

doi:10.3969/j.issn.1001-893x.2016.11.014

引用格式:刘相,刘泽平.可变报文格式编解码模块设计与实现[J].电讯技术,2016,56(11):1260-1266.[LIU Xiang,LIU Zeping.Design and im-
plementation of a variable message format coding and decoding module[J].Telecommunication Engineering,2016,56(11):1260-1266.]

可变报文格式编解码模块设计与实现*

刘 相**,刘泽平

(华北计算技术研究所,北京 100083)

摘 要:为适应现代战争数字化的军事需求,解决战场环境下需要传输的信息种类和数量日益增多以及网络带宽受限等问题,设计了可变报文格式(VMF)编解码模块,作为底层消息传送的技术支撑。利用动态链接库、JSON(Java Script Object Notation)、Snappy(压缩算法)等技术为上层应用提供消息的编解码服务,设计基于JSON的可变长报文模板、比特流集成压缩算法、基于JSON的数据字典结构,实现报文的“按需发送”,根据不同的发送需求,动态地组合报文内容。实验显示所设计的模块编码解码只需3 ms左右,而参考系统编解码则需耗时200 ms以上,可见速率提升明显,数据流长度也进一步缩短。

关键词:可变格式报文;编解码模块;JSON;数据字典;比特流集成压缩

中图分类号:TN919.3 **文献标志码:**A **文章编号:**1001-893X(2016)11-1260-07

Design and Implementation of a Variable Message
Format Coding and Decoding Module

LIU Xiang,LIU Zeping

(North China Institute of Computing Technology, Beijing 100083, China)

Abstract: In order to adapt to the military requirement of modern war digitization and solve the problems of increasing amounts and types of information need to be transmitted in the battlefield environment and constrained network bandwidth, a decoding module of variable message format(VMF) is designed as the technical support of underlying message. Using dynamic link library, Java Script object notation(JSON), Snappy and other technology, it can provide the service of message codec for the upper application. The design is based on the template of JSON, bit stream integrated compression algorithm and the data dictionary structure JSON, and then realizes the message sent "on demand" and the combination of dynamic message content according to the different delivery requirements. The experiment shows that coding and decoding only takes about 3 ms in the designed module, while the original decoding system takes over 200 ms. Obviously, it can increase the rate, and further shorten the length of the encoded data stream.

Key words: variable message format(VMF); coding and decoding module; Java script object notation(JSON); data dictionary; bit stream integrated compression

1 引 言

随着世界各国军事科技的快速发展,传统的通信和指挥控制方式已经难以满足指挥员对战场信息

获取和新型武器对信息系统的需求,传统的作战指挥信息的传输处理方式已不能满足信息的高实时性和安全性的要求。传统的作战指挥信息采用固定格

* 收稿日期:2016-07-22;修回日期:2016-09-08 Received date:2016-07-22;Revised date:2016-09-08
基金项目:国家部委专项资金资助项目(403060101)
Foundation Item:The Special Fund Project of National Ministries and Commissions(No.403060101)

** 通信作者:George_xl@ 163.Com Corresponding author:George_xl@ 163.Com

式报文进行传输, 固定格式报文内容均为预先定义的, 而作战过程中很多信息却是暂时不需要传输的, 报文在传输过程中会产生大量的数据冗余, 浪费了有限的网络带宽^[1]。可变长报文编解码技术的出现解决了传输过程中的数据冗余等问题。为了适应未来信息化战争的要求, 美军从 20 世纪 50 年代就开始研制使用战术数据链, 先后发展了 Link1、Link4A/4C、Link11、Link14、Link16、Link22 等数据链, 并从 16 号数据链使用的战术数字信息链技术中发展了可变报文格式 (Variable Message Format, VMF) 技术, 并于 2004 年正式发布为独立的 VMF 标准, 作为美军大力发展的数据链报文格式^[2]。

近年来, 我国有不少科研机构针对性地开展了报文编解码领域的理论技术研究, 并针对某类系统、某种特定目标建立了报文协议模型和数据字典原型^[3-6]。但为了应对未来更为复杂的战场环境, 在支持报文自由组报、平台迁移、报文快速解析、信息安全等方面, 仍需进一步研究。

本文在对报文协议进行充分归纳研究的基础上, 进行编解码模块设计, 将抽象化的战术数据报文标准的功能、语义、结构及语法规则等方面内容转换为结构良好的报文模板及计算机可识别的统一的数据结构, 作为与上层应用交互的数据接口; 再结合跨平台技术与加密机制, 设计实现了可变报文格式的编解码模块; 并通过多种类型的报文展示验证了编解码模块的应用效果, 证明了技术方法的可行性。

2 需求分析

VMF 作为一种主要应用于指挥控制系统中的信息格式标准, 保障指挥对象在带宽有限的战场环境能够安全实时地传输指挥控制信息。典型的编解码模块需要包括自由组报、计算机自动识别处理、可以进行面向比特的数字信息操作。因此, 指挥控制信息系统中的编解码模块应具备以下五方面能力:

(1) 自由组报的能力

能在组报时实现报文的可变长, 根据作战需求在组报的过程中自适应地增删报文数据字段, 确保发送的报文中不会有任何冗余的字段。

(2) 快速解析处理报文能力

在作战环境中, 能够通过高级程序语言快速解析报文结构, 处理报文信息, 具备在短时间内处理大量报文的能力, 使接收端不会出现大量报文囤积的情况。

(3) 信息安全能力

能基于数据元素字典和集成压缩算法, 对原始报文信息或数据进行两层加密, 确保在战场上传递的军事机密不会轻易地被敌方截获。

(4) 跨平台能力

具备灵巧的体系结构, 能用较小的空间存储报文结构以及数据元素字典的完整信息, 报文相关文件不需人工转换便能应用于多种系统或平台。

(5) 报文数据源扩展能力

具备开放的体系结构, 基于报文模板参数库, 能自适应地增加、修改想定的报文种类、相关参数。

3 系统方案设计

3.1 模块的体系结构

研究工作主要针对报文编解码的各种技术, 通过 VMF 报文通用描述格式、数据字典、集成压缩算法实现对可变长报文的编解码功能, 设置比特计数器实现对报文各字段的按位操作; 接收任意可变形式的报文进行解析, 将解析完成的结果存放在一个结构统一的双链表进行管理, 这样对报文的操作就转换成对双链表的操作, 方便对 JSON (Java Script Object Notation) 文件解析完成后对报文的后续操作, 并通过集成压缩算法对编码解码完成后的数据流进行处理, 进一步降低传输数据量。编解码流程如图 1 所示。

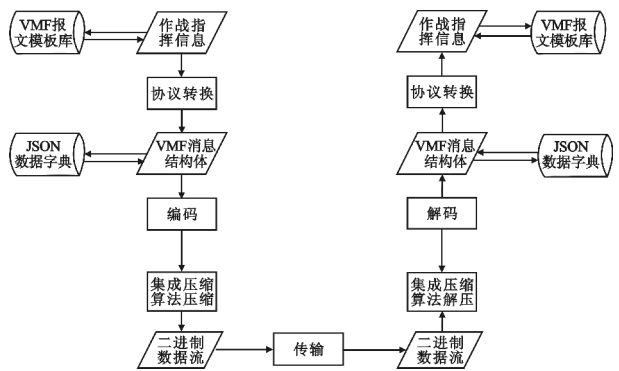


图 1 编解码流程
Fig. 1 Encoding and decoding process

编解码模块可以分成三部分实现: 一是使用数据结构将抽象的报文协议描述出来; 二是编解码程序的设计和实现, 将待编码的报文信息存储到数据结构中; 三是压缩算法在编解码模块内部的集成, 在接口内部封装 LZ4、Snappy、Zlib 3 种压缩算法, 计算机编码过程中自动选择具有最高压缩率的算法对报

文信息进行处理,压缩后的数据流长度进一步缩短,有效节省了传输带宽。

3.2 关键技术

3.2.1 基于 DLL 的通用型标准接口

报文编解码模块为上层应用提供标准动态链接库接口函数,上层应用只需要调用接口就可以完成对报文数据的编码和解码工作。在实际应用中,需对 VMF 协议进行分析、总结,从而提炼出统一的数据结构,如图 2 所示。

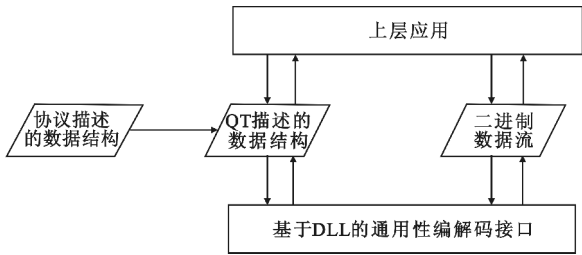


图 2 基于 DLL 接口的完整调用流程
Fig. 2 Complete call flow based on DLL interface

3.2.2 基于 JSON 的 VMF 消息结构设计

VMF 消息语法主要通过数据标识符 (DATA)、字段出现标识符 (FPI)、字段重复次数标识符 (FRT)、组出现标识符 (GPI)、组重复次数标识符 (GRT)来定义,VMF 根据包含、省略、重复字段、组 (即字段组合)等控制消息来实现消息的长度和格式的自由可变。

可变格式报文消息由若干个数据字段、指示符、组、重复组、组嵌套组成,因此,VMF 协议中应用的数据类型往往错综复杂,需要提供一种统一、支持多种数据类型的数据结构描述方式,这里采用 JSON 语言对报文进行描述。JSON 文档适合作为一种跨平台协议标准的基本格式。当对上层应用传来的数据进行编解码时,依据数据的功能类型在文件系统中查找对应的 JSON 文件,在 JSON 文件中的数据元素对报文进行组帧,完成后即可以进行比特编码操作。JSON 语言定义的数据元素标识使得编解码接口对于 VMF 数据结构解析速率更快,定义简洁高效,格式规范。

(1)DATA 基本数据元素

最基本的数据元素包括了字段名称、DFI、DUI 等属性,类似于下列结构:

```
{ "FIELD":  
  { "NAME" : "6013",  
    "DFI" : "4003",
```

```
"DUI" : "023",
```

```
}
```

```
}
```

在实际操作中通过唯一的 DFI/DUI 值实现对数据元素的快速索引。NAME 存储数据元素名称。

(2)FPI 数据域出现标识

```
{ "FIELD":  
  { "NAME" : "6013",  
    "DFI" : "4003",  
    "DUI" : "023",  
    "-minOccurs" : "0",  
    "-maxOccurs" : "1",  
  },  
}
```

属性 minOccurs 表示该数据元素最小出现次数,maxOccurs 表示该数据元素最大出现次数,这里的 minOccurs 只能为 0,maxOccurs 只能为 1。通过 minOccurs 与 maxOccurs 的组合实现该数据元素的省略或出现。

(3)GPI 组出现标识

```
{ "GROUP":  
  { "-NAME" : "坐标",  
    "-minOccurs" : "0",  
    "-maxOccurs" : "1",  
  },  
  .....}
```

包含组名称、组最小出现次数、组最大出现次数等属性,组内包含若干数据元素字段。

(4)FRT 数据域重复次数标识

```
{ "FIELD":  
  { "NAME" : "6013",  
    "DFI" : "4003",  
    "DUI" : "023",  
    "-minOccurs" : "0",  
    "-maxOccurs" : "100",  
  },  
}
```

区别于 FPI,FRT 中的 minOccurs 可以是小于等于 maxOccurs 的任意正整数,maxOccurs 为大于等于 0 的任意正整数。通过 FRT 可以实现对多个相同类型的数据元素在 VMF 消息中的增删组合。

(5)GRT 组重复次数标识

```
{ "GROUP":  
  { "-NAME" : "坐标",  
    "-minOccurs" : "0",  
    "-maxOccurs" : "100",  
  },  
  .....}
```

GRT 中的 minOccurs 和 maxOccurs 实现对多个相同类型的组在 VMF 消息中的组合。为了程序中报文解析操作,增加了组开始标志(BGI)和组结束标志(BFI)。

3.2.3 基于 JSON 的数据字典技术

针对报文解析过程中出现的数据存储无序、查询和调用不方便等问题,采用 JSON 文件来保存数据字典中对于数据模型的所有约束信息。

表 1 和表 2 为数据字典库中关于国家及地区的数据表。在此定义 1 个数据结构,将其数据结构名设为 CountryDic。DFI 的编号是“6013”,DUI 的编号是“001”,最大值以及最小值的变量为“unsigned short”型,这决定了消息字段的取值范围。步长分子(PACE_NUMERATOR)和步长分母(PACE_DENOMINATOR)为“long”型变量,两者决定了数据元素值的精度范围。

表 1 消息字典数据表
Tab.1 Message dictionary table

字段名称	举例
DFI	6013
DUI	001
NAME(名称)	国家
TYPE(类型)	unsignedchar
MAX_VALUE(最小值)	0
MIN_VALUE(最大值)	255
PACE_NUMERATOR(步长分子)	1
PACE_DENOMINATOR(步长分母)	1

表 2 编码字典表
Tab.2 Coding dictionary table

比特数值	域值名称
0	未知
1	中华人民共和国
2	美国
⋮	⋮
255	待定

采用 JSON 来描述“国家”数据元素:

```
{ "ELEMENT":  
{  
  "DFI":  
  { "ID" : " 6013 ",  
    "DUI" : {  
      "ID " : "001",  
      "TYPE" : "unsigned char",  
      " MIN_VALUE" : "0",
```

```
"MAX_VALUE" : "255",  
"PACE_NUMERATOR" : "1",  
"PACE_DENOMINATOR" : "1"  
}}}  
"VALUE" : {  
  "-NAME" : "国家",  
  "ID" : "0", "#text" : "未知",  
  "ID" : "1", "#text" : "中华人民共和国"  
}}}
```

(1)ELEMENT 对象:数据字典的根元素。作为 JSON 文档的第一层对象,“ELEMENT”包含“DFI”对象和“VALUE”对象,“DFI”元素是这个数据字典的最主要部分,它含有表 1 里面的所有信息,“VALUE”元素有表 2 里的所有信息。

(2)DUI 对象:第三层目录,包含了表 1 中除 DFI、DUI 之外的其他属性。”DUI”元素中字段 ID 为”001”时,域值名称为”国家”。

(3)TYPE 对象:包含了域值字段的数据类型,主要的数据类型有 char、unsigned char、int、double、float 等。

(4)MAX_VALUE 和 MIN_VALUE 对象:当数据元素的数据类型为字符类型时,MAX_VALUE 和 MIN_VALUE 用来表示数据元素字段的长度范围,单位为 b。

(5)PACE_NUMERATOR 对象和 PACE_DENOMINATOR 对象:有当数据元素字段数据类型为浮点数时,使用 PACE_NUMERATOR 和 PACE_DENOMINATOR,当需要表示的浮点数的精度是万分之一时,PACE_NUMERATO 的值为 1,PACE_DENOMINATOR 的值为 10 000,比特数值 1 用来表示 0.000 1。

(6)VALUE 对象:和 DUI 一样,都是第三层对象,表 2 的全部引用数据都被包括在里面,包括当数据元素字段作为编码字段时的编码值及其所表示的含义,并将表 1 中“名称”作为数据元素名称。

对于不同类型的元素编码的编码规则如下:

(1)当数据类型(TYPE)为整型(char、unsigned char、bool、short、unsigned short、int、unsigned int、long 等)时,PACE_NUMERATOR 和 PACE_DENOMINATOR 的值恒为 1。假定当前要编码的数值为 M,对于 M 可以进行如下表示:

$$M=MIN_VALUE+N,$$

(1)

$$N=d_{n-1}2^{n-1}+d_{n-2}2^{n-2}+\cdots+d_02^0。$$

(2)

式中:N 是 M 与该元素最小值的差;MIN_VALUE>0 时,N 要比 M 小,对 N 进行编码要比 M 短;d_{n-1}、d_{n-2}

... d_0 为转换后得到的二进制序列值。

(2) 当数据类型 (TYPE) 为浮点数类型时:

$$N = (M - MIN_VALUE) \cdot (PACE_NUMERATOR \cdot PACE_DENOMINATOR)^{-1}.$$

(3) 式中: M 为所需要编码的十进制数值; $PACE_NUM$ $RATOR$ 为步长分子; $PACE_DENOMINATOR$ 为步长分母; N 为待转换为二进制序列的值, 最后通过公式 (1) 将 N 转换为二进制序列。

举例: 假定某侦查站于 2016 年 6 月 18 日 15 时 30 分 30 秒, 在方位为东经 109°、北纬 32°发现异常目标, 通信车通过可变报文格式将这一消息实时传递给指挥部。根据上述编码规则编码后数据如表 3 所示。

表 3 报文数据结构实例
Tab. 3 Example of message data structure

字段名	长度/b	数值	二进制
GPI(时间)	1	1	1
年	12	2016	011111100000
月	4	6	0110
日	5	18	10010
时	6	15	001111
分	8	30	011110
秒	8	30	011110
GPI(坐标)	1	1	1
经度	9	109	100100001
纬度	9	32	011011110

3.2.4 编解码算法

编解码算法首先都需要通过 JSON 文件得到报文字段的 DFI、DUI 和名称, 在 JSON 数据字典中找到相应的字典元数据, 将解析文件得到的元数据全部存储到双链表节点中, 构造出来一个链表。双向链表中每个节点代表报文体中的一个数据元素, 链表节点内存有所有报文节点的信息。最终形成的双链表结构如图 3 所示。

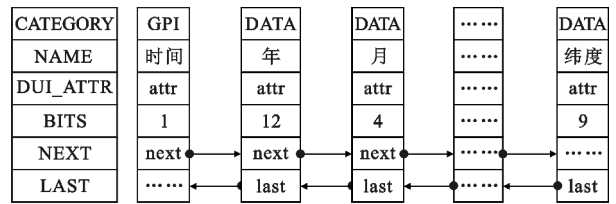


图 3 消息的链表结构
Fig. 3 Linked list structure of message

在编码算法中, 对存储有报文结构的双向链表

• 1264 •

进行遍历, 根据数据域的标识类型, 当标识类型为 DATA, 则从字节数组中取值后添加到比特数组; 若为 FPI, 则 FPI 的值为 1 时, 对下一个数据域编码, 否则跳过下一个数据域。GPI 的值为 1 时, 对下一组编码, 否则跳过该组。FRT、GRT 的值决定该数据域、组重复出现的次数。

3.2.5 集成压缩算法

压缩是在保证有用数据不损失的前提下, 缩减二进制数据量以达到传输过程节约网络带宽、提升传输效率的目的。压缩算法对于数据类型具有很强的依赖性, 同一种压缩算法对于不同的数据压缩后的压缩比可能差异很大, 针对这个问题, 我们提出在报文编解码系统中集成 3 种压缩算法, 让计算机自动选择出具有最佳效果的压缩算法, 算法流程如图 4 所示。本系统中集成了 Snappy (在谷歌内部被广泛使用的 C++解压缩包)、LZ4 (世界上速度最快的压缩工具之一) 和 Zlib (压缩率高的压缩包), 编码完成后的二进制数据流长度为 Length, 通过 Snappy、LZ4、Zlib 压缩后长度 Snappy_length、LZ4_length、Zlib_length, 计算机自动选择具有最优压缩效果的数据流进行传输。接收方收到发送方传输过来的数据流后选择对应的解压算法解压之后再交由编解码接口编码。此外, 压缩算法也作为一种加密手段提升了报文安全性。

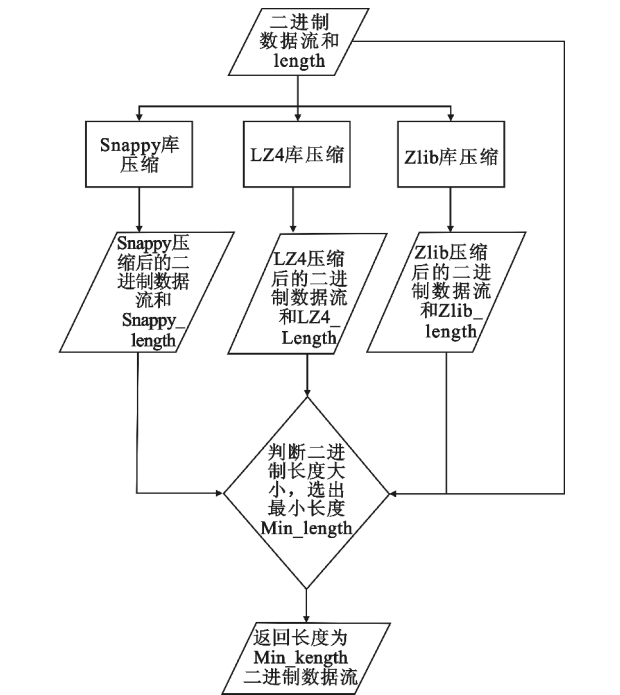


图 4 集成压缩算法流程
Fig. 4 Flow chart of the integrated compression algorithm

4 系统应用效果

4.1 实验方法与数据

定义了报文 4 (420x00x10) 想定模型: 报文长度为 420 B; 00 表示算术类型; 01 表示字符类型; 02 表示既有算术类型又有字符类型; 10 表示有重复数据; 11 表示无重复数据。

为进一步检验编解码模块在各种情况下数据处理效能, 针对报文长度、数据类型、有无重复数据设计了 9 种报文应用场景, 建立了报文 1 (210x00x10)、报文 2 (210x01x10)、报文 3 (210x02x10)、报文 4 (420x00x11)、报文 5 (420x01x11)、报文 6 (420x02x11)、报文 7 (420x00x11)、报文 8 (420x01x11) 和报文 9 (420x02x11) 等 9 个想定报文模型。

4.2 结果分析

实验采用与参考系统中进行对比的方法, 收集整理 9 组报文数据在不同系统的编解码模块中运行 1 000 次的编码和解码时间以及编解码后报文长度, 取最后的平均值作为实验结果, 如表 4 所示。

表 4 编解码时间与报文长度表

Tab. 4 Encoding and decoding time and message

报文序号	原始报文长度	编码时间/ms		解码时间/ms		编码后报文长度/B	
		参考系统	本系统	参考系统	本系统	参考系统	本系统
1	210	220	2.318	230	3.295	28	28
2	210	214	2.541	204	2.954	156	156
3	210	216	2.135	222	3.152	103	103
4	420	238	2.185	204	1.984	55	55
5	420	217	2.292	203	2.678	300	131
6	420	218	3.195	209	1.284	199	139
7	420	207	1.314	221	1.395	55	55
8	420	224	4.310	218	4.277	300	222
9	420	233	2.318	206	1.984	199	199

从图 5 的实验结果可知, 对于长度较短的报文 1、2、3, 编码后报文压缩比相同; 对于长度较长的报文, 参考编解码模块与本系统编解码对于算术类型报文编码效果接近于一致, 都具有极高的压缩效果。然而, 当两种编解码模块对相同的字符类型报文编码时, 参考编解码模块编码效果极差, 本系统编码后的压缩比比参考系统压缩比高出 40%。对比报文 5 和报文 8 的压缩比发现, 当报文中存在一定量重复值时, 本系统的编码解码效率更高。由此可知, 本系

统的编解码模块既延续了之前编解码模块擅长对算术类型报文的优点, 又在对字符型报文数据编码能力方面得到了增强。

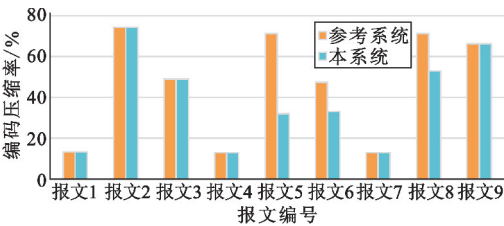


图 5 编码模块压缩率

Fig. 5 Compression ratio of coding module

从图 6 可知, 参考系统的编解码模块编码解码动辄耗时几百毫秒, 这也意味着原有的编解码模块 1 s 只能编码解码几条报文。可能会存在这样的情况: 几十个乃至上百个系统同时向一个系统发送报文消息, 由于报文的解析处理速率不够快, 将会导致接收端系统囤积大量待解析的报文。而改进后的编解码模块编解码耗时只需几毫秒, 意味着相同时间内能处理的报文数量变得更多, 能够有效缓解报文囤积的情况。

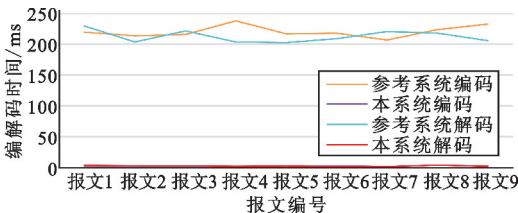


图 6 编码模块耗时

Fig. 6 Consuming time for encoding and decoding

5 结束语

基于 VMF 的报文编解码模块为 VMF 报文提供高性能标准化编码和解码方案, 解决了实际应用中报文数据结构定义、数据字典结构定义、双链表报文解析、报文数据源动态扩展等关键问题, 实现了报文的自由组报功能。本文提出了在报文编解码模块中集成压缩算法, 进一步减轻了多余报文字段造成传输带宽浪费的情况, 减少了传输时延, 对报文的使用也更加灵活方便, 按位操作报文, 而不是局限于字节, 同时也提高了报文信息的安全性和对报文的解析速率, 适合条件受限的环境下使用。相较于以往的编解码技术, 本文中的编解码技术进一步降低了报文长度, 但并不代表没有进一步提升解析效率的

空间,这一领域也将成为下一步研究的重点。

参考文献:

- [1] 张晓红. 基于 VMF 的作战指挥信息短报文实现方法[J]. 舰船电子工程,2008,28(7):66-68.
ZHANG Xiaohong. Implement method of short message for army command based on VMF[J]. Ship Electronic Engineering,2008,28(7):66-68. (in Chinese)
- [2] 杨猛,田畅. VMF 的发展前景展望[J]. 四川兵工学报,2009(1):111-112.
YANG Meng, TIAN Chang. The development prospect of VMF[J]. Journal of Sichuan Ordnance,2009(1):111-112. (in Chinese)
- [3] 罗捷,潘江峰,耿修堂. 某协议报文编解码模块的设计与实现[J]. 火炮发射与控制学报,2013(3):37-40.
LUO Jie, PAN Jiangfeng, GENG Xiutang. Design and realization of message coding & decoding module of some protocol[J]. Journal of Gun Launch & Control,2013(3):37-40. (in Chinese)
- [4] 廖龙灵,南建设. 体系结构设计技术及应用[J]. 电讯技术,2014,54(4):408-411.
LIAO Longling, NAN Jianshe. Techniques and applications of architecture design[J]. Telecommunication Engineering,2014,54(4):408-411. (in Chinese)
- [5] 耿腊元,毛玉泉,丁笑亮,等. 基于 VMF 的近距离空中支援应用研究[J]. 舰船电子工程,2009(7):9-12.

GENG Layuan, MAO Yuquan, DING Xiaoliang, et al. Research on close air support application based on variable format[J]. Ship Electronic Engineering,2009(7):9-12. (in Chinese)

- [6] 康瑞琪,葛利嘉,张振宇. 短波通信多组空时编码中等效 MIMO 模型 ZF 解码[J]. 电讯技术,2016,56(8):906-912.

KANG Ruiqi, GE Lijia, ZHANG Zhenyu. Effective MIMO model zero-forcing decoding for multi-group space-time coding in HF communications [J]. Telecommunication Engineering,2016,56(8):906-912. (in Chinese)

作者简介:



刘 相(1992—),男,四川南充人,2014 年获学士学位,现为硕士研究生,主要研究方向为体系结构设计及仿真;

LIU Xiang was born in Nanchong, Sichuan Province, in 1992. He received the B. S. degree in 2014. He is now a graduate student. His research concerns system structure design and simulation.

Email:George_xl@ 163. Com

刘泽平(1966—),男,重庆人,硕士,研究员,主要研究方向为军事信息系统总体及系统集成。

LIU Zeping was born in Chongqing, in 1966. He is now a senior engineer of professor with the M. S. degree. His research concerns military information systems and systems integration.