, 覃 燕. 飞行试验机载网络数据采集与分析技术 [J]. 现代电子技术, 2011, 34 (11): 144-147.

[5] 王建军, 彭国金, 陈惠君. 飞行试验多源 iNET 数据精密分析技术研究 [J]. 计算机测量与控制, 2016, (6): 248-250.

[6] 朱铁夫. 飞行试验图形监视系统设计 [D]. 西安: 西北工业大学, 2002.

[7] 赵晓阳. 基于 WEB 技术的飞行试验数据管理系统设计与实现 [D]. 成都: 电子科技大学, 2014.

[8] 彭国金. 基于类 PCM 结构过程的试飞多科目 iNET 处理技术 [J]. 计算机测量与控制, 2015, 23; 205 (10): 306-308.

[9] 滕 飞. 基于网络化的机载测试系统关键技术研究 [D]. 西安: 西安电子科技大学, 2013.

[10] 彭国金, 王建军, 刘嫚婷. 基于预留缓存的飞行试验 iNET 数据处理技术 [J]. 现代电子技术, 2016, 39 (12): 48-50.

[11] 房 瑾, 霍朝晖, 聂 睿. 机载网络数据实时处理软件的设计与实现 [J]. 电子设计工程, 2015, (12): 60-62.

[12] 宗思瑶. 机载实时数据处理系统设计与实现 [D]. 西安: 西安电子科技大学, 2013.

[13] 杨廷梧, 田宝泉. 飞行试验新型遥测机载网络化采集与记录系统架构 [J]. 测控技术, 2013, 32 (5): 59-63.

[9] 马 彬. 基于规则学习和朴素贝叶斯算法的软件缺陷预测模型研究 [D]. 长春: 吉林大学, 2019.

[10] 张 燕. 基于朴素贝叶斯的英语成绩预测模型构建研究 [J]. 自动化技术与应用, 2019, 38 (10): 67-69.

[11] 耿 倩. 贝叶斯算法在图书馆读者智能分析中的应用 [J]. 自动化技术与应用, 2018, 37 (5): 14-16.