

ASN.1及其UPER编码在LTE中的应用研究

文小强¹ 李校林²

1. 重庆邮电大学 通信新技术应用研究所 重庆 400065

2. 重庆邮电大学 重庆信科设计有限公司 重庆 400065

0 引言

随着通信系统的不断发展,通信系统中协议结构的可选项、元素类型和扩展项等也变得越来越复杂。为了适应复杂数据结构传输的需要,ITU-T建议的X.680和X.681中定义了抽象语法描述(ASN.1)。国际标准8825定义了把ASN.1表示的数据结构值编码为适合传输的字节序列的转换语法。ASN.1描述时一般使用BER(basic encoding rules)或PER(packaged encoding rules)编码规则。同PER编码相比,BER编码因其在大小上的开销过大而受人诟病,和真实编码数据相比,平均需要增加50%的额外数据,也正是因为这个原因推动了PER的诞生。相同协议,PER编码与BER相比在大小上至少有40%~60%的改进^[1]。ITU-T在90年代初制订用于ASN.1的PER编码方式中,PER编码又分为对齐方式(aligned PER)和非对齐方式(unaligned PER)这2种情况。其中,非对齐方式的编码都是按比特来的,各个数据项的编码之间没有填充的比特,因此它比对齐方式的编码更为精简。本文采用非对齐方式编码方式并结合LTE层3消息中的无线资源控制(RRC)协议,详细进行了具体ASN.1描

述的分析及编码实现。

1 ASN.1概述

ASN.1是一种标记语言,它可以描述通信系统之间和各种应用程序之间要交换和传递的结构化信息,因为不同通信系统接收数据后数据重组的顺序不同,所以需要定义传送数据的标识方法,即传送语法。它是一种标准的、与具体的网络环境无关的语法格式。对于传送语法的要求有:必须能够定义各种复杂的类型,必须能够精确说明这些类型的值,需要提供一种以上的编码规则,这种编码规则能够确定会话层用何种比特模式来标识应用层的数据的值。

ASN.1可分为2个部分,即语法规则和编码规则。语法规则从数据类型、内容顺序或结构等方面来描述消息的内容,编码规则如何编码实际消息中的数据。

1.1 数据类型

根据标准X.680 ASN.1定义的数据结构类型分为简单类型和复合类型。基本类型是ASN.1描述数据结构的基础,复合类型是通过基本类型的组合而形成的较复杂类型^[2]。

基本类型包括 BOOLEAN, INTEGER, REAL, BIT STRING, OBJECT IDENTIFIER, SET, ENUMERATED, OBJECT

收稿日期: 2009-12-02

IDENTIFIER和 NULL等。BOOL是布尔量; INTEGER定义任意长度的整数, 相当于整数值的集合; REAL类型相当于实数值的集合; BIT STRING是以比特为单位的二进制字符串; OCTET STRING是以字节为单位的字符串; ENUMERATED是一个定义变量值的集合, 变量的值只在列举出来的范围内; OBJECT IDENTIFIER则用来唯一标识 ISO/ IUT定义的对象; NULL是位置符, 一般出现在 CHOICE SEQUENCE和 SET等复合类型中作为一项, 其本身没有意义, 不需要编码。

复合类型有 SEQUENCE SET CHOICE SEQUENCE OF和 SET OF等。其中 SEQUENCE和 SET类型可表示一组不同数据类型的集合, 相当于 C语言中的 struct 结构体; CHOICE类型是表示选择一组数据类型中的一个, 相当于 C语言中的 union 联合体; SEQUENCE OF和 SET OF类型定义了同一数据类型元素的集合, 相当于 C语言中的数组。二者的差别是 SEQUENCE OF内各项元素是有序的, 而 SET OF结构中各项元素的顺序是任意的。

1.2 数据表示

ASN 1描述中一般固有的数据类型用大写字母表示 (如 INTEGER SEQUENCE和 CHOICE等); 用户定义的数据类型以大写字母开头, 但是至少含有一个非大写字母 (如 DRB-Identify等); 标识符可以包括大写和小写字母、数字和下划线, 但以小写字母开头。

1) ASN 1的类型定义, 其语法为: $\langle \text{type name} \rangle:: = \langle \text{type} \rangle$ 。例如: $\text{NumberOfReportsSent}:: = \text{INTEGER}$ $\text{DedicatedInfnAS}:: = \text{OCTET STRING}$ 等。

2) ASN 1的子类型定义。ASN 1可以定义一种子类型, 将变量限制为特定值或特定范围。其语法为: $\langle \text{subtype name} \rangle:: = \langle \text{type} \rangle (\langle \text{constraint} \rangle)$ 。例如 $\text{RRC-TransactionIdentifier}:: = \text{INTEGER} (0..3)$ 。

3) ASN 1的赋值。其语法为: $\langle \text{value name} \rangle:: = \langle \text{value} \rangle$ 。例如 $\text{num NumberOfReportsSent}:: = 30$ $\text{dedicatedInfnASone DedicatedInfnAS}:: = \text{'FFFF0000'}$ H等。复合类型 SEQUENCE可定义实例: $\text{UserAccount}:: = \text{SEQUENCE} \{ \text{name PrintableString password PrintableString};$ 其可以赋值为 $\text{jessicaAccountUserAccount}:: = \{ \text{name "jessica",}$

$\text{password "admin"} \}$; 复合类型 SEQUENCE OF可定义实例: $\text{Students}:: = \text{SEQUENCE OF PrintableString}$ 其可赋值为 $\text{centralClass Students}:: = \{ \text{"Lilley", "Hameiney", "Yuyang"} \}$, 其中 PrintableString为字符串类型。

2 LTE中层三消息及其编码规则

LTE是 3G的下一代演进技术, 是近 2年来 3GPP启动的最大的新技术研发项目, 这种以 OFDM/FDMA为核心的技术可以被看作“准 4G”技术。与 3G相比, LTE更具技术优势, 具体体现在高数据速率、分组传送、延迟降低、广域覆盖和向下兼容等方面。

在 LTE中使用 ASN 1的协议有 E-UTRA无线资源控制 (RRC)协议, E-UTRAN S1应用协议和 E-UTRAN X2应用协议。E-UTRAN S1应用协议和 E-UTRAN X2应用协议使用 BASIC-PER编码规则, E-UTRA无线资源控制 (RRC)协议使用 UPER编码规则^[9], 并主要使用了 ASN 1的 BOOL INTEGER BIT STRING OCTET STRING ENUMERATED NULL SEQUENCE CHOICE和 SEQUENCE OF类型。

3 LTE中的 UPER编码

非对齐紧缩编码 (UPER)规则包括 3个部分: Preamble(前缀)、length(长度)和 contents(内容)。其编码格式如图 1所示。Preamble它只出现在 ENUMERATED SEQUENCE SET CHOICE这 4种数据结构的编码中。用来记录结构中是否有扩展项 (extension)、选择项 (optional)或缺省项 (default), length对 STRING SEQUENCE OF SET OF等数据类型进行编码时, 需对其长度按非对齐方式进行编码; contents若数据是基本结构类型, 如 BOOLEAN INTEGER REAL和 BIT STRING等, 即可直接进行编码; 若数据是符合结构类型, 如 SEQUENCE SET CHOICE和 SEQUENCE OF等, 则属嵌套编码。此时 contents中也包含有 Preamble length和 contents 3个部分。

PER的编码原则就是用尽可能简单的规则得到最简洁的编码, 因此非对齐紧缩编码 (UPER)与对齐紧缩编码 (PER)一个明显的不同在于: PER方式下, 在数据编码时, 需要在要求对齐的数据前插

“0”以达到保持 8 位组对齐的目的,而 UPER 方式则编码更为紧凑,不会检查 8 位组的对齐情况,只在整个数据编码结束后,才进行补位。但是 PER 的使用必须有一个前提,那就是网络通信的收发双方都清楚正在使用的用 ASN.1 句法描述的数据结构,当然这只需要收发双方都遵循统一的通信协议即可。UPER 编码结构如图 1 所示。

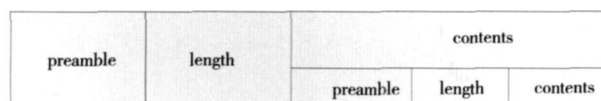


图 1 UPER 编码结构

与 LTE 层三消息密切相关的数据类型编码方法如下。

1) 范围受限的整数类型 INTEGER (lb... ub)。UPER 编码规则采取的是对数据的偏移值 (offset) 进行编码。所谓偏移值就是实际值减去下界的值。即: lb 为范围的下限, ub 为上限,由 UPER 可视约束给定,范围 $N = ub - lb + 1$ 。则需要的值为 "N", 实际进行编码的值为 (N - lb)。

例如: 当 ASN.1 描述的数据 MSB-integer INTEGER (0... 9) 的值为 4 时, UPER 编码为 0100。INTEGER (0... 255) 的值为 16 时, UPER 编码为 00010000。而当 maxCD INTEGER (1... 16383) 的值为 1000 时, UPER 编码是 00010010101111。

2) BOOLEAN 类型。布尔型编码非常简单,只需 1 bit 即可。用“0”表示 FALSE,“1”表示 TRUE。

3) BIT STRING 类型。比特串的编码分为长度固定和长度可变 2 种情况。

① 长度固定的编码: 直接将比特串与前面编码的比特流级联即可, LTE 层三消息的只有长度固定比特串这种情形。

② 长度可变的编码: 按前述整形编码的规则,以描述长度范围最少比特的原则在比特串的前面加上用于描述长度所需的比特,然后将待编码的比特串加在后边。

例如:

CellIdentity ::= BIT STRING (SIZE (28))

若 CellIdentity 的值为全 1, 则被编码为 '11111111 11111111 11111111 1111'。实际编码时将这一串编码比特级联到比特流中即可。

4) ENUMERATED 类型。ENUMERATED 类型的编码类似于整形,以描述枚举内全体成员所需比特数最少的原则进行编码。且 ENUMERATED 类型中的第 1 个成员类型名称对应的常数下标从 0 开始,第 2 个成员类型名称对应的常数下标为 1,依次编号。

例如:

ReestablishmentCause ::= ENUMERATED {
reconfigurationFailure, handoverFailure,
otherFailure, spare}

若 ReestablishmentCause 的值取其中的 handoverFailure, 则被编码为 01, 其中 ReestablishmentCause 类型里面有 4 个成员类型名称, 需要用 2 个比特表示, 而 handoverFailure 处于第 2 个成员类型名称, 下标为 1, 即编码为 01。

5) CHOICE 复合类型。CHOICE 结构中有许多元素, 每一个元素均有一个索引号, 第 1 个元素的索引号为 0, 第 2 个元素的索引号为 1, 依次类推, 直到最后一个元素 N。编码时首先判断在 CHOICE 结构内是否有扩展标记 (...), 如果选择的是扩展标记前面的项, 则在 preamble 前增加一个比特位 '0' 表示所选项在原列表内; 否则, 增加一个比特位 '1' 表示所选为扩展内容。接着按索引号 + 所选元素编码方式编码。即编码方法为: preamble + 索引号 + 所选元素编码。

例如:

ThresholdUTRA ::= CHOICE {
utra-RSCP INTEGER (-5..91),
utra-EcNo INTEGER (0..49),
...
}

若 ThresholdUTRA 的取值如下

{
utra-RSCP - 1
}

则编码表示为 000000100, 其中第 1 个 0 表示所选项在原列表内, 第 2 个 0 表示选择了 CHOICE 中的第一个元素 utra-RSCP 接着的 7 bit 即 0000100 表示 utra-RSCP 的值为 -1。

6) SEQUENCE 复合类型。SEQUENCE 结构中可能有扩展标记 (...), 也有可能有些元素为可选项

或者默认项。编码时首先判断在 CHOICE结构内是否有扩展标记(…),如果选择的是扩展标记前面的项,则在 preamble前增加一个比特位'0'表示所选项在原列表内;否则,增加一个比特位'1'表示所选为扩展内容;其次看 SEQUENCE结构内是否有可选项(OPTIONAL)或者默认项(DEFAULT),若有(假设个数为'n'),则添加 n个比特到 preamble中,这 n个比特中的第 1个比特对应第 1个可选项或者默认项;最后 1个比特对应最后 1个可选项或者默认项,若对应比特的值为 0则表示该 SEQUENCE中不包括该可选项或者默认项。若对应比特为 1,则表示该 SEQUENCE中包括该可选项或者默认项。编码顺序为:先编扩展项,后编可选项或默认项,接着是结构中各个元素的内容编码。

例如:

```
DRB-ToAddMod ::= SEQUENCE {
    epsBearerId INTEGER (0..15) OPTIONAL,
    logicalChannelId INTEGER (3..10) OPTIONAL,
    logicalChannelConfig LOGICALCHANNELCONFIG OPTIONAL,
    ...
}
```

若 DRB-ToAddMod的取值如下

```
{
    epsBearerId 5
    logicalChannelId 6
}
```

则编码表示为 01100101011。其中第 1个 0表示无扩展项;第 2个 1表示 epsBearerId选用;第 3个 1表示 logicalChannelId选用;第 4个 0表示 logicalChannelConfig不选用;接着的 0101表示 epsBearerId的值为 3最后的 011表示 logicalChannelId的值为 6。

7) SEQUENCE SIZE(lb, ub) OF复合类型。其编码形式为:成员数目编码+各个成员内容编码;成员数目按受限整数方式进行编码。

例如:

```
MSI ::= SEQUENCE (SIZE (6..21)) OF MSIDigit
MSIDigit ::= INTEGER (0..9)
```

若 MSI的取值为 SEQUENCE(8) OF 8时,则编码为: 0010 10001000 10001000 10001000 10001000 其中前面 4个比特 0010表示此时 SIZE里面取的范围为 8即一共有 8个数组。接着的第 1个 4个比特 1000表示第 1个数组内容为 8。接下来的第 2个 4个比特 1000表示第 2个数组的内容为 8同上,直至第 8个 4个比特 1000表示第 8个数组的内容为 8。

4 总结

由于基于 ASN.1描述的非对齐紧缩编码规则(UPER)具有效率高、占用带宽小的特点,所以 LTE^[3]协议栈采用了 ASN.1描述,并在层三消息使用了 ASN.1的 UPER编码方式,这也意味着 ASN.1特别适合于网络中应用的那些复杂的、可变化的及可扩展的数据结构,它将在 OS框架结合和 3GPP中得到更广泛的应用^[2]。本文介绍了抽象语法描述(ASN.1)的概念和数据表示方式,并针对 LTE层三消息所用到的数据结构,详细举例实现了 UPER的编码过程。

参考文献:

- [1] 吕谦,黄本雄. ASN.1及其两种编码方式(BER和PER)的对比研究[J]. 数据通信, 2003(3): 18-21.
- [2] 李小文,李贵勇,陈贤亮. TD-SCDMA第三代移动通信系统、信令及实现[M]. 北京:人民邮电出版社, 2003: 243-250.
- [3] 徐昌彪,武岳,鲜永菊. LTE-Advanced系统中改进的载波聚合 HARQ进程映射机制[J]. 重庆邮电大学学报:自然科学版, 2010, 22(2): 170-173, 213.

作者简介:

文小强(1985—),男,湖北荆门人,硕士研究生,主要研究方向为 LTE协议栈, E-mail: mazing27@163.com.