

文章编号:1001-9081(2010)04-0881-03

## 基于 SigComp 的会话初始化协议压缩算法

李秉智, 杨富良

(重庆邮电大学 计算机科学与技术学院, 重庆 400065)

(yangfl\_cqupt@163.com)

**摘要:**独立媒体服务(IMS)是第三代合作伙伴计划(3GPP)在版本5中提出的支持IP多媒体业务的子系统,而基于文本的SIP消息过大成为其在IMS无线环境应用下的瓶颈,因此采用会话初始化协议(SIP)来建立和维护多媒体会话。在SigComp框架下,将改进后的LZSS算法与算术编码相结合对SIP信令进行压缩。实验表明新的算法有较高的压缩率,对改善IMS的SIP会话建立延时有一定的参考价值。

**关键词:**会话初始化协议; LZSS算法; 算术编码; 信令压缩

**中图分类号:** TP391 **文献标志码:** A

## SIP compression algorithm based on SigComp

LI Bing-zhi, YANG Fu-liang

(College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** Independent Media Services (IMS) is a subsystem to support IP Multimedia services, working in the 3rd Generation Partnership Project (3GPP) of version 5. IMS uses Session Initiation Protocol (SIP) to establish and maintain multi-media session; however, oversized SIP message based on text is becoming a bottleneck under IMS wireless circumstances. Under the framework of SigComp, a new algorithm based on improved Lempel-Ziv-Storer-Szymanski (LZSS) algorithm and arithmetic combination was presented to compress SIP signaling. Simulation results prove that the new algorithm has higher compression efficiency. It is valuable for IMS SIP session setup delay.

**Key words:** Session Initiation Protocol (SIP); Lempel-Ziv-Storer-Szymanski (LZSS) algorithm; arithmetic coding; signaling compress

### 0 引言

会话初始化协议<sup>[1]</sup> (Session Initiation Protocol, SIP) 是由 IETF 组织于 1999 年提出的一个在基于 IP 网络中实现实时通信的一种应用层控制(信令)协议,用于在 IP 网络上建立、修改以及终止多媒体会话或呼叫。由于其具有易理解、可扩展、使用灵活、对移动性的支持等特点,使得第三代合作伙伴计划(3rd Generation Partnership Project, 3GPP)在版本 5 中采用了 SIP 作为会话控制协议来设计独立媒体服务(Independent Media Services, IMS)。然而,与 HTTP 类似, SIP 是一种基于文本的设计用于带宽丰富的有线链路的协议,消息体有几百到几千字节甚至更大<sup>[2]</sup>,所以当 SIP 用于带宽资源有限、速率较低的无线网络时,庞大的消息造成了会话建立时延的增加,最大延时可达 20 s 左右,对于实时通信应用来说是无法容忍的。本文在 SigComp 的框架下,根据 SIP 消息的特点,对无损压缩效果较好的 LZSS 算法进行改进,并结合算术编码进一步提高算法的压缩效果,以降低传输延时,缩短 SIP 会话建立时间。

为了解决一些基于文本的应用协议(如 SIP、RTSP 等)消息过大而引起的会话建立延时问题, IETF 提出了 SigComp 机制,用健壮无损的压缩算法来减小协议消息,以改善时延。SigComp 的核心通用解压虚拟机(Universal Decompression Virtual Machine, UDVM)类似于 Java 虚拟机,不同的是它独立于各种不同的压缩算法,通过解压指令它能解析任意压缩部分的压缩算法,本文就是对其压缩算法进行研究,提出一种压缩效果较好的 SIP 协议压缩算法。

SigComp 是介于应用层与传输层之间的一个新层,主要由压缩程序、UDVM、压缩分发器、解压分发器和状态控制器等组成,体系结构如图 1 所示。

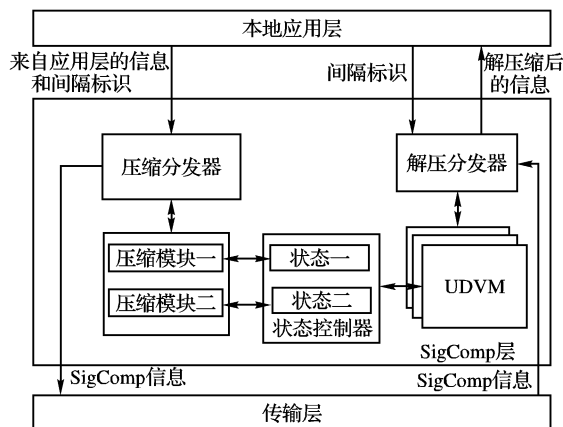


图1 SigComp 体系结构

工作流程如下:

- 1) 压缩分发器作为应用程序和传输层的接口,负责从应用层接受消息,根据间隔表示符调用相应的压缩程序,并将压缩好的 SigComp 消息发送到传输层;
- 2) 解压分发器从传输层接收 SigComp 消息,根据附带信息调用 UDVM 解压,并将解压后的消息发送给应用程序;
- 3) 状态是为 SigComp 消息检索恢复而保存的数据,状态控制器在获得应用程序批准后存取这些数据实体。

收稿日期:2009-10-12;修回日期:2009-12-07。

作者简介:李秉智(1946-),男,河北唐山人,教授,博士生导师,主要研究方向:三网合一、计算机网络与通信; 杨富良(1985-),男,福建三明人,硕士研究生,主要研究方向:计算机网络与通信。

## 1 无损压缩算法及 LZSS 的选择

压缩算法大体上可以分为两类:有损压缩和无损压缩。所谓的有损压缩是利用了人类对图像或声波中的某些频率成分不敏感的特性,允许压缩过程中损失一定的信息以换取更大的压缩比。无损压缩是利用数据的统计冗余进行压缩,可完全恢复原始数据而不引起任何失真,经常使用的无损压缩方法有 Huffman 编码、游程(Run-length)编码、LZ(系列)编码和算术编码等。

对于 SIP 信令的压缩要求是还原后与原来信息完全一致,只能采用无损压缩方法,基于字典的无损压缩算法对文本信息具有压缩速度快、效果较好等特点,而根据文献[3]的结论,LZ 系列的表现较优。其中 LZSS 和 LZW 是 2 个比较实用的算法。LZW 是 LZ78 的改进,它通过初始化学字典来包含输入字符表中的所有字符,然后对短语进行编码,输出的是短语在字典中的索引。LZSS 则是在窗口匹配模式和输出编码的结构方面对 LZ77 进行了改进。LZSS 算法建立了一棵二叉搜索树,如果短语在前向缓冲区中处理完毕并移入窗口时,就将该短语在窗口中的偏移量添加到二叉搜索树中<sup>[4]</sup>,这样能让查找匹配的时间复杂度由 LZ77 中的窗口长度  $n$  降为  $\lg n$ 。LZ77 的输出编码由偏移量、匹配长度和紧跟在短语后面的字符 3 部分组成,而 LZSS 则允许编码和未能匹配的原始字节随意混合在一起,使用 1 个比特作为标志位,标明输出的是〈距离,长度〉码,还是原始字节。例如对字符串“123123...”具体的编码过程如图 2(a)所示。

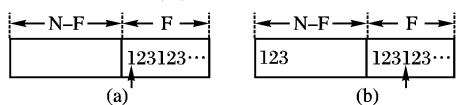


图2 LZSS 编码过程

开始,滑动窗口 N-F 为空,编码位置为向前缓冲区 input = “123123...”的起始位置 1,在 N-F 的窗口中没有找到匹配,则输出标识位 0(标识输出的为未经压缩单个符号)和相应的字符:output = “0 00000001”,指针向前移动一位。同理对于 input 中的字符“2”和“3”得到相应的 output = “0 00000001 0 00000010 0 00000011”,此时指针指向“1”,如图 2(b)所示。

找到了相应的 N-F 窗中的匹配,且匹配字符串长度为 3,大于最小匹配 THRESHOLD(假设为 2)长度,则输出标识位 1(标识输出为〈距离,长度〉对),此时 output = “0 00000001 0 00000010 0 00000011 1 0011 011”,同时指针指向匹配字符串后一位。如此循环,直到 input 中没有未编码的字符为止。

分析 SIP 消息的特点可知,SIP 消息的大小一般只有几百到几千字节且都由 ASCII 组成,格式比较固定,对于类似“SIP”、“Call-ID”等指令短语出现的频率比较高,某些字段在会话建立过程中是不变的,例如“Call-ID”等。而 LZW 算法和 LZSS 算法的差异主要表现在两个方面:1) LZW 算法的压缩速度要优于 LZSS,而对于传输时间几千毫秒来说,压缩/解压上的几毫秒到几十毫秒几乎可以忽略不计<sup>[5]</sup>;2) LZSS 的压缩比要高于 LZW,特别对于 SIP 这类短文本消息,优势更加突出。因为 LZSS 的自适应速度要明显优于 LZW,例如,对于 SIP 消息中经常出现的短语“branch”,LZW 需要 6 次后才能使整个短语编入字典,而 LZSS 在第二次出现词短语的时候就能完全拷贝。因此,本文选用 LZSS 算法并加以改进,以取得更好的压缩效果。

## 2 对 LZSS 的改进

虽然 LZSS 对 SIP 消息的压缩有其优势,但仍有可以改进的地方。

1) 对 LZSS 中的滑动窗口事先加载字符。对于 LZSS 的滑动窗口,开始时是空的,这样不利于压缩过程开始阶段压缩,特别是对于 SIP 这类短文本信息,对压缩效果影响较大。所以,可以预先加载部分字符到窗口中<sup>[6]</sup>,使得压缩程序在开始时就能够找到匹配字符进行压缩,取得较好的压缩效果。

2) 对 LZSS 滑动窗口大小的设定。对于上下文相关度越低的文件,相应选择较大的滑动窗口能取得较好的压缩比,但也相应增加了匹配时间与编码长度。对于 SIP 这类型的文本来说,本文采用 4 KB 的滑动窗口,其偏移值基本上是均匀分布的,并且可以用固定的位数来表示。实践证明,这个窗口大小是合适的。

3) LZSS 与算术编码相结合。使用基于字典编码的 LZSS 算法虽然速度较快,但压缩比并不是十分理想,而在一个压缩系统中,熵编码往往是必需的环节,因此可以考虑将 LZSS 与熵编码相结合。而最常用的两种熵编码是 Huffman 编码和算术编码。算术编码是一种到目前为止编码效率最高的统计熵编码方法,它比 Huffman 编码效率提高 10% 左右,而且可以以分数比特逼近信源的熵,统计模型可以与算法很好地分离,易于实现自适应模式。因此本文中采用 LZSS 与算术编码相结合的方法来压缩 SIP 消息。

算数编码的编码过程如下所示。首先将当前区间  $[L, H)$  初始化为  $[0, 1)$ 。其次对于文件中的每一个事件做如下处理:1) 将当前区间细分为一些子区间,一个事件子区间的大小同这个事件在这个文件中再一次出现的估计概率成正比;2) 选择下一个将发生事件的相应子区间,并使它成为新的当前区间;3) 输出足够的比特位,使得当前区间同所有其他的可能最后区间区分开来;4) 子区间的长度等于各个事件的概率乘积,并且需要一些方法指示文件末端。

在处理文中的每一个事件时,只需要计算实际发生事件的相关子区间,使用两个“累加”概率即可:

$$\text{累加概率 } P_x = \sum_{m=1}^{i-1} p_m$$

$$\text{下一次累加概率 } P_y = P_x + p_i = \sum_{m=1}^i p_m$$

由此,新的区间为:  $[L + P_x(H - L), L + P_y(H - L))$ 。

改进后的 LZSS 与算术编码相结合后的编码流程如图 3 所示。

## 3 改进的 LZSS 对 SIP 信令压缩仿真结果

为了对改进的 LZSS 算法的压缩效果进行验证,本文采用了如图 4 所示的典型的 SIP 会话环境<sup>[7]</sup>,对 23 条 SIP 信令消息进行压缩。

为了比较,实验分别采用了 LZW、LZSS 和改进后的 LZSS 三种算法对相同的 SIP 消息进行压缩,定义  $n$  个 SIP 消息的压缩率,为对其性能进行比较,如图 5 所示。

$$\text{Ratio}_n = \frac{n \text{ 个消息压缩后大小}}{n \text{ 个消息总大小}} \quad (1)$$

仿真结果分析:随着会话建立过程的延续,SIP 消息个数的增加,式(1)中定义的消息压缩率不断降低;由曲线的斜率可知,LZSS 的收敛速度快于 LZW;改进后的 LZSS,由于采用了熵编码,较 LZSS 和 LZW 算法压缩率有很大提高,改进后的 LZSS 算法的压缩率接近 LZSS 压缩率的一半。因此,改进后

的 LZSS 适合于 SIP 消息的压缩,有利于 SIP 在不同带宽场合

下的应用。

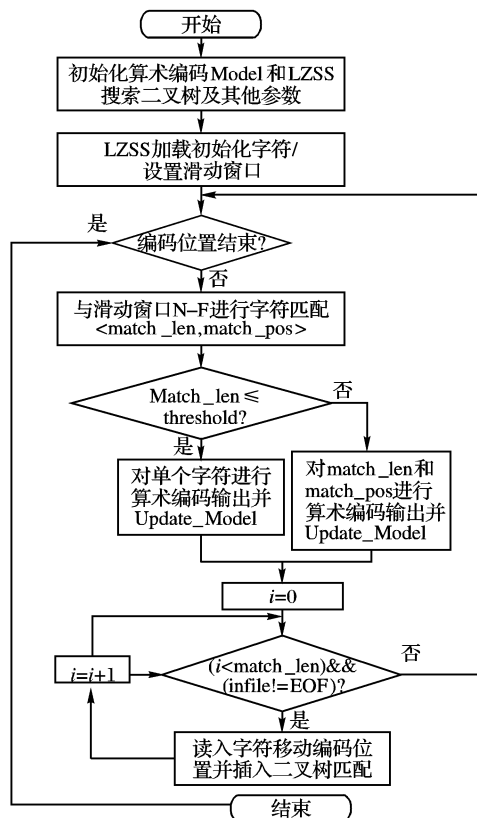


图3 改进后的 LZSS 编码流程

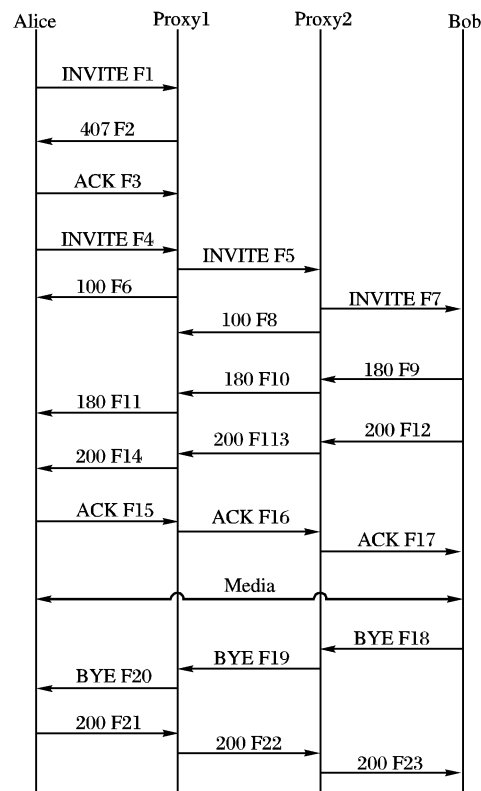


图4 典型 SIP 会话建立过程

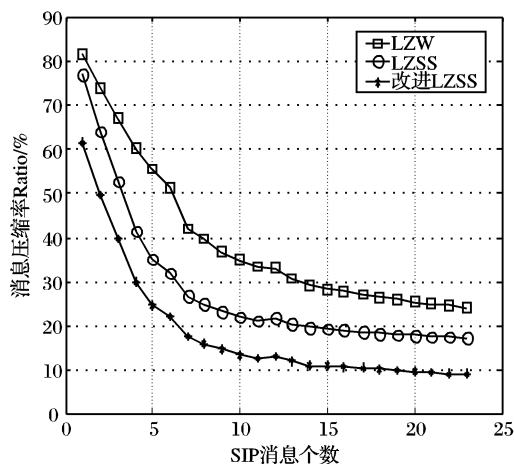


图5 SIP 消息压缩率曲线

## 4 结语

本文采用改进的 LZSS 算法对 SIP 消息进行压缩,取得了较好的压缩比。但是这种方法仍有许多可以改进的地方,例如,对 LZSS 的码值存储可以进行优化处理,对算术编码的复杂过程可以做进一步简化等。

## 参考文献:

- [1] ROSENBERG J, SCHULZRINNE H, CAMARILLO G, *et al.* RFC3261, SIP: Session initiation protocol [S]. 2002.
- [2] PRICE R, BORMANN C, CHRISTOFFERSSON J, *et al.* RFC3320, Signaling compression (SigComp) [S]. 2003.
- [3] JIN H, NAHENDRAN A C. Using sigcomp to compress SIP/SDP message [C]// IEEE International Conference on Communications. [S. l.]: IEEE, 2005: 3107 - 3111.
- [4] WITREN I H, BELL T C. The zero-frequency problem: Estimating the probabilities of novel events in adaptive text compression [J]. IEEE Transaction on Information Theory, 1991, 37 (4): 1085 - 1094.
- [5] DAI BIN, WANG FU-RONG, KE JIN-SHUI. Performance analysis of signaling using SigComp scheme in narrowband system [C]// IEEE Consumer Communications and Networking Conference. [S. l.]: IEEE, 2006, 1: 376 - 379.
- [6] 王忠效, 姜丹. Lempel-Ziv 77 压缩算法及其实现的研究 [J]. 计算机研究与发展, 1996, 33(5): 1 - 4.
- [7] JOHNSTON A, DONOVAN S, SPARKS R. RFC3665, Session initiation protocol (SIP) basic call flow examples [S]. 2003.

(上接第 880 页)

- [8] DEB B, BHATNAGAR S, NATH B. ReInForm: Reliable information forwarding using multiple paths in sensor networks [C]// Proceedings of IEEE International Conference on Local Computer Networks. Washington, DC: IEEE Computer Society, 2003: 406 - 415.
- [9] FELEMBAN E, LEE C G, EKICI E. MMSPEED: Multi-path multi-speed protocol for QoS guarantee of reliability and timeliness in wireless sensor networks [J]. IEEE Transactions on Mobile Computing, 2006, 5(6): 738 - 754.
- [10] YE F, ZHONG G, LU S W, *et al.* Gradient broadcast: A robust data delivery protocol for large scale sensor networks [J]. ACM Wireless Networks, 2005, 11(3): 285 - 298.
- [11] CHEN D Z, VARSHNEY P K. On-demand geographic forwarding for data delivery in wireless sensor networks [J]. Computer Communication, 2007, 30(14/15): 2954 - 2967.
- [12] RIZZO L. Effective erasure codes for reliable computer communication protocols [J]. ACM Computer Communication Review, 1997, 27(2): 24 - 36.