

## 2. 适应性播放时延

上面的例子显示了当使用固定播放时延来设计播放策略时所引起的重要的时延与丢包折中。若初始播放时延设置得比较大,大多数分组能在它们的截止时间内到达,因此存在的丢失将可忽略不计;然而,对于如 VoIP 这样的会话式服务,长时延即使不是不能忍受的,至少也是令人厌恶的。在理想情况下,我们希望播放时延最小化,使丢包低于一定百分比的限制。

处理这种折中的自然方法是估计网络时延和网络时延变化,并且在每个话音突峰期的开始相应地调整播放时延。在话音突峰期开始时适应性调整播放时延将导致发送方静默期的压缩和拉长;然而,静默期的少量压缩和拉长在谈话中是不易觉察的。

根据 [Ramjee 1994],我们现在描述一种接收方可以用于适应性调整它的播放时延的通用算法。为此,令

$t_i$  = 第  $i$  个分组的时间戳 = 该分组在发送方产生的时间

$r_i$  = 分组  $i$  被接收方接收的时间

$p_i$  = 分组  $i$  在接收方播放的时间

第  $i$  个分组的端到端网络时延是  $r_i - t_i$ 。由于网络时延抖动,这个时延在不同的分组之间会发生变化。令  $d_i$  表示接收到第  $i$  个分组时的平均网络时延的估计值。这个估计值根据如下的时间戳来构造:

$$d_i = (1 - u)d_{i-1} + u(r_i - t_i)$$

式中  $u$  是一个固定的常数 (例如,  $u = 0.01$ )。这样  $d_i$  是观察到的网络时延  $r_1 - t_1, \dots, r_i - t_i$  的一个平滑均值。这个估计值为最近观察到的网络时延设置了比过去一段时间观察到的网络时延有更大的权重。这种估值的形式不应该是完全陌生的;相似的思想在第 3 章讨论估计 TCP 往返时间时就使用过。令  $v_i$  表示与估计平均时延的平均时延绝对偏差的估计值。这个估计值也可从这些时间戳构建:

$$v_i = (1 - u)v_{i-1} + u|r_i - t_i - d_i|$$

为每个接收的分组计算估计值  $d_i$  和  $v_i$ , 尽管它们仅能用于为任何话音突峰期的第一个分组确定播放点。

一旦计算完了这些估计值,接收方为分组播放应用下列的算法。如果分组  $i$  是一个话音突峰期的第一个分组, 它的播放时间  $p_i$  计算如下:

$$p_i = t_i + d_i + Kv_i$$

这里  $K$  是一个正的常数 (例如  $K = 4$ )。  $Kv_i$  项的目的是给将来设置足够大的播放时间, 以便话音突峰期中只有一小部分到达的分组由于迟到而丢失。在一个话音突峰期中任何后续分组的播放点被计算为对于这个话音突峰期的第一个分组播放时间点的偏移。特别是, 令

$$q_i = p_i - t_i$$

表示从话音突峰期的第一个分组产生到它播放的时间长度。如果分组  $j$  也属于这个话音突峰期, 它播放的时刻是

$$p_j = t_j + q_i$$

刚才描述的算法在假设接收方能够辨别一个分组是否是话音突峰期中的第一个分组的情况下很有意义。这能够通过检查每个接收到的分组中的信号能量来做到。