

GSM 长短信 PDU 编码分析

王国军

(甘肃兰州供电公司,兰州 730050)

摘 要:目前广泛应用的移动短信业务普遍地存在超过 70 个汉字分条发送的问题,给接收用户的阅读带来不便。文章详细分析了 GSM 长短信发送和接收的 PDU 编码方式,为开发者实现长短信收发提供实例指导。

关键词:GSM 长短信; PDU 编码;接收与合并

中图分类号:TP393, TN929.5 **文献标识码:**A

GSM Long-SMS PDU Coding Analyses

WANG Guo-jun

(Gansu Lanzhou Power Supply Company,Lanzhou 730050, China)

Abstract:Now widely used mobile messaging business commonly have the problem when the message more than 70 characters it will be splited more than one SMS to send , it bring inconvenience to the receiver's reading. This article particular analyzes the PDU coding of GSM long-SMS sending and receiving, to provide a instance guide for developers achieve long-SMS.

Key words: GSM Long-SMS; PDU Coding;Receive and Merger

随着现代通信技术的飞速发展,手机短信业务已在各行业广泛应用。但是由于 GSM 协议规定的短信体长度不超过 140 个字节,所以以前许多短信应用系统发送时均按单条长度 140 字节(70 个汉字)进行拆分,移动终端设备收到该短信后也是逐条显示。这样一来用户查看短信不方便,二来由于拆分后的短信经过短信网关、短信中心的存储转发后到达手机终端时可能会未按原发送的逻辑顺序接收到,使得接收方对于信息的理解出现偏差。尽管 1996 年颁布的 GSM03.40 v4.13.0 已支持长短信,但由于以前大部分手机终端不支持长短信从而限制了长短信的应用,考虑到 2005 年后生产的手机设备已普遍支持长短信的收发,各类应用系统支持长短信发送已很有必要。本文将结合作者在开发短信平台时对长短信协议的研究与短信控件开发升级,对长短信发送与接收时的编码进行实例分析。

1 长短信的定义

长短信即 GSM 03.40 中定义的 Concatenated Short Messages,由一组相互独立的不超过普通短信长度的子短信组成,在网络传输中被视为多条普通短信,而在终端上被合并显示——用户看到的是一条完整的长短

信。其实现方式是在短信报文中将 UDHI 设为 1,标识短信内容具有头结构,同时将用户数据区即短信内容中的前 6 个字节设置为长短信协议头(见表 1),标识各子短信之间的关系,以便接收端进行识别与合并。由于用户区共 140 个字节,此协议头占去 6 个字节后,只剩 134 个字节用于实际的短信内容,故长短信按 67 个汉字进行拆分。

表 1 长短信协议头格式

| 字节序号 | 含 义 | 说 明 |
|------|---|---------------------------|
| 1 | User Data Header Length 用户数据长度 | 05: 表示协议头的总长度,不包含本字节 |
| 2 | Information Element Identifier 信息单元标识 | 00: 代表用 0-0F 个字节进行编号的长短信。 |
| 3 | Length of the sub-header · 信息单元头的长度 | 03: 信息单元头的长度,不包含本字 |
| 4 | Reference Number 参考号 | 用于短信接收方区别不同短信的唯一标识 |
| 5 | Maximum number of short messages in the concatenated short message 长短信拆分后的条数 | 长短信拆分后的条数 |
| 6 | Sequence number of the current short message 当前短信的序列号 | 本条短信在此长短信中的第几条 |

收稿日期:2010-08-30

作者简介:王国军(1978-),男,本科,甘肃民勤人,主要从事负责软件系统维护,软件开发。

2 发送短信编码分析

在进行具体分析前,我们先对发送短信的 PDU 编码格式进行了解,见表 2:

表 2 发送短信时的 PDU 格式

| SCA | PDU-type (参见表 3) | MR | DA | PID | DCS | VP | UDL | UD |
|-------------------|---------------------|------|-------------------------------------|-----|-----|----|-----|----|
| PDU-type 的说明见表 3: | | | | | | | | |
| 表 3 PDU-type 说明 | | | | | | | | |
| 位 | 发送 | 接收 | 说明 | | | | | |
| 7 | RP | RP | 0:PDU 串中不带回复地址 1:PDU 串中带回复地址 | | | | | |
| 6 | UDHI | UDHI | 0:用户数据区只是短信 1:用户数据区开头带用户报头 | | | | | |
| 5 | SRR | SRI | 0:不需要/不会反馈状态报告 1:需要/会反馈状态报告 | | | | | |
| 4,3 | VPF | VPF | 设置 VP 域的格式,一般为 10,表示相对格式,VP 域占 1 字节 | | | | | |
| 2 | RD | MMS | MMS:长短信标志,0:普通短信,1:长短信 | | | | | |
| 1,0 | MTI | MTI | 短信类型:发送:01 接收:00 | | | | | |

下面通过一个具体的例子来进行编码的分析:将苏轼著名的词作《江城子·乙卯正月二十日夜记梦》通过短信发送,发送与手机号码均为:13139247510,通过作者的控件将发送与接收的 PDU 编码串保存为文本文件进行分析。

第一条短信:

051000D91683131297415F00008FF8C0500033902016C5F57CE5B5000B74E59536F6B6367084E8C534165E5591C8BB068A6002082CF8F7C002053415E74751F6B7B4E24832B832B30024E0D601D91CFFF0C81EA96BE5FD83002534391CC5B64575FFF0C65E059048BDD51C451C930027EB54F7F76F890225E944E0D8BC6FFDC5C186EE19762FF0C9B135982971C3002591C67655E7D68A65FFD8FD8

我们对照表 2 来看以上编码,见表 4。

第二条短信:

0051000D91683131297415F00008FF500500033902024E6130025C0F8F697A97FF0C6B6368B35986300276F8987E65E08A00FF0C60DF67096CEA5343884C300265995F975E745E7465AD80A05904FF0C660E6708591CFF0C77ED677E51883002

其中 051000D91683131297415F00008FF 同第一条,之后的 50 代表用户数据长度为 80 字节,0500033902 同

第一条,之后的 02 代表此短信是长短信的第 2 条,之后从 4E16 开始为用户数据区,代表短信内容“乡。小轩窗,正梳妆。相顾无言,惟有泪千行。料得年年断肠处,明月夜,短松冈。”

表 4 短信编码解析

| 含义 | 编码 | 说明 |
|---|--------------------|---|
| SCA (Service Center Address 短信中心地址) | 00 | 00: 表示使用 SIM 卡上设置的短信中心号码。目前的 GSM 网络在 SIM 卡注册时即将短信中心号码通知终端设备并保存于 SIM 卡,故一般按此即可,不需再每次指定短信中心 |
| PDU-type (PDU 类型) | 51 | 参见表 2 51 151H=0161001BQ 代表有报头短信 11 111H=0001001BQ 代表普通短信 |
| MR (Message Reference 消息参考) | 00 | 0-255 之间的十六进制表示,一般为 00H |
| DA (Destination Address 目标地址) | 91 | 0DH=13D:被叫号码长度。被叫号码长度的十六进制表示,对于本例来讲,不包括 '91' 与 'F'。 91: 目标号码是国际号码(带 + 号)如果是 81 则是本地号码。参见 GSM 07.08 |
| PID (Protocol Identifier 协议标识) | 00 | 00: 普通短信类型,点对点式 02: 长媒美醜錕的 0 0 呢通 GSM 美 |
| DCS (Data Coding Scheme 数据编码方案) | 08 | 08: UCS2,Unicode 编码,用于汉字; 00: 7bit,用于英文字符; 04: 8bit,用于图片等数据。 |
| VP (Validity Period 有效期) | FF | 短信有效期,在 PDU-Type 的 VPF 为 10 时需提供,FF 代表最长周期 |
| UDL (User Data Length 用户数据长度) | 8C | 此字节后的用户数据长度。8CH=140D 对于长短信,包含用户报头长度。 |
| UDHL (User Data Header Length 用户数据报头长度) | 05 | 代表后面的 5 个字节为报头 |
| Information Element Identifier | 00 | 标志这是个分拆短信 |
| Length of the sub-header | 03 | 代表后面还有 3 个字节的报头 |
| Reference Number | 39 | 如果手机同时收到不同号码发送的长短信,通过此标识号(再参考短信的时间戳)可以唯一区分并合并,实际编程中我采用系统时间的秒 + 10 的办法来生成此号 |
| Maximum Number | 02 | 代表此长短信共分 2 条发送 |
| Sequence Number | 01 | 代表此短信是本长短信的第 1 条 |
| UD (User Data) | 6C5F 57CE 5FD 8FD8 | 江城子·乙卯正月二十日夜记梦 苏轼 十年生死两茫茫,不思量,自难忘。千里孤坟,无处话凄凉。纵使相逢应不识,尘满面,鬓如霜。夜来幽梦忽还 |

3 接收短信编码分析

接受短信的 PDU 编码格式见表 4:

表 4 接收短信时的 PDU 格式

| SCA | PDU-type | OA | PID | DCS | SCTS | UDL | UD |
|--|----------|----|-----|-----|------|-----|----|
| 第一条短信: | | | | | | | |
| 0891683108901305F0440D91683131297415F0000 | | | | | | | |
| 8018091718352238C0500033902016C5F57CE5B5000B | | | | | | | |
| 74E59536F6B6367084E8C534165E5591C8BB068A600 | | | | | | | |
| 2082CF8F7C002053415E74751F6B7B4E24832B832B30 | | | | | | | |
| 024E0D601D91CFFF0C81EA96BE5FD83002534391CC | | | | | | | |
| 5B64575FFF0C65E059048BDD51C451C930027EB54F7 | | | | | | | |
| F76F890225E944E0D8BC6FF0C5C186EE19762FF0C9B | | | | | | | |
| 135982971C3002591C67655E7D68A65FFD8FD8 | | | | | | | |

我们参照表 4 来进行分析,见表 5:

表 5 接受短信分析

| 含义 | 编码 | 说明 |
|---|----------------|--|
| SCA(Service Center Address 短信中心地址) | 08 | 08H=16D 短信中心地址长度。((短信中心号码类型 + 短信中心号码长度)/2 的十六进制表示) |
| | 91 | 短信中心号码是国际号码(带 + 号) |
| | 683108901305F0 | 短信中心地址:+8613800931500 |
| PDU-type | 44 | 44H=01000100B,有协议头标志;普通短信一般为 04H=00000100B |
| | 0D | 来源地址长度 |
| OA(Originator Address 来源地址) | 91 | 来源地址格式为国际格式 |
| | 683131297415F0 | 来源号码为:+8613139247510 |
| PID | 00 | 普通 GSM 类型,点到点方式 |
| DCS | 08 | Unicode 编码 |
| SCTS(Service Center Time Stamp 短信中心时间戳) | 01809171835223 | 10-08-19 17:38:25 |

续表

| 含义 | 编码 | 说明 |
|--------------------------------|----|-----------------------|
| UDL | 8C | 用户信息长度(含用户报头)为 140 字节 |
| UDHL | 05 | 用户报头的长度(不包括本字节) |
| Information Element Identifier | 00 | 标志这是个分拆短信 |
| Length of the sub-header | 03 | 代表后面还有 3 个字节的报头 |
| Reference Number | 39 | 唯一标志(用于把两条短信合并) |
| Maximum number | 02 | 一共两条 |
| Sequence number | 01 | 本条是第一条 |

第二条短信:

0891683108901305F0440D91683131297415F0000 15日D00
 8018091718352238C0500033902016C5F57CE5B5000B 00025C
 7A97FF0C6B6368B35986300276E8087E65E08A801F0 5E08A0
 C60DF6709609E6A5B43884C308265995F976E745E7465 5E745E
 AD80A05904FF0C660E6708591CFFF0C977ED677E51883 677E5
 002

除表示长度的字节值 50(十进制 80),表示当前短信是第 2 条的字节 02 及具体短信内容(4E61 至结束)外,其余内容与第一条相同。

4 总结

短信发送相关的协议很多,但是没有实例看得清楚。通过本例实际分析,可以帮助开发者轻松地升级普通短信至长短信,提升短信管理系统功能,拓展系统应用范围,也方便手机用户查看短信。

参考文献:

- [1] GSM 03.38 Version 5.3.0[S], ETSI TC-SMG.
- [2] GSM 03.40 V7.3.0[S], ETSI TC-SMG.
- [3] GSM 07.05 Version 5.5.0[S], ETSI TC-SMG.
- [4] 3GPP TS 23.040 V9.2.0(2010-03)[S], <http://www.3gpp.org>.
- [5] SMS with the SMS PDU-mode[S], Siemens AG.

作者： 王国军 , WANG Guo-jun
作者单位： 甘肃兰州供电公司,兰州,730050
刊名： 电脑与信息技术
英文刊名： COMPUTER AND INFORMATION TECHNOLOGY
年,卷(期)： 2011,19(2)

参考文献(5条)

1. GSM 03.40 V7.3.0.ETSI TC-SMG
2. GSM 03.38 Version 5.3.0.ETSI TC-SMG
3. SMS with the SMS PDU-mode
4. 3GPP TS 23.040 V9,2.0 2010
5. GSM 07.05 Version 5.5.0.ETSI TC-SMG

本文读者也读过(10条)

1. 雷勇 PD分析与手机短信控件开发 [期刊论文]- 电力系统通信 2004,25(12)
2. 张文奇 . 段斌 . 王健 . 肖衡 Linux平台上基于PKCS#7的PKCS#7 Signed-data的实现 [期刊论文]- 计算机应用 2003,23(11)
3. 张灵敏 . ZHANG Ling-min 挖掘短信中的语文资源 [期刊论文]- 兰州教育学院学报 2010,26(4)
4. 刘敏 水文短信发布平台的研制与应用 [会议论文]-2007
5. 周丹 关于超长短信技术的研究探索 [会议论文]-2007
6. 关克 . 陈阳 . 宋柏 . GUAN KeCHEN YangSONG Bai基于AR的传感器校验仪GS数据传输的研究 [期刊论文]- 制造业自动化 2012,34(5)
7. 杨维永 . 林为民 . 陈亚东 . YANG Wei-yong LIN Wei-min . CHEN Ya-dongLinux系统下高性能加密系统框架研究与优化[期刊论文]- 计算机与现代化 2010(4)
8. 王进军 . 徐述 加密设备接口标准安全性分析 [期刊论文]- 计算机安全 2005(3)
9. 李远征 . 任传伦 . 吕慧勤 基于PK的Linux用户认证系统 [期刊论文]- 网络安全技术与应用 2004(5)
10. 李桂枝 . 韩江洪 . 刘小平 . LI Gui-zhi . HAN Jiang-hong . LIU Xiao-ping 嵌入式Mode在专线模式下的数据通信研究[期刊论文]- 计算机工程 2012,38(2)

本文链接：http://d.g.wanfangdata.com.cn/Periodical_dnyxxjs201102016.aspx