

(12) 发明专利

(10) 授权公告号 CN 101567853 B

(45) 授权公告日 2011. 08. 10

(21) 申请号 200910142922. 5

CN 101136772 A, 2008. 03. 05, 全文.

(22) 申请日 2009. 05. 13

审查员 郭海波

(73) 专利权人 中兴通讯股份有限公司

地址 518057 广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦法律部

(72) 发明人 丁向军 王魏强 夏洪升

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

代理人 田红娟 龙洪

(51) Int. Cl.

H04L 12/56 (2006. 01)

H04L 12/18 (2006. 01)

(56) 对比文件

CN 101296184 A, 2008. 10. 29, 全文.

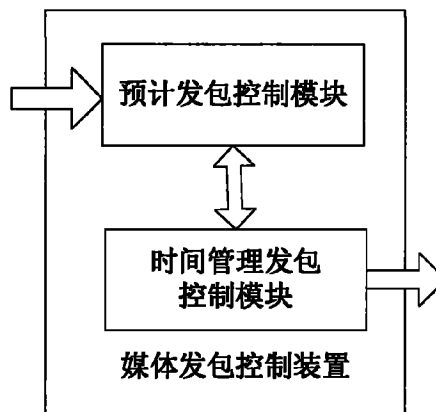
权利要求书 3 页 说明书 11 页 附图 9 页

(54) 发明名称

音频媒体发包控制装置、方法及音频媒体服务器

(57) 摘要

本发明提供一种音频媒体发包控制装置, 该装置包括预计发包控制模块及时间管理发包控制模块, 其中, 所述预计发包控制模块, 用于确定要发送的媒体报文的预计发送时间, 还用于将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上; 所述时间管理发包控制模块, 用于根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文, 并在所述预计发送时间将所述媒体报文发送出去; 若该指针数组元素所指的链表上没有媒体报文, 则延时到所述指针数组元素的预计发送时间。本发明音频媒体发包控制装置、方法及音频媒体服务器, 可以提高媒体报文打包间隔精度。



1. 一种音频媒体发包控制装置,其特征在于:该装置包括预计发包控制模块及时间管理发包控制模块,其中,

所述预计发包控制模块,用于确定要发送的媒体报文的预计发送时间,还用于将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上;

所述时间管理发包控制模块,用于根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文,并在所述预计发送时间将所述媒体报文发送出去;若该指针数组元素所指的链表上没有媒体报文,则延时到所述指针数组元素的预计发送时间。

2. 如权利要求 1 所述的装置,其特征在于:

所述媒体报文的预计发送时间是这样确定的:读取所述时间管理发包控制模块正在处理的当前指针数组元素的预计发送时间;在读取的预计发送时间的基础上增加一定的时间余量作为所述媒体报文的预计发送时间,以保证所述时间管理发包控制模块浏览到挂接在对应的指针数组元素的所述媒体报文。

3. 如权利要求 1 或 2 所述的装置,其特征在于:所述指针数组元素的索引及所述指针数组元素指向的链表上的媒体报文的预计发送时间具有依次对应关系,相邻指针数组元素的预计发送时间间隔是固定的。

4. 如权利要求 1 或 2 所述的装置,其特征在于:所述预计发包控制模块及时间管理发包控制模块采用多线程处理技术实现。

5. 一种媒体发包控制方法,其特征在于,该方法包括:

挂接步骤,指确定要发送的媒体报文的预计发送时间并将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上;

发送步骤,指根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文并在所述预计发送时间将所述媒体报文发送出去;若该指针数组元素所指的链表上没有媒体报文,则延时到该元素的预计发送时间。

6. 如权利要求 5 所述的方法,其特征在于:所述挂接步骤由多个预计发包线程实现,每个线程浏览其负责的通道,当发现某个通道要求发送媒体报文时,执行以下步骤:

A:读取时间管理发包控制线程中当前正在处理的当前指针数组元素的索引  $index$  和预计发送时间  $foresndtime$ ;

B:确定要发送的第一个媒体报文的挂接索引  $nextpkt\_tindex$  及预计发送时间  $nextpkt\_foresndtime$ ,其中,

$nextpkt\_tindex = (index+n) \% \text{指针数组的元素个数}$ ;

$nextpkt\_foresndtime = foresndtime+n* \text{每个指针数组索引代表的时间间隔}$ ;

其中, $n$  为预先设定的索引余量,以保证所述时间管理发包控制线程浏览到挂接在对应的指针数组元素的所述媒体报文;

C:确定后续媒体报文的挂接索引  $nextpkt\_tindex$  及预计发送时间  $nextpkt\_foresndtime$ ,直到所有媒体报文处理完毕,其中:

$nextpkt\_tindex = nextpkt\_tindex+( \text{打包间隔} / \text{每个索引代表的时间间隔} ) \% ( \text{指针数组的元素个数} )$ ,其中,所述打包间隔是每个索引代表的时间间隔的整数倍;

$nextpkt\_foresndtime = nextpkt\_foresndtime+( \text{打包间隔} )$ 。

7. 如权利要求 5 所述的方法,其特征在于:所述发送步骤由多个时间管理发包控制线

程实现,各时间管理发包控制线程循环执行以下步骤:

A:浏览指针数组当前索引的链表上是否有要发送的媒体报文,若有,则读取所述媒体报文,并把所述媒体报文的地址和载荷信息以及这批媒体报文的预计发送时间参数传到 sendmsg 函数;sendmsg 函数在预计发送时间把该批媒体报文批量发送出去,然后在预计发送时间返回;

若在当前索引的链表上没有要发送媒体报文,所述时间管理发包控制线程调用延时函数 DelayTime 延时到当前索引链表的预计发送时间;

B:更新当前索引 index 以及当前索引的链表上媒体报文的预计发送时间,其中,  
当前索引:  $index = (index+1) \% (\text{指针数组最大元素个数})$ ;  
当前索引对应的预计发送时间:  $foresndtime+ = \text{每个索引代表的时间间隔}$ 。

8. 如权利要求 6 或 7 所述的方法,其特征在于:所述挂接步骤中,每挂接一个媒体报文,对应链表结构的媒体报文计数器加 1;所述时间管理发包控制线程根据所述媒体报文计数器确定当前索引的链表上是否有要发送的媒体报文。

9. 一种音频媒体服务器,其特征在于,所述音频媒体服务器包括音频媒体发包控制装置,该装置包括预计发包控制模块及时间管理发包控制模块,其中,

所述预计发包控制模块,用于确定要发送的媒体报文的预计发送时间,还用于将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上;

所述时间管理发包控制模块,用于根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文并在所述预计发送时间将所述媒体报文发送出去;若该指针数组元素所指的链表上没有媒体报文,则延时到该元素的预计发送时间。

10. 如权利要求 9 所述的音频媒体服务器,其特征在于:

所述媒体报文的预计发送时间是这样确定的:读取所述时间管理发包控制模块正在处理的指针数组元素的预计发送时间;在读取的预计发送时间的基础上增加一定的时间余量作为所述媒体报文的预计发送时间,以保证所述时间管理发包控制模块浏览到挂接在对应的指针数组元素的所述媒体报文。

11. 如权利要求 9 或 10 所述的音频媒体服务器,其特征在于:所述音频媒体服务器还包括信令转发单元、主控单元及媒体存储转发单元,其中,

所述信令转发单元:用于实现服务器内部与外部信令的统一转发;

所述主控单元:用于接收并解析所述信令转发单元转发的控制信令生成相应的控制消息来控制媒体存储转发单元;

所述媒体存储转发单元,具有对外网口,用于根据所述主控单元下发的控制消息实现录音及语音播放功能,其中,由所述音频媒体发包控制装置实现媒体报文的发包控制。

12. 如权利要求 11 所述的音频媒体服务器,其特征在于:所述媒体存储转发单元实现的各种操作采用多线程处理技术实现。

13. 如权利要求 11 所述的音频媒体服务器,其特征在于:所述音频媒体服务器还包括媒体处理单元,所述媒体处理单元包括呼叫代理模块、收发号模块及会议处理模块,其中,

呼叫代理模块,用于实现主控单元和媒体处理单元之间控制消息的交互,若控制消息为收发号消息则通知收发号模块处理,若控制消息为转码或者会议消息则通知会议处理模块处理;

收发号模块,用于进行收号器资源设置,以及对媒体流中的有效号码进行有效性识别;还用于根据主控单元的控制消息产生号码;

会议处理模块,用于处理来自各个参与者的媒体流,把不同参与者的输入语音帧合成在一起作为混合语音帧,且输出给每个参与者的语音帧是把混和语音帧滤掉该参与者的输入语音帧;

所述主控单元,还用于生成对所述媒体处理单元的控制消息;

所述媒体存储转发单元还用于将含有双音多频 DTMF 号码的媒体报文的目的地地址转换成收发号器通道对应的地址。

## 音频媒体发包控制装置、方法及音频媒体服务器

### 技术领域

[0001] 本发明属于计算机应用领域的网络通信技术,具体涉及一种基于 Linux/vxworks 的音频媒体发包控制装置、方法及音频媒体服务器。

### 背景技术

[0002] 媒体服务器是下一代网络 (NGN, Next Generation Network) 中的重要设备,其在 NGN 网络中的位置如图 1 所示。现有音频媒体服务器中媒体处理系统涉及到对大量数据的解析、流化、打包间隔的精确控制等关键技术。针对这些关键技术,目前业内媒体服务器媒体处理系统主要用两种方式:硬件方式-专用 DSP (Digital Singnal Processing, 数字信号处理) 和软件方式-HMP (Host Media Processing, 主机媒体处理) 技术。但是它们的缺陷在于:硬件方式成本高、软件方式虽然成本降低了,但是容量却太小,一旦增加容量又造成流化和打包间隔精确控制性能降低。专用 DSP 技术是通过专有芯片中的特定算法来进行媒体数据的处理,一般来说专有 DSP 芯片媒体处理能力强,通道密度也即容量大,每个通道的价格一定的话,通道密度越高成本也就越高;而 HMP 技术是未来媒体服务器的发展方向,但相对于专有 DSP,主机媒体处理能力一般,因此在大容量的情况下,打包精度控制会显著下降,大容量和打包精度成为其发展的瓶颈。

### 发明内容

[0003] 本发明要解决的技术问题是提供一种音频媒体发包控制装置、方法及音频媒体服务器,提高媒体报文打包间隔精度。

[0004] 为解决上述技术问题,本发明提供一种音频媒体发包控制装置,该装置包括预计发包控制模块及时间管理发包控制模块,其中,

[0005] 所述预计发包控制模块,用于确定要发送的媒体报文的预计发送时间,还用于将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上;

[0006] 所述时间管理发包控制模块,用于根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文,并在所述预计发送时间将所述媒体报文发送出去;若该指针数组元素所指的链表上没有媒体报文,则延时到所述指针数组元素的预计发送时间。

[0007] 进一步地,

[0008] 所述媒体报文的预计发送时间是这样确定的:读取所述时间管理发包控制模块正在处理的当前指针数组元素的预计发送时间;在读取的预计发送时间的基础上增加一定的时间余量作为所述媒体报文的预计发送时间,以保证所述时间管理发包模块浏览到挂接在对应的指针数组元素的所述媒体报文。

[0009] 进一步地,

[0010] 所述指针数组元素的索引及所述指针数组元素指向的链表上的媒体报文的预计发送时间具有依次对应关系,相邻指针数组元素的预计发送时间间隔是固定的。

[0011] 所述预计发包控制模块及时间管理发包控制模块采用多线程处理技术实现。

[0012] 为解决上述技术问题,本发明还提供了一种媒体发包控制方法,该方法包括:

[0013] 挂接步骤,指确定要发送的媒体报文的预计发送时间并将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上;

[0014] 发送步骤,指根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文并在所述预计发送时间将所述媒体报文发送出去;若该指针数组元素所指的链表上没有媒体报文,则延时到该元素预计发送时间。

[0015] 进一步地,

[0016] 所述挂接步骤由多个预计发包线程实现,每个线程浏览其负责的通道,当发现某个通道要求发送媒体报文时,执行以下步骤:

[0017] A:读取时间管理发包控制线程中当前正在处理的当前指针数组元素的索引 index 和预计发送时间 foresndtime;

[0018] B:确定要发送的第一个媒体报文的挂接索引 nextpkt\_tindex 及预计发送时间 nextpkt\_foresndtime,其中,

[0019]  $nextpkt\_tindex = (index+n) \% \text{指针数组的元素个数}$ ;

[0020]  $nextpkt\_foresndtime = foresndtime+n* \text{每个指针数组索引代表的时间间隔}$ ;其中,n 为预先设定的索引余量,以保证所述时间管理发包控制线程浏览到挂接在对应的指针数组元素的所述媒体报文;

[0021] C:确定后续媒体报文的挂接索引 nextpkt\_tindex 及预计发送时间 nextpkt\_foresndtime,直到所有媒体报文处理完毕,其中:

[0022]  $nextpkt\_index = nextpkt\_index + (\text{打包间隔} / \text{每个索引代表的时间间隔}) \% (\text{指针数组的元素个数})$ ,其中,所述打包间隔是每个索引代表的时间间隔的整数倍;

[0023]  $nextpkt\_foresndtime = nextpkt\_foresndtime + (\text{打包间隔})$ 。

[0024] 所述发送步骤由多个时间管理发包控制线程实现,各时间管理发包控制线程循环执行以下步骤:

[0025] A:浏览指针数组当前索引的链表上是否有要发送的媒体报文,若有,则读取所述媒体报文,并把所述媒体报文的地址和载荷信息以及这批媒体报文的预计发送时间参数传到 sendmsg 函数;sendmsg 函数在预计发送时间把该批媒体报文批量发送出去,然后在预计发送时间返回;

[0026] 若在当前索引的链表上没有要发送媒体报文,所述时间管理发包控制线程调用延时函数 DelayTime 延时到当前索引链表的预计发送时间;

[0027] B:更新当前索引 index 以及当前索引的链表上媒体报文的预计发送时间,其中,

[0028] 当前索引:  $Index = (Index+1) \% (\text{指针数组最大元素个数})$ ;

[0029] 当前索引对应的预计发送时间:  $foresndtime+ = \text{每个索引代表的时间间隔} (10ms)$ 。

[0030] 进一步地,

[0031] 所述挂机步骤中,每挂接一个媒体报文,对应链表结构的媒体报文计数器加 1;所述时间管理控制线程根据所述媒体报文计数器确定当前索引的链表上是否有要发送的媒体报文。

[0032] 为解决上述技术问题,本发明还提供一种音频媒体服务器,所述音频媒体服务器

包括该音频媒体发包控制装置,该装置包括预计发包控制模块及时间管理发包控制模块,其中,

[0033] 所述预计发包控制模块,用于确定要发送的媒体报文的预计发送时间,还用于将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上;

[0034] 所述时间管理发包控制模块,用于根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文并在所述预计发送时间将所述媒体报文发送出去;若该指针数组元素所指的链表上没有媒体报文,则延时到该元素预计发送时间。

[0035] 进一步地,

[0036] 所述媒体报文的预计发送时间是这样确定的:读取所述时间管理发包控制模块正在处理的指针数组元素的预计发送时间;在读取的预计发送时间的基础上增加一定的时间余量作为所述媒体报文的预计发送时间,以保证所述时间管理发包模块浏览到挂接在对应的指针数组元素的所述媒体报文。

[0037] 进一步地,所述音频媒体服务器还包括信令转发单元、主控单元及媒体存储转发单元,其中,

[0038] 所述信令转发单元:用于实现服务器内部与外部信令的统一转发;

[0039] 所述主控单元:用于接收并解析所述信令转发单元转发的控制信令生成相应的控制消息来控制媒体存储转发单元;

[0040] 所述媒体存储转发单元,具有对外网口,用于根据所述主控单元下发的控制消息实现录音及语音播放功能,其中,由所述音频媒体发包控制装置实现媒体报文的发包控制。

[0041] 进一步地,所述媒体存储转发单元实现的各种操作采用多线程处理技术实现。

[0042] 进一步地,所述音频媒体服务器还包括媒体处理单元,所述媒体处理单元包括呼叫代理模块、收发号模块及会议处理模块,其中,

[0043] 呼叫代理模块,用于实现主控单元和媒体处理单元之间控制消息的交互,若控制消息为收发号消息则通知收发号模块处理,若控制消息为转码或者会议消息则通知会议处理模块处理;

[0044] 收发号模块,用于进行收号器资源设置,以及对媒体流中的有效号码进行有效性识别;还用于根据主控单元的控制消息产生号码;

[0045] 会议处理模块,用于处理来自各个参与者的媒体流,把不同参与者的输入语音帧合成在一起作为混合语音帧,且输出给每个参与者的语音帧是把混和语音帧滤掉该参与者的输入语音帧;

[0046] 所述主控单元,还用于生成对所述媒体处理单元的控制消息;

[0047] 所述媒体存储转发单元还用于将含有多音双频 DTMF 号码的媒体报文的地址转换成收发号器通道对应的地址。

[0048] 相较于现有技术,本发明发包控制装置和方法利用发包线程管理带时间信息的指针数组的方法实现对媒体报文指针数组挂载索引和预计发送时间的准确有效的控制,能高效的把该索引位置的所有媒体报文在预定发送时刻批量的发送出去,实现了对打包间隔的精确控制,能够提供高精度 Qos(Quality of Service) 性能,另外采用多线程处理技术,具有大容量的特点。

[0049] 本发明音频媒体服务器,基于 HMP 实现,成本低,不仅实现对打包间隔的精确控

制,容量可与专用 DSP 媲美,可以提供基本和增强业务中的媒体处理功能,包括业务音提供、会议、交互式应答、通知和高级语音业务等功能。

#### 附图说明

- [0050] 图 1 是 NGN 网络中媒体服务器的位置的示意图。
- [0051] 图 2 是本发明媒体服务器各个功能模块的示意图。
- [0052] 图 3 是音源文件结构的示意图。
- [0053] 图 4 是音源文件解析流程图。
- [0054] 图 5 是本发明采用的多线程处理技术示意框图。
- [0055] 图 6 是本发明音频媒体发包控制装置的模块结构示意图。
- [0056] 图 7 是本发明音频媒体发包控制方法的示意图。
- [0057] 图 8 是本发明预计发包控制模块的处理流程图。
- [0058] 图 9 是本发明时间管理发包控制模块处理流程图。
- [0059] 图 10 是 DTMF 号码收发号流程图。
- [0060] 图 11 是会议电话流程图。

#### 具体实施方式

[0061] 如图 2 所示,本发明音频媒体服务器包括信令转发单元、与信令转发单元连接的主控单元、控制 / 媒体交换单元、媒体存储转发单元、媒体处理单元,其中主控单元、媒体存储转发单元及媒体处理单元均通过控制 / 媒体交换单元相互交互控制流或媒体流,媒体存储转发单元、媒体处理单元是本发明媒体处理技术实现的载体。

[0062] 信令转发单元:主要实现系统内部网络与外部网络的隔离,设备的控制信令 SIP(Session Initiation Protocol,会话初始化协议)/H248/MGCP(Media Gateway Control Protocol,媒体网关控制协议)统一通过信令转发单元转发,实现设备对外单一的 IP(Internet Protocol) 地址呈现;

[0063] 主控单元:用于接收并解析控制信令生成相应的控制消息来控制媒体存储转发单元和媒体处理单元对放音、录音、收发号、会议、转码进行相应通道处理;

[0064] 控制信令是按国际标准格式的消息,而控制消息是把控制信令中的各个字段解释翻译后,组成的适合于系统内部的传递格式的消息。

[0065] 控制 / 媒体交换单元:用于完成服务器内部控制消息交换 / 媒体数据包交换,其实现交换机的功能,是其他模块之间通信的桥梁,在此不再赘述。

[0066] 媒体存储转发单元:完成海量的多媒体数据存储,包括实现语音播放功能的语音库、实现录音功能时对用户录音的存储。媒体存储转发单元设置有对外网口,当媒体不需要额外转换处理时,直接通过单元上的对外网口收发。

[0067] 通过对外网口发送的媒体报文是流化后的媒体报文,该报文是符合国际标准格式的报文。

[0068] 在实现录音业务时,通过对外网口接收媒体流如果终端(比如固定电话终端)想录音,并且在录制的过程中,没有按任何 DTMF 号码(0-9、\*、#),那么此时媒体流中就不会含有 DTMF 号码;如果录制完毕时(比如:以 # 号表示录音完毕的结束键),此时按 #,那么此时



接收的媒体流中就含有 DTMF 号码 (# 号), 探测到 # 号, 并上报给主控单元, 主控单元就会下发控制消息来关闭该通道停止录音。

[0069] 媒体存储转发单元包括内部 call 代理模块、录音媒体处理模块、放音媒体处理模块、发包控制模块、收包控制模块及存储模块, 其中,

[0070] 内部 call 代理 (呼叫代理) 模块, 用于实现与主控模块之间控制消息的交互, 这些消息是通过控制 / 媒体交换单元的控制平面实现的。若分析接收的控制消息为录音则通知录音处理模块; 若为放音, 则通知放音处理模块。

[0071] 录音媒体处理模块, 用于从把收包控制模块接收来的媒体报文排序后, 提取其有效载荷存储为指定的文件格式 (WAV/AMR)。

[0072] 放音媒体处理模块, 用于根据放音控制消息中的参数从存储模块找到音源文件, 并对音源文件进行解析、读取音源的有效数据、打包流化为适合于网络传输的媒体报文; 若根据音源数据流化后的媒体报文编码格式与终端需要的编码格式不一致, 需要发送到媒体处理单元进行处理;

[0073] 发包控制模块, 用于通过内外网口的发送媒体报文, 具体地, 该发包控制模块即发包控制装置, 是本发明的核心, 具体将在下文中进行详细说明。

[0074] 收包控制模块, 用于通过内外网口的接收媒体报文, 如果从外网口接收到的媒体报文与需要的存储的音源文件格式 (WAV/AMR) 不一致就需要转发到媒体处理单元, 由媒体处理单元进行转码后发送到收包控制模块的内网口, 然后给录音媒体处理模块处理, 还用于将含有 DTMF 号码的媒体报文的地址转换成收发号器通道对应的地址。

[0075] 存储模块, 用于存储音源文件。

[0076] 媒体处理单元: 完成语音编解码方式转换、实现 DTMF (Dual Tone Multi-Frequency, 双音多频) 收发号以及会议混音功能。

[0077] 媒体处理单元有以太网口和 TDM (Time Division Multiplex and Multiplexer, 时分复用) 接口, 来实现 IP 用户和 PSTN (Public Switched Telephone Network, 公用交换电话网) 用户的各自需要。

[0078] 媒体处理单元包括 call 代理模块、收发号模块及会议处理模块, 其中,

[0079] call 代理模块, 用于实现主控单元和媒体处理单元之间控制消息的交互, 若控制消息为收发号消息则通知收发号模块处理, 若控制消息为转码或者会议消息则通知会议处理模块处理

[0080] 收发号模块, 用于进行收发号器资源设置, 以及对媒体流中的有效号码进行有效性识别; 还用于根据主控单元的控制消息产生号码;

[0081] 会议处理模块, 用于处理来自各个参与者的媒体流, 把不同参与者的输入语音帧合成在一起作为混合语音帧; 输出给每个参与者的语音帧是把混和语音帧滤掉该参与者的输入语音帧, 这样就避免该用户听到自己的回声。

[0082] 本发明音频媒体服务器在 HMP 方式改进的基础上提供低成本和大容量来克服上述 DSP 和 HMP 的不足, 一种基于 Linux/vxworks 操作系统的适用于各种网络用户的媒体处理技术实现方案, 本发明媒体服务器在应用服务器的控制下, 提供各种业务所需的媒体处理功能, 包括: 放音, 录音, DTMF 收号, 会议、转码等媒体处理技术。本发明媒体处理技术的要点如下:

[0083] 一、对于音源文件格式的解析和构造：

[0084] 由于不同类型的文件对音频采样数据的组织形式不同，所以需要根据相应的格式进行分析。一般来说，文件包含对数据的描述部分和数据部分。目前，对音源文件的处理支持 WAV (Wave Form, 波形) 和 AMR (Adaptive Multi-Rate, 自适应多速率) 等文件类型，如图 3 所示，当然根据要求也可以支持其他的音源类型，完全由软件控制。

[0085] WAV 文件头作为多媒体中使用的声波文件格式之一，它是以 RIFF (Resource Interchange File Format, 资源交换文件格式) 格式为标准的。AMR 全称 Adaptive Multi-Rate, 自适应多速率编码，主要用于移动设备的音频，压缩比较大。

[0086] WAV 文件和 AMR 文件文件的解析流程框图如图 4，构造流程与此类似。其中，

[0087] 对 WAV 文件的解析过程：

[0088] 读取 WAV 文件的文件头；提取文件头信息，包括数据编码方式、频道数、采样频率、每个采样需要的 bit 数等信息；跳过文件描述信息；从音频数据部分读取数据。

[0089] 对 AMR 文件的解析过程：

[0090] 读取文件的 magic number；读取首帧数据；根据帧头计算帧数据的大小；读帧头；判断是否坏帧，若坏帧则读取下一个帧头，否则读取本帧语音数据。

[0091] 二、语音数据的流化和流化数据的语音构造

[0092] 流化处理指把读取的音源文件中的数据进行 RTP (Real-time Transport Protocol, 实时传送协议) 打包处理，构造成可在 IP 网上传输的符合 RFC3550 标准的媒体报文。这其中包括构造 IP/UDP (User Datagram Protocol, 用户数据报协议, )/RTP 头和 payload (有效载荷) 数据。

[0093] 若语音数据的编码格式与终端用户所要求的编解码格式相同，则不需要转换处理，直接由媒体存储转发单元的放音处理模块进行流化处理后通过对外网口发送；

[0094] 若语音数据的编码格式与终端用户所要求的编解码格式 (如 G. 711a/u、G. 729、G. 723) 不一致，则需要进行语音数据转换，流化处理的过程由媒体存储转发单元和媒体处理单元共同完成，两者之间的媒体流交互通过控制 / 媒体交换单元的媒体平面实现。

[0095] 可以通过以下两种方式实现流化处理过程：

[0096] 方式一：媒体存储转发单元的放音处理模块将语音数据还原成 PCM8 或 PCM16 的格式，媒体处理单元通过算法将其转为对应的 G. 711/G. 729/G. 723 的编码数据，最后媒体存储转发单元的放音处理模块基于实时传送协议 RTP 进行构包处理。

[0097] 方式二：媒体存储转发单元的放音处理模块将音源文件中的语音数据，比如 a/u-law 或者 AMR 帧，直接流化处理成 G. 711a/u 或者 RFC3267 payload 格式的 RTP 报文，媒体处理单元的会议处理模块再将 RTP 处理后数据转码为目的编解码 G. 729/G. 723 方式。本发明推荐使用第二种方式。

[0098] 当然对于 WAV 文件的流化，需要知道 WAV 文件中的音源数据编码格式 (音源数据编码格式在文件头中)。IP/UDP 的源和目的地址信息由主控单元带下来的参数决定，payload 数据的长度由主控单元带下来的打包时长信息决定。

[0099] 对于 AMR 文件的流化，需要参考 rfc3267 标准把存储格式变为传输格式的流化数据。这里涉及到不同的 payload 格式的问题：字节对齐格式和带宽效率格式。需要根据信令中 SDP 描述选择不同的格式，再进行音源语音帧在 payload 中的格式变换。

[0100] 流化数据的语音构造相反：就是把流化数据中的 payload 变化为音源数据进行存储。在这个过程中需要对接收的 RTP 数据进行排序和抖动处理，并进行格式转换以便语音数据按正确格式存于相应的文件类型。

[0101] 三、大容量的多线程处理技术

[0102] 媒体存储转发单元的各模块实现的各个操作，包括录音处理时文件头信息构造、rtp 报文载荷提取、有效语音存储及放音处理时的音源解析、数据读取、流化、打包以及发包，都采用多线程处理技术，即每种操作有多个线程并行处理，从而实现多通道大容量的特点。

[0103] 操作文件最常用的一个模型是同步阻塞 I/O 模型。在这个模型中，用户空间的应用程序执行一个系统调用，这会导致应用程序阻塞。如图 5 所示，以读操作为例，本发明利用线程族（也就是起多个线程）来模拟异步读操作；读请求线程决定是否发起读操作，读文件线程读取数据；由多个这样的读请求线程和读文件线程构成读请求线程族和读文件线程族；这样通过并发读文件线程就可以模拟异步读文件操作了，另外图中还显示了进行打包处理的构造 RTP 包线程族及执行发包处理的发包线程族。写文件的模拟也是类似的。在需要执行文件操作时，在读写请求线程与读写文件线程之间用消息队列的方式在两个线程族之间通信。

[0104] 一旦某些通道需要发送媒体流，那么对于所读取的语音数据要不停的打包和发包处理。如果全由一个进程或线程来处理，肯定会造成“拥塞”的情况，这就对大容量造成了限制；而多个线程并发处理的运用对这些问题的解决奠定了基础。

[0105] 四、高精度发包控制技术

[0106] 媒体服务器中有大量的媒体流从媒体存储转发单元的发包控制模块即发包控制装置的内外网口发送。为了对流化数据（rtp 数据）的打包间隔进行精确控制，本发明发包控制装置通过发包线程管理带时间信息的指针数组的方法实现，如图 6 所示，本发明发包控制装置包括预计发包控制模块及时间管理发包控制模块，其中：

[0107] 预计发包控制模块用于缓存放音媒体处理模块流化后的媒体报文，确定预计发送时间并挂接到指针数组对应发送时刻的元素所指的链表上。

[0108] 每个网口都对应一定数量的缓存，用来存储媒体报文；当没有媒体报文要处理时，所有的缓存都在 idlelist 上，若有媒体报文需要处理，预计发包控制模块就会从空闲链表 idlelist 上申请缓存用来装载媒体报文，然后把该报文挂接到该网口对应的时间管理发包控制模块管理的指针数组元素上。等时间管理发包控制模块把该数组元素上的报文发送完毕后，又会把这些缓存挂接到该网口的空闲链表 idlelist 上。

[0109] 所述媒体报文的预计发送时间是这样确定的：读取所述时间管理发包控制模块正在处理的当前指针数组元素的预计发送时间；在读取的预计发送时间的基础上增加一定的时间余量作为所述媒体报文的预计发送时间，以保证所述时间管理发包模块浏览到挂接在对应的指针数组元素的所述媒体报文。

[0110] 所述指针数组元素的索引及所述指针数组元素指向的链表上的媒体报文的预计发送时间具有依次对应关系，相邻指针数组元素的预计发送时间间隔是固定的。

[0111] 时间管理发包控制模块用于根据指针数组元素索引依次读取指针数组元素所指链表上的媒体报文，并在预计发送时间发送出去；如果指针数组元素链表头没有携带任何

媒体报文,那么需要延时到该元素预计发送时间。

[0112] 所述预计发包控制模块及时间管理发包控制模块采用多线程处理技术实现。对应地,如图 7 所示,本发明媒体发包控制方法包括以下步骤:

[0113] 步骤 701:挂接步骤,指确定要发送的媒体报文的预计发送时间并将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上;

[0114] 步骤 702:发送步骤,指根据指针数组元素索引依次读取指针数组元素所指的链表上的媒体报文并在所述预计发送时间将所述媒体报文发送出去;若该指针数组元素所指的链表上没有媒体报文,则延时到该元素预计发送时间。

[0115] 以下结合附图对模块功能及步骤进行详细的说明:

[0116] 预计发包控制模块的流程图如图 8 所示,当预计发包控制模块的预计发包线程族由调度进程创建起来后进入永远的 while 循环,在这个 while 循环中会浏览该线程负责所有的通道,一旦发现某个通道要求发送媒体报文就会进入打包函数中,在打包函数中会查看一下要发送的媒体报文是否第一个,如果是,那么按以下步骤进行:

[0117] A、读取时间管理发包控制线程中当前正在处理的指针数组的索引 index 和预计发送时间 foresndtime;为了能够把该第一个媒体报文发送出去,那么第一个媒体报文的索引和预计发送时间必须是 index 和 foresndtime 之后的索引和时间,所以需要在 index 和 foresndtime 上保留一定的余量(索引余量及时间余量),以便时间管理发包控制线程会浏览到(时间管理发包控制线程按索引递增浏览的);目前第一个媒体报文的索引 nextpkt\_tindex 及预计发送时间 nextpkt\_foresndtime 的计算方法为:

[0118]  $nextpkt\_tindex = (index+n) \% \text{指针数组的元素个数};$

[0119]  $nextpkt\_foresndtime = foresndtime+n* \text{每个指针数组索引代表的时间间隔};$

[0120] 其中,n 为预先设定的索引余量(比如  $n = 5$ ),以保证所述时间管理发包控制线程浏览到挂接在对应的指针数组元素的所述媒体报文;对应的时间余量即索引余量 n 与每个指针数组索引代表的时间间隔的乘积。

[0121] 目前每个指针数组索引代表的时间间隔即相邻指针数组元素的预计发送时间间隔是固定的(如 10ms),索引和预计发送时间其实是依次对应的关系的,第一个媒体报文的发送索引是当前索引往后移 n 个索引,对应的预计发送时间是当前预计发送时间往后移 n 个索引对应的时间间隔。索引号找到后会该媒体报文挂接到指针数组的对应索引的链表上,同时该链表结构中的媒体报文计数器会加 1。

[0122] B、计算该通道的下一个要发送的媒体报文的索引和预计发送时间,算法如下:

[0123]  $nextpkt\_index = nextpkt\_index + (\text{打包间隔} / \text{每个索引代表的时间间隔}) \% (\text{指针数组的元素个数})$

[0124]  $nextpkt\_foresndtime = nextpkt\_foresndtime + (\text{打包间隔})$

[0125] 这里的打包间隔是每个索引代表的时间间隔的整数倍,最简单的情况就是打包间隔与每个索引代表的时间间隔是相同的。

[0126] 该通道接下来的发送的第二个媒体报文就利用以上计算的索引号和预计发送时间找到指针数组的对应索引的链表上,同时该链表结构中的媒体报文计数器会加 1。

[0127] C、后续的媒体报文的所挂接的索引号和预计发送时间都重复 B 步骤进行,直到媒体报文处理完毕。

[0128] 根据以上方式确定媒体报文的指针数组元素索引可以保证的媒体报文都会被发送,不会产生丢包的现象。每个指针元素上存放的媒体报文没有量的限制,即使同时有多个通道的媒体报文需要发送,只要根据以上原则确定指针数组元素索引,以上多个通道的媒体报文可以存放在同一指针数组元素索引对应的链表上,从而为媒体报文的批量发送提供了前提条件。

[0129] 时间管理发包控制模块的流程图如图 9,当时间管理发包控制模块的时间管理发包控制线程族由调度进程创建起来后,就会获取当前系统时间,并把该时间圆整到秒,作为指针数组第一个索引 0 链表上携带的媒体报文的预计发送时间 T0,当前索引  $index = 0$ ,然后永远的 while 循环,在这个 while 循环中会:

[0130] A、浏览指针数组当前索引的链表上是否有要发送的媒体报文(通过媒体报文计数器),一旦发现当前索引的链表上有要发送媒体报文就会把这些媒体报文取下来,然后把这些媒体报文的地址和载荷信息以及这批媒体报文的预计发送时间参数传到 sendmsg 函数中,由该函数在预预计发送时间把该批报文批量发送出去,然后在预计发送时间返回;若在当前索引的链表上没有要发送媒体报文,那么会调用延时函数 DelayTime 延时到该索引链表的预计发送时间。

[0131] B、更新当前索引 index 以及当前索引的链表上媒体报文的预计发送时间。更新方法:

[0132] 当前索引:

[0133]  $Index = (Index + 1) \% (\text{指针数组最大元素个数})$

[0134] 当前索引对应的预计发送时间:

[0135]  $foresndtime + = \text{每个索引代表的时间间隔}(10\text{ms})$

[0136] C、重复步骤 A、B,永远的循环下去

[0137] 代码实现如下:

```
[0138] void*TimeQueueThread(void*arg)
[0139] {
[0140]     gettimeofday(&tpSndTime, NULL);
[0141]     tpSndTime.tv_sec+ = 3;
[0142]     tpSndTime.tv_usec = 0;
[0143]     tpLastSndTime = tpSndTime;
[0144]     wQStartIndex = 0;
[0145]     PointerArray[wQStartIndex].QForeCastSndTime
[0146]         = tpSndTime;
[0147]     while(1)
[0148]     {
[0149]         wQIndex = wQStartIndex;
[0150]         wEleNum = PointerArray[wQIndex].wElementNum
[0151]         if(! wEleNum)
[0152]             {
[0153]                 DelayTime();
```

```

[0154]         }
[0155]         Else
[0156]         {
[0157]             Sendmsg() ;
[0158]         }
[0159]         NextTv (&tpLastSndTime, MinPktTime*1000, &tpSndTime) ;
[0160]         tpLastSndTime = tpSndTime ;
[0161]         wQStartIndex+ = 1 ;
[0162]         wQStartIndex = wQStartIndex % QUEUE_MAX ;
[0163]         PointerArray[wQStartIndex]. QForeCastSndTime
[0164]             = tpSndTime ;
[0165]     }

```

[0166] 因为对终端用户来说,打包间隔如果抖动过大那么就会严重影响语音质量,这个打包间隔精确度是 Qos 的重要指标,直接影响运营商对产品质量的评估。可以说,在文件读取数据和打包能够及时供给的情况下, Qos 就完全依赖打包间隔精确度了。

[0167] 本发明发包控制装置,通过对媒体报文的在指针数组挂载索引和预计发送时间的准确有效的控制,能高效的把该索引位置的所有媒体报文在预定发送时刻批量的发送出去。

[0168] 五、DTMF 收发号功能,支持带内方式和 RFC2833 方式

[0169] DTMF 收发号功能是在媒体处理单元的收发号模块上实现的,如图 10 所示,当媒体处理单元的收发号模块收到主控单元要求打开收发号器收发号码时,首先进行一系列的收号器资源的设置(包括创建收发号通道;设置语音通道参数、DTMF 号码的探测方式及激活收号器资源),然后媒体存储转发单元通过 nat(Network Address Translation,网络地址转换)把外网来的含有 DTMF 号码的媒体流转发到对应的媒体处理单元的收号资源上,由收号资源对该媒体流中的有效号码进行探测和有效性的识别(符合国标 GB9038-88);发送号码,就是根据 GB9038-88 产生有效的 DTMF 号码通过媒体存储转发单元的 NAT 转发到目的终端上去。以上是描述了 IP 侧的收发号过程,TDM 侧的收发号,需要通过 E1/T1 把 PCM(Pulse Code Modulation,脉码调制)流引入到媒体处理单元的收号资源上作同样的流程处理。只不过配置收发号时配置参数不一样。

[0170] 主控模块是通过控制消息对收发号器的参数进行配置,让收发号器处于工作状态,当想收号的,媒体流是通过媒体存储转发单元的 nat 转发进来的,由于外网设备只能看到媒体存储转发单元的外网口,其媒体报文的地址也是外网口的地址(ip+port),因此需要媒体存储转发单元把该媒体报文的地址在外网口替换成收发号器通道对应的地址,然后发送到收发号器通道上,而外网口的地址与收发号器通道对应的地址的映射关系就是在 nat 表保存着的;当想发送号码时需要根据主控发号消息的指示产生号码。

[0171] 六、支持 3 方及多方会议功能,最大可支持 64 方混音

[0172] 会议/转码都是在媒体处理单元的会议处理模块上实现的;会议是允许多个用户参与到一个会议中,每个用户可以听到其他用户的谈话。会议的参与者可以是 TDM 用户也可以是 IP 用户。会议/转码中有关键的处理方法,该方法能够把不同参与者的输入语音帧

合成在一起作为混合语音帧;输出给每个参与者的语音帧是把混和语音帧滤掉该参与者的输入语音帧,这样就避免该用户听到自己的回声。每个参与者可以是不同的编解码;转码只是会议的一种特殊形式。

[0173] 会议的参与者可以设置为“哑音”状态,此时混音器会把该参与者作为静音处理,但是该参与者可以听到其他参与者的声音;会议的参与者可以设置为“保持”状态,此时混音器把该参与者作为静音处理,并且该参与者也听不到其他参与者的声音;会议的参与者可以设置为“通告保持”状态,此时混音器把该参与者作为静音处理,该参与者可听到源参与者的声音;会议的参与者可以设置为“源参与者”状态,此时混音器把该参与者的输入输到其他“通告保持”的参与者,使其他“通告保持”的参与者都能够听到“源参与者”的声音;“源参与者”典型的用来播放音乐、新闻、以及通告等等。

[0174] 创建会议的实现流程如图 11,当会议处理模块接收到主控单元发送的创建会议消息时,若会议 ID 已经存在则上报主控单元,否则创建会议 ID 及参与者 ID,并设置参与者参数、使能参与者对应的通道,会议进行中执行混合语音帧、过滤参与者自身的语音帧的处理过程;当收到参与者加入会议的消息时,同样执行创建参与者 ID 及设置参与者参数、使能参与者对应的通道的过程;若在会议进行中收到主控单元下发的撤销会议参与者的消息,则使会议参与者对应的通道失效并撤销对应参与者的 ID,当参与者只剩下一个时,撤销会议 ID。

[0175] 本发明发包控制装置和方法,利用发包线程管理带时间信息的指针数组的方法实现对媒体报文指针数组挂载索引和预计发送时间的准确有效的控制,能高效的把该索引位置的所有媒体报文在预定发送时刻批量的发送出去,实现了对打包间隔的精确控制,能够提供高精度 Qos(Quality of Service) 性能,另外采用多线程处理技术,具有大容量的特点。

[0176] 本发明音频媒体服务器,基于 HMP 实现,成本低,不仅实现对打包间隔的精确控制,容量可与专用 DSP 媲美,可以提供基本和增强业务中的媒体处理功能,包括业务音提供、会议、交互式应答、通知和高级语音业务等功能。

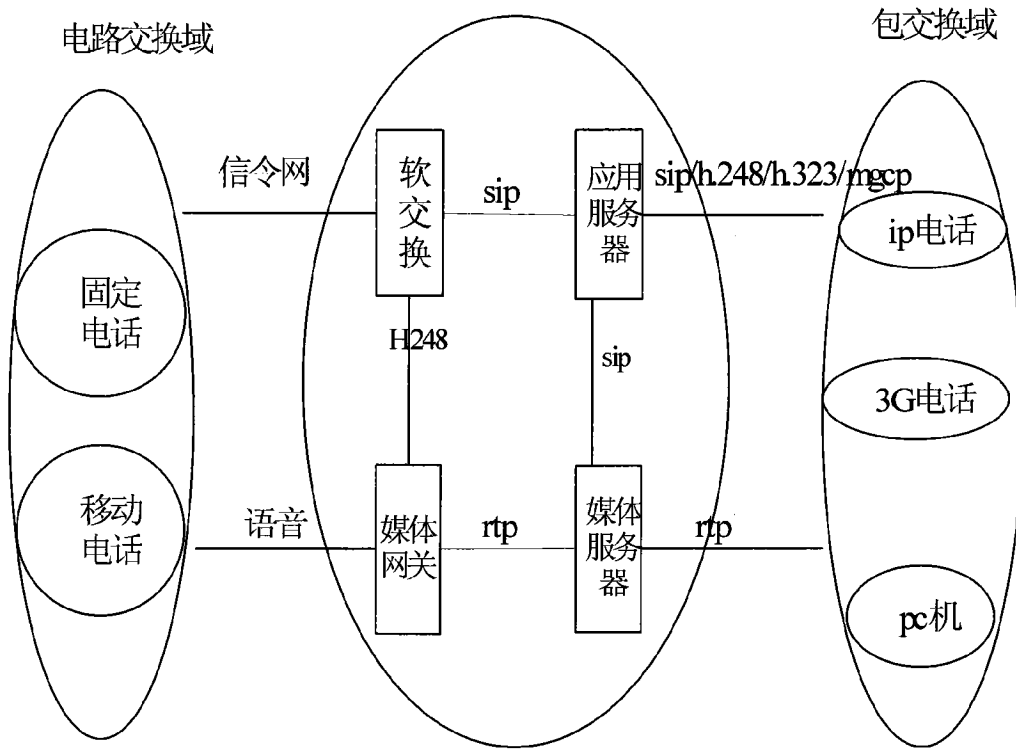


图 1

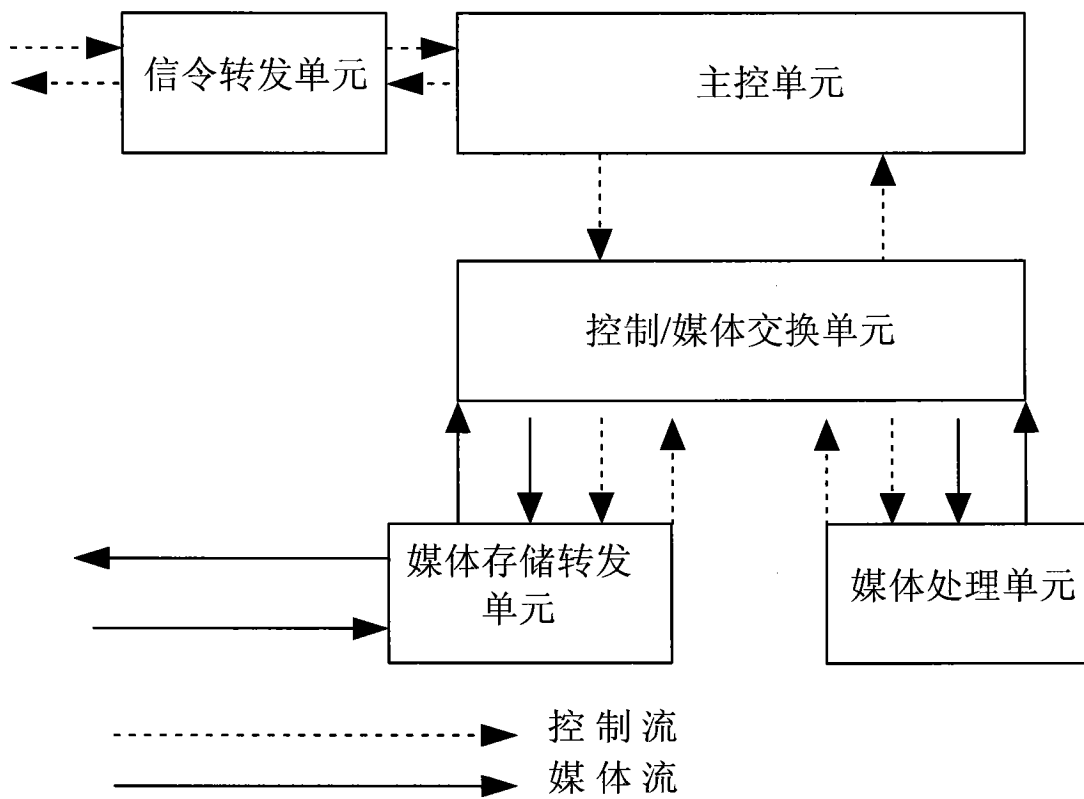
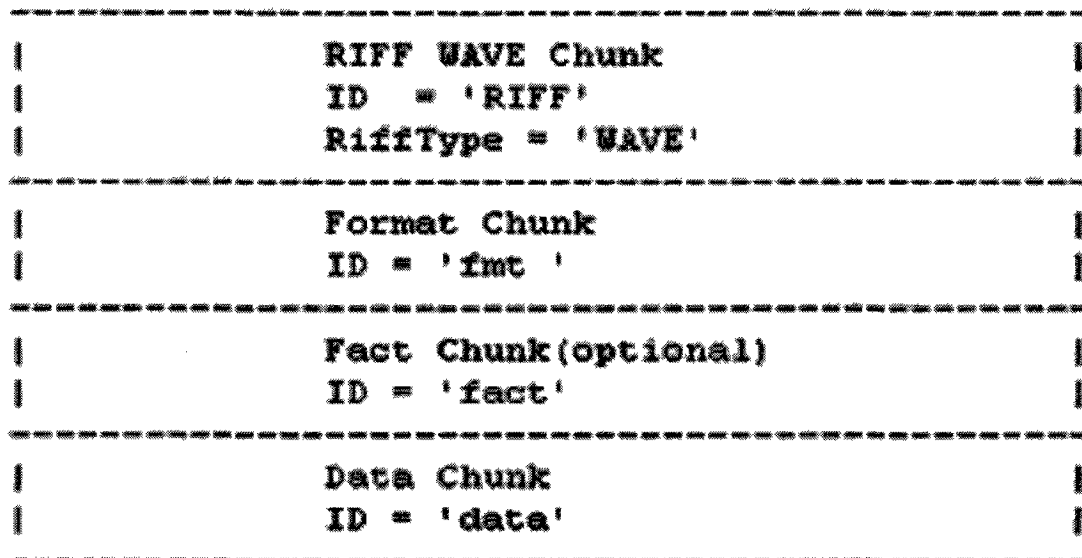
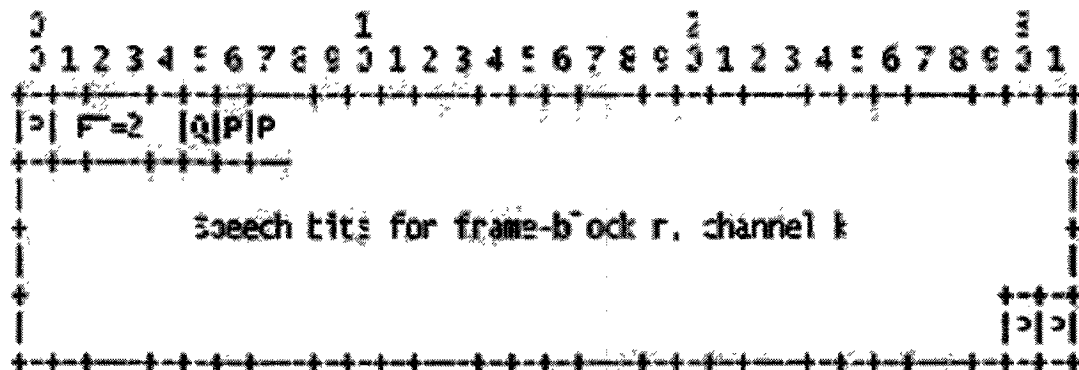
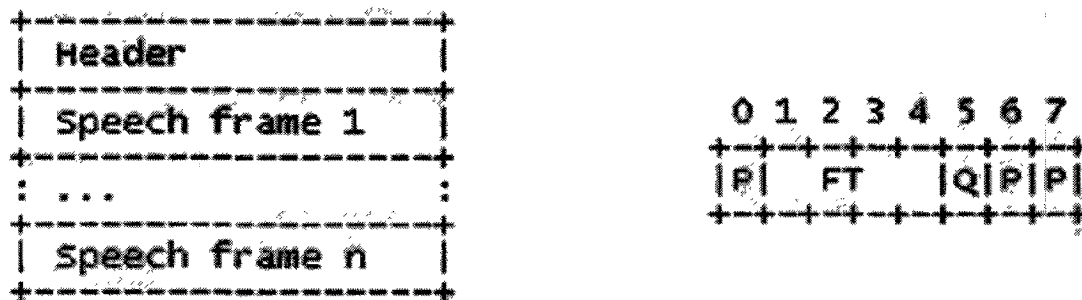


图 2





WAV文件格式



AMR文件格式

图 3

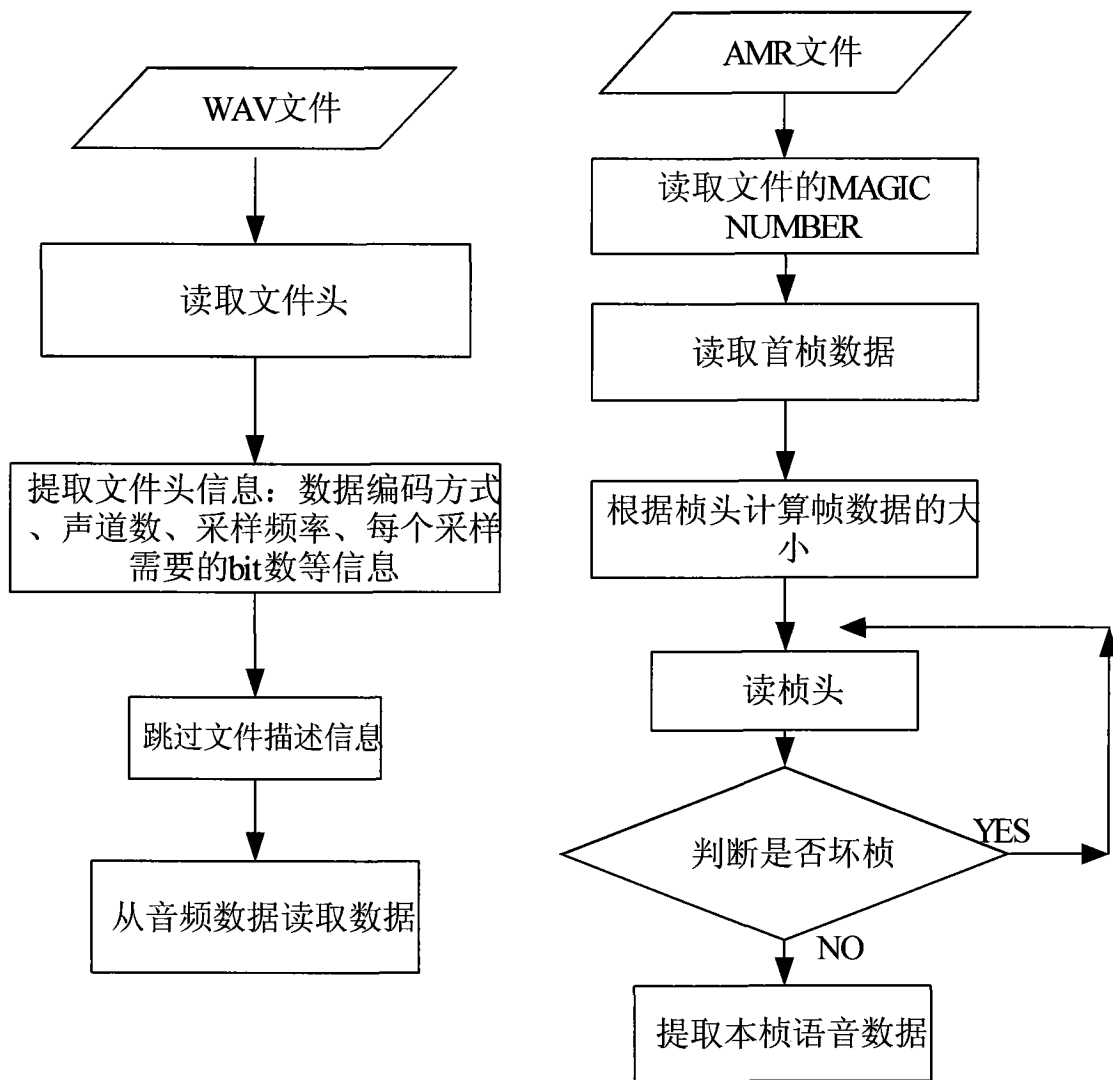


图 4

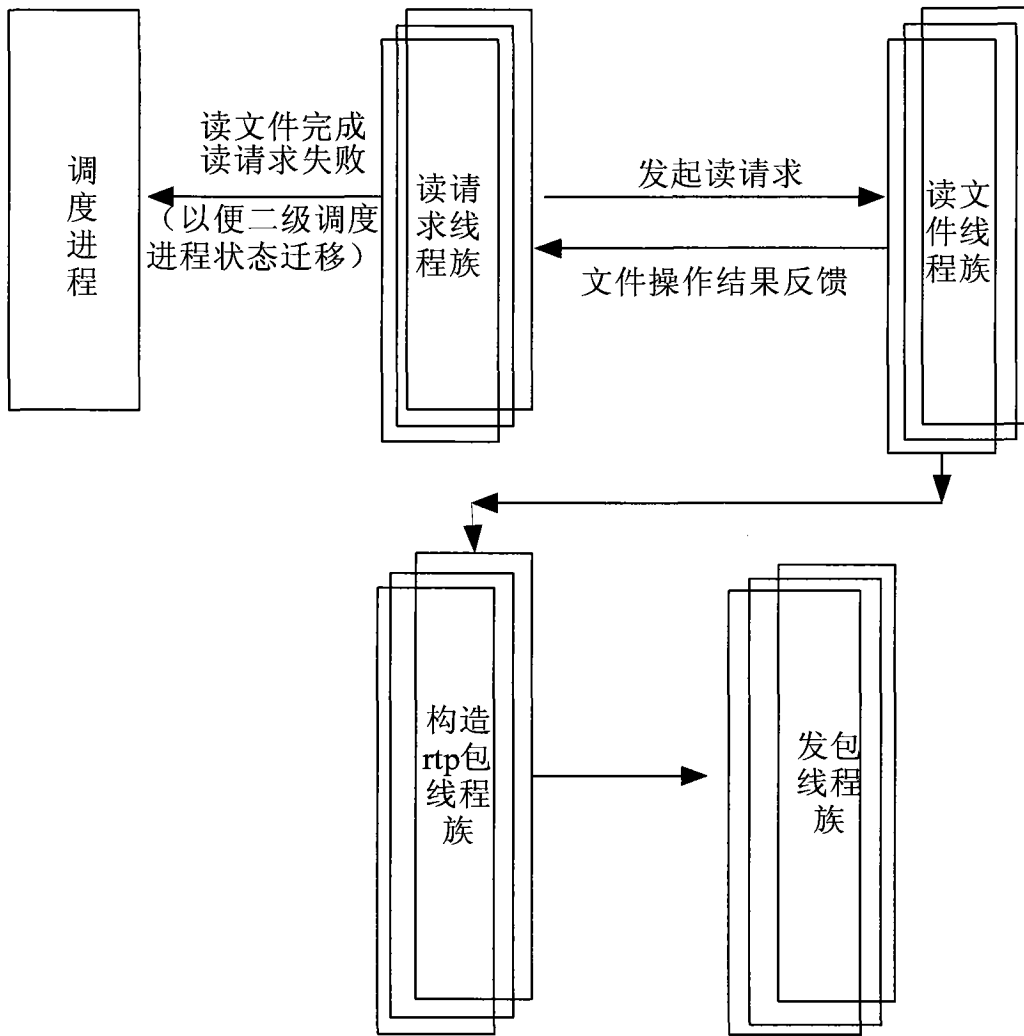


图 5

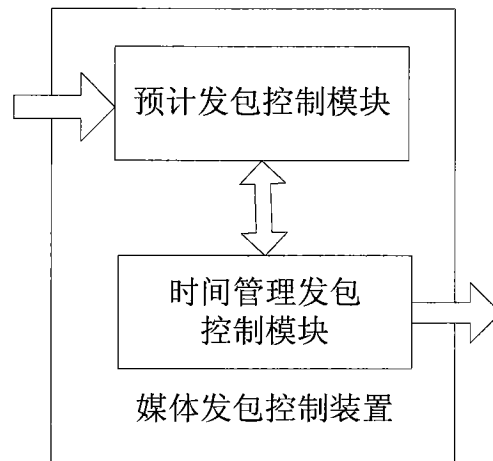


图 6

701确定要发送的媒体报文的预计发送时间并将所述媒体报文挂接到与所述预计发送时间对应的指针数组元素所指的链表上；

702：根据指针数组元素索引依次读取指针数组元素所指的链表上的的媒体报文，并在所述以及发送时间指的链表上的媒体报文并在所述预计发送时间将所述媒体报文发送出去；若该指针数组元素所指的链表上没有媒体报文，则延时到该元素预计发送时间。

图 7

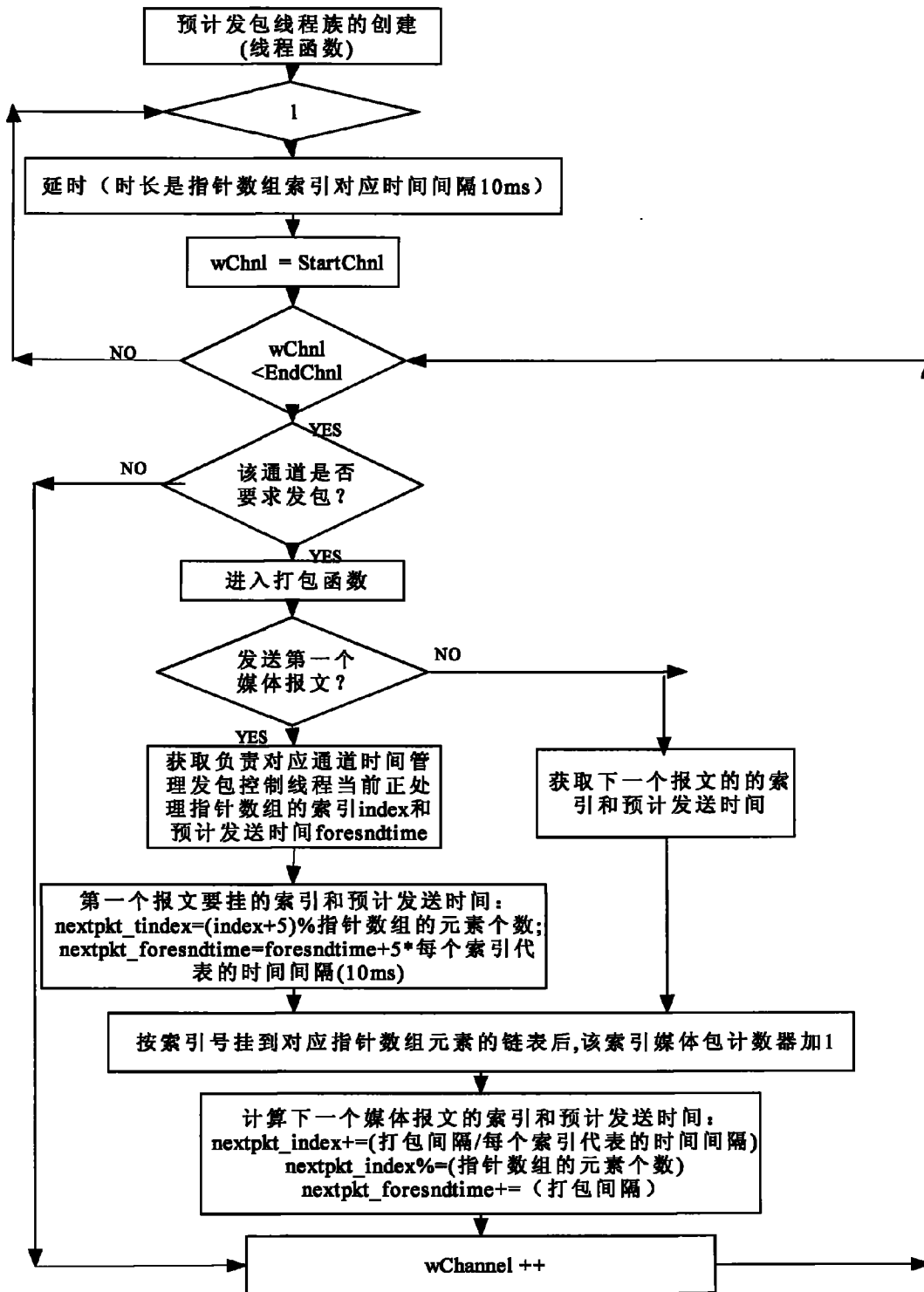


图 8

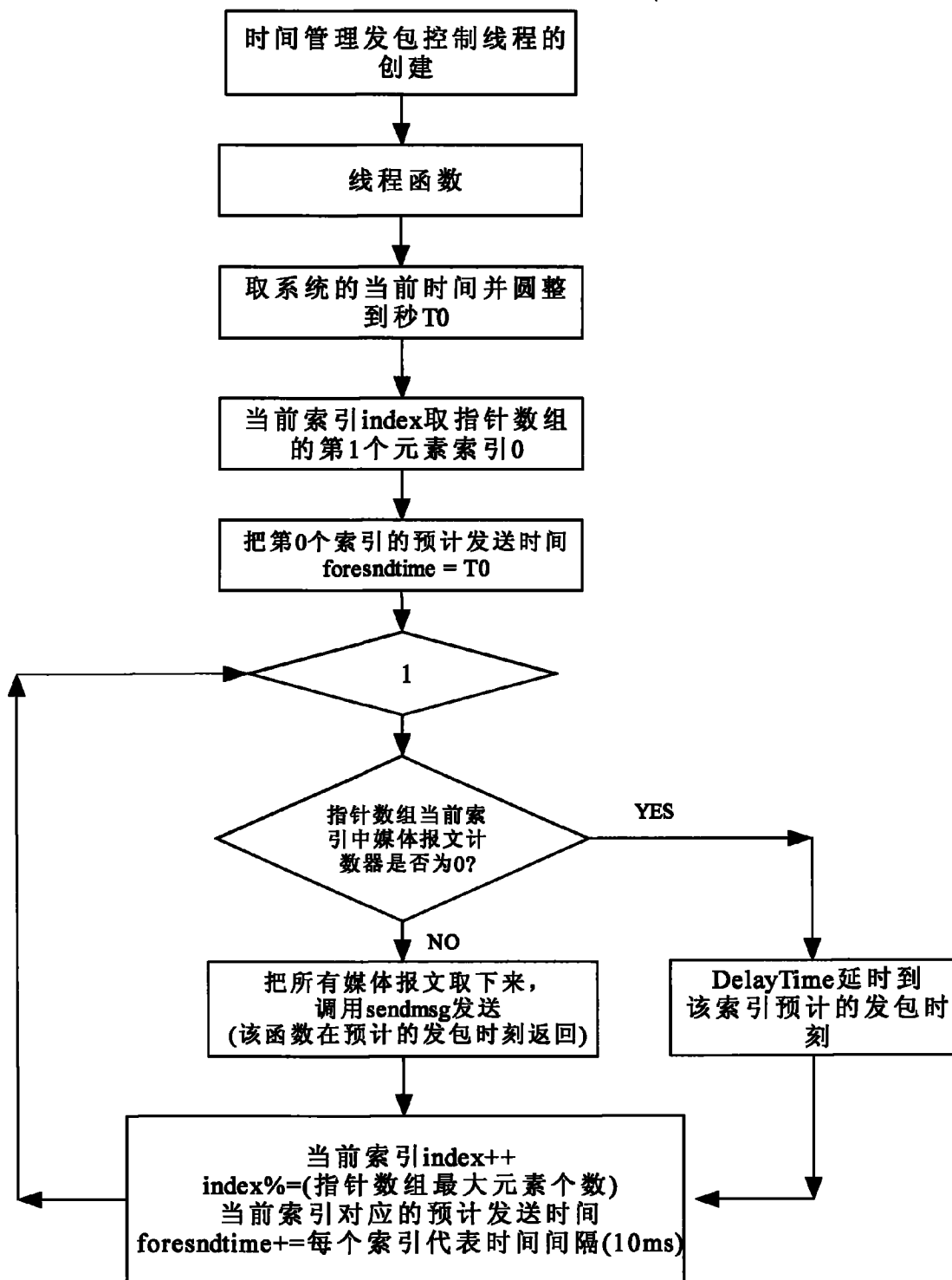


图9

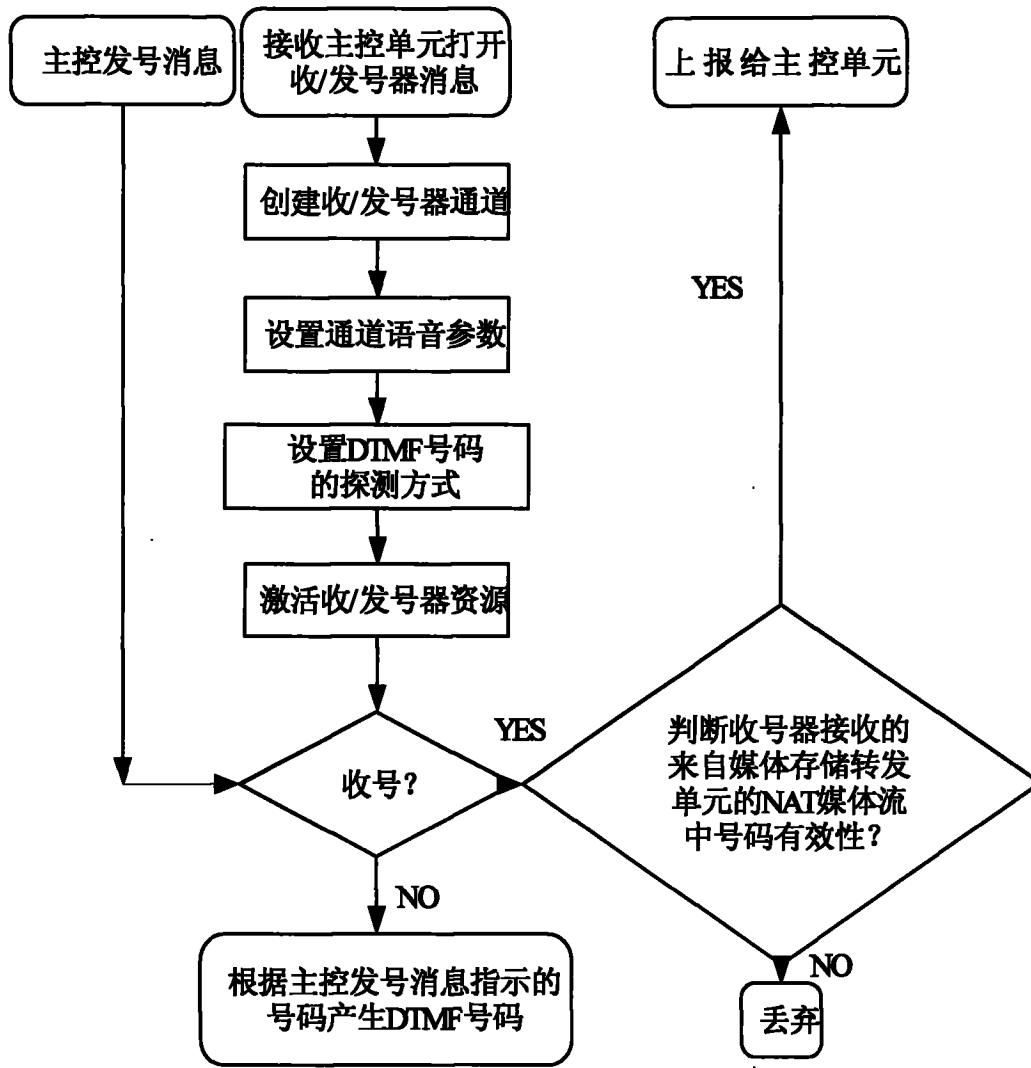


图 10

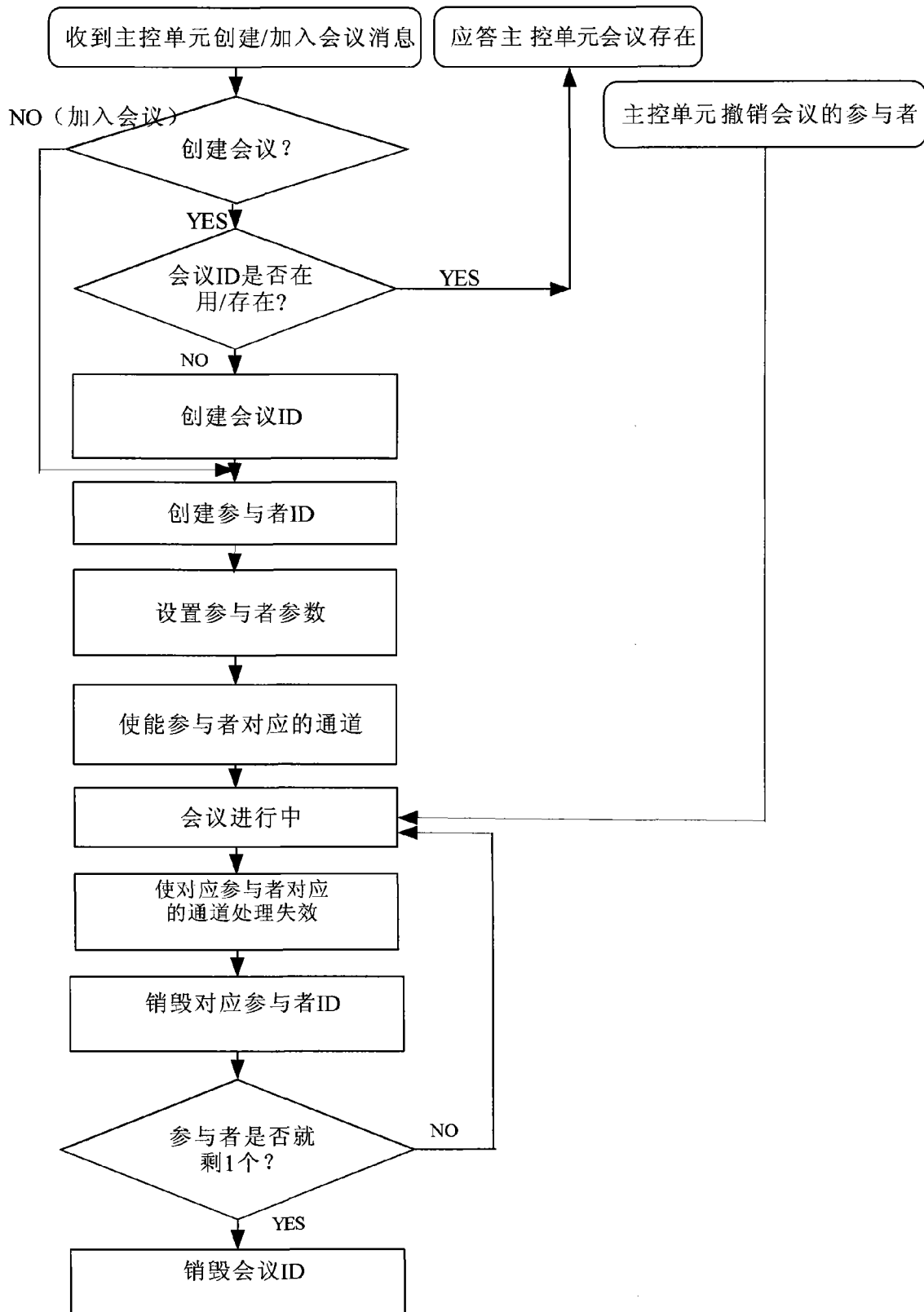


图 11