



(12) 发明专利

(10) 授权公告号 CN 111666484 B

(45) 授权公告日 2023. 10. 13

(21) 申请号 201910172262.9

(22) 申请日 2019.03.07

(65) 同一申请的已公布的文献号
申请公布号 CN 111666484 A

(43) 申请公布日 2020.09.15

(73) 专利权人 腾讯科技(深圳)有限公司
地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72) 发明人 林章楷

(74) 专利代理机构 深圳翼盛智成知识产权事务
所(普通合伙) 44300
专利代理师 黄威

(51) Int. Cl.
G06F 16/9535 (2019.01)
G06F 16/957 (2019.01)

(56) 对比文件

- CN 107451831 A, 2017.12.08
- CN 104247441 A, 2014.12.24
- CN 105608627 A, 2016.05.25
- CN 108021639 A, 2018.05.11
- WO 2018153271 A1, 2018.08.30
- CN 108337172 A, 2018.07.27
- CN 101262497 A, 2008.09.10

审查员 李雪晴

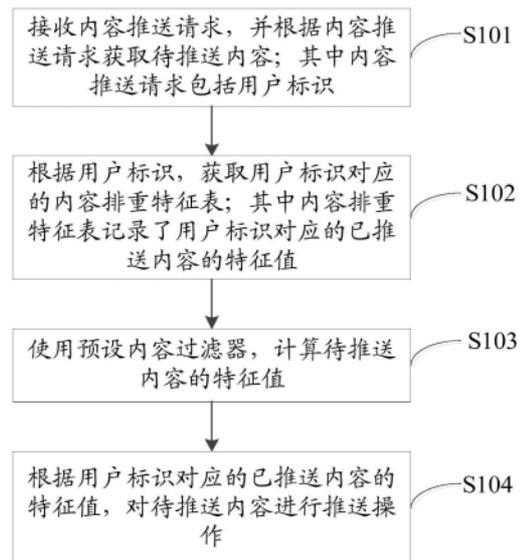
权利要求书3页 说明书13页 附图5页

(54) 发明名称

内容推送方法、推送装置及计算机可读存储
介质

(57) 摘要

本发明提供一种内容推送方法,其包括:接收内容推送请求,并根据内容推送请求获取待推送内容;根据用户标识,获取用户标识对应的内容排重特征表;使用预设内容过滤器,计算待推送内容的特征值;当用户标识对应的已推送内容的特征值不包括待推送内容的特征值时,对待推送内容进行推送操作。本发明还提供一种内容推送装置,本发明的内容推送方法及装置通过内容排重特征表记录已推送内容的特征值,从而可有效快捷的对已推送内容进行识别,有效的对推送内容进行有效去重。



1. 一种内容推送方法,其特征在于,包括:

接收内容推送请求,并根据所述内容推送请求获取待推送内容;其中所述内容推送请求包括用户标识;

根据所述用户标识,获取所述用户标识对应的内容排重特征表,其中所述内容排重特征表记录了所述用户标识对应的已推送内容的特征值;

使用预设内容过滤器,计算所述待推送内容的特征值;

根据所述用户标识对应的已推送内容的特征值,对所述待推送内容进行推送操作;

其中,所述内容排重特征表通过以下步骤更新:

获取所述用户标识对应的已推送内容以及所述用户标识对应的内容排重特征表,所述内容排重特征表包括一个内容排重特征主表以及至少一个内容排重特征次表;

使用预设内容过滤器,计算所述用户标识对应的已推送内容的特征值;

使用所述用户标识对应的已推送内容的特征值,对所述用户标识对应的内容排重特征表进行更新操作,包括:

使用所述用户标识对应的已推送内容的特征值,对所述用户标识对应的内容排重特征主表进行更新操作;

当更新后的内容排重特征主表满足预设条件,对最晚更新时间最早的内容排重特征次表进行复位操作;

将所述更新后的内容排重特征主表设置为内容排重特征次表,将复位操作后的内容排重特征次表设置为内容排重特征主表。

2. 根据权利要求1所述的内容推送方法,其特征在于,所述内容推送方法还包括:

当所述用户标识对应的已推送内容的特征值不包括所述待推送内容的特征值时,对所述待推送内容进行推送操作;

当所述用户标识对应的已推送内容的特征值包括所述待推送内容的特征值时,则重新获取待推送内容,并返回计算所述待推送内容的特征值的步骤。

3. 根据权利要求2所述的内容推送方法,其特征在于,所述预设内容过滤器为包含多个哈希函数的布隆过滤器,所述内容排重特征表包括多个特征位,所述特征位包括已推送特征值以及未推送特征值;

所述使用预设内容过滤器,计算所述待推送内容的特征值的步骤包括:

使用所述布隆过滤器的哈希函数计算所述待推送内容的特征值在所述内容排重特征表的特征位;

所述当所述用户标识对应的已推送内容的特征值不包括所述待推送内容的特征值时,对所述待推送内容进行推送操作的步骤包括:

当所述待推送内容的任意一个特征值在所述内容排重特征表的特征位为未推送特征值,对所述待推送内容进行推送操作;

当所述待推送内容的所有特征值在所述内容排重特征表的特征位均为已推送特征值,则重新获取待推送内容,并返回计算所述待推送内容的特征值的步骤。

4. 根据权利要求1所述的内容推送方法,其特征在于,所述预设内容过滤器为包含多个哈希函数的布隆过滤器,所述内容排重特征表包括多个特征位,所述特征位包括已推送特征值以及未推送特征值;

所述使用预设内容过滤器,计算所述用户标识对应的已推送内容的特征值的步骤包括:

使用所述布隆过滤器的哈希函数计算所述用户标识对应的已推送内容的特征值在所述内容排重特征表的特征位;

所述使用所述用户标识对应的已推送内容的特征值,对所述用户标识对应的内容排重特征表进行更新操作的步骤包括:

将所述已推送内容的特征值在内容排重特征表的特征位设定为已推送特征值。

5. 根据权利要求1所述的内容推送方法,其特征在于,所述当更新后的内容排重特征主表满足预设条件,对最晚更新时间最早的内容排重特征次表进行复位操作的步骤包括:

当更新后的内容排重特征主表中设定数量的特征位为已推送特征值,对最晚更新时间最早的内容排重特征次表进行复位操作。

6. 根据权利要求1所述的内容推送方法,其特征在于,所述内容推送方法还包括:

如具有内容排重特征表的用户数量大于设定值时,对使用频率最低的内容排重特征表进行删除操作或转存操作。

7. 根据权利要求1所述的内容推送方法,其特征在于,所述内容推送方法还包括:

将最近未访问时间大于设定值的内容排重特征表进行删除操作或转存操作。

8. 一种内容推送装置,其特征在于,包括:

第一待推送内容获取模块,用于接收内容推送请求,并根据所述内容推送请求获取待推送内容;其中所述内容推送请求包括用户标识;

排重特征表获取模块,用于根据所述用户标识,获取所述用户标识对应的内容排重特征表,其中所述内容排重特征表记录了所述用户标识对应的已推送内容的特征值;

待推送内容特征值计算模块,用于使用预设内容过滤器,计算所述待推送内容的特征值;

内容推送模块,用于根据所述用户标识对应的已推送内容的特征值,对所述待推送内容进行推送操作;

已推送内容获取模块,用于获取所述用户标识对应的已推送内容以及所述用户标识对应的内容排重特征表,所述内容排重特征表包括一个内容排重特征主表以及至少一个内容排重特征次表;

已推送内容特征值计算模块,用于使用预设内容过滤器,计算所述用户标识对应的已推送内容的特征值;

排重特征表更新模块,用于使用所述用户标识对应的已推送内容的特征值,对所述用户标识对应的内容排重特征表进行更新操作,包括:

使用所述用户标识对应的已推送内容的特征值,对所述用户标识对应的内容排重特征主表进行更新操作;

当更新后的内容排重特征主表满足预设条件,对最晚更新时间最早的内容排重特征次表进行复位操作;

将所述更新后的内容排重特征主表设置为内容排重特征次表,将复位操作后的内容排重特征次表设置为内容排重特征主表。

9. 根据权利要求8所述的内容推送装置,其特征在于,所述内容推送模块具体用于当所

述用户标识对应的已推送内容的特征值不包括所述待推送内容的特征值时,对所述待推送内容进行推送操作;

所述内容推送装置还包括:

第二待推送内容获取模块,用于当所述用户标识对应的已推送内容的特征值包括所述待推送内容的特征值时,则重新获取待推送内容。

10. 根据权利要求9所述的内容推送装置,其特征在于,所述预设内容过滤器为包含多个哈希函数的布隆过滤器,所述内容排重特征表包括多个特征位,所述特征位包括已推送特征值以及未推送特征值;

所述待推送内容特征值计算模块具体用于使用所述布隆过滤器的哈希函数计算所述待推送内容的特征值在所述内容排重特征表的特征位;

所述内容推送模块具体用于当所述待推送内容的任意一个特征值在所述内容排重特征表的特征位为未推送特征值,对所述待推送内容进行推送操作;

所述第二待推送内容获取模块具体用于当所述待推送内容的所有特征值在所述内容排重特征表的特征位均为已推送特征值,则重新获取待推送内容。

11. 根据权利要求8所述的内容推送装置,其特征在于,所述预设内容过滤器为包含多个哈希函数的布隆过滤器,所述内容排重特征表包括多个特征位,所述特征位包括已推送特征值以及未推送特征值;

所述已推送内容特征值计算模块具体用于使用所述布隆过滤器的哈希函数计算所述用户标识对应的已推送内容的特征值在所述内容排重特征表的特征位;

所述排重特征表更新模块具体用于将所述已推送内容的特征值在内容排重特征表的特征位设定为已推送特征值。

12. 一种计算机可读存储介质,其内存储有处理器可执行指令,所述指令由一个或一个以上处理器加载,以执行如权利要求1-7中任一的内容推送方法。

13. 一种电子设备,其特征在于,包括处理单元和存储器,所述存储器存储有多条指令;所述处理单元从所述存储器中加载指令,以执行如权利要求1-7中任一的内容推送方法。

内容推送方法、推送装置及计算机可读存储介质

技术领域

[0001] 本发明涉及内容推送,特别是涉及一种内容推送方法、推送装置及计算机可读存储介质。

背景技术

[0002] 随着互联网技术的发展,各种互联网多媒体内容的推送越来越多,比如抖音、微视等短视频应用的短视频推送,极大丰富了用户的业余生活,并且吸引到了大量的用户观看以及点赞。

[0003] 作为各种推送内容应用的用户,希望看到的是感兴趣且新鲜的推送内容;作为各种推送内容应用的服务商,希望尽可能将用户感兴趣的内容推送至用户,并且尽量不让用户接收到相同的推送内容。

[0004] 因此需要一种可对推送内容进行有效去重的内容推送方法及内容推送装置。

发明内容

[0005] 本发明实施例提供一种可对推送内容进行有效去重的内容推送方法及内容推送装置;以解决现有的内容推送方法及内容推送装置的不能对推送内容进行有效去重的技术问题。

[0006] 本发明实施例提供一种内容推送方法,其包括:

[0007] 接收内容推送请求,并根据所述内容推送请求获取待推送内容;其中所述内容推送请求包括用户标识;

[0008] 根据所述用户标识,获取所述用户标识对应的内容排重特征表,其中所述内容排重特征表记录了所述用户标识对应的已推送内容的特征值;

[0009] 使用预设内容过滤器,计算所述待推送内容的特征值;

[0010] 根据所述用户标识对应的已推送内容的特征值,对所述待推送内容进行推送操作。

[0011] 本发明实施例提供一种内容推送装置,其包括:

[0012] 第一待推送内容获取模块,用于接收内容推送请求,并根据所述内容推送请求获取待推送内容;其中所述内容推送请求包括用户标识;

[0013] 排重特征表获取模块,用于根据所述用户标识,获取所述用户标识对应的内容排重特征表,其中所述内容排重特征表记录了所述用户标识对应的已推送内容的特征值;

[0014] 待推送内容特征值计算模块,用于使用预设内容过滤器,计算所述待推送内容的特征值;

[0015] 内容推送模块,用于根据所述用户标识对应的已推送内容的特征值,对所述待推送内容进行推送操作。

[0016] 在本发明实施例所述的内容推送装置中,所述内容排重特征表包括一个内容排重特征主表以及至少一个内容排重特征次表;

[0017] 所述排重特征表更新模块包括：

[0018] 排重特征主表更新单元，用于使用所述用户标识对应的已推送内容的特征值，对所述用户标识对应的内容排重特征主表进行更新操作；

[0019] 排重特征次表复位单元，用于当更新后的内容排重特征主表满足预设条件，对最晚更新时间最早的内容排重特征次表进行复位操作；

[0020] 排重特征主表设置单元，用于将所述更新后的内容排重特征主表设置为内容排重特征次表，将复位操作后的内容排重特征次表设置为内容排重特征主表。

[0021] 在本发明实施例所述的内容推送装置中，所述排重特征次表复位单元具体用于，当更新后的内容排重特征主表中设定数量的特征位为已推送特征值，对最晚更新时间最早的内容排重特征次表进行复位操作。

[0022] 在本发明实施例所述的内容推送装置中，所述内容推送装置还包括：

[0023] 内容删除或转存模块，用于如具有内容排重特征表的用户数量大于设定值时，对使用频率最低的内容排重特征表进行删除操作或转存操作；和将最近未访问时间大于设定值的内容排重特征表进行删除操作或转存操作。

[0024] 本发明实施例还提供一种计算机可读存储介质，其内存储有处理器可执行指令，所述指令由一个或一个以上处理器加载，以执行上述的内容推送方法。

[0025] 相较于现有技术，本发明的内容推送方法及装置通过内容排重特征表记录已推送内容的特征值，从而可有效快捷的对已推送内容进行识别，有效的对推送内容进行有效去重；解决了现有的内容推送方法及内容推送装置的不能对推送内容进行有效去重的技术问题。

附图说明

[0026] 图1为本发明的内容推送方法的第一实施例的流程图；

[0027] 图2为本发明的内容推送方法的第二实施例的流程图；

[0028] 图3为本发明的内容推送方法的第三实施例的流程图；

[0029] 图4为本发明的内容推送装置的第一实施例的结构示意图；

[0030] 图5为本发明的内容推送装置的第二实施例的结构示意图；

[0031] 图6为本发明的内容推送装置的第三实施例的结构示意图；

[0032] 图7a为本发明的内容推送方法及内容推送装置的具体实施例对应的内容推送服务器的结构示意图；

[0033] 图7b为本发明的内容推送方法及内容推送装置的具体实施例的内容推送过程的流程图；

[0034] 图8为本发明的内容推送装置所在的电子设备的工作环境结构示意图。

具体实施方式

[0035] 请参照图式，其中相同的组件符号代表相同的组件，本发明的原理是以实施在一适当的运算环境中来举例说明。以下的说明是基于所例示的本发明具体实施例，其不应被视为限制本发明未在此详述的其它具体实施例。

[0036] 在以下的说明中，本发明的具体实施例将参考由一部或多部计算机所执行之作业

的步骤及符号来说明,除非另有说明。因此,其将可了解到这些步骤及操作,其中有数次提到为由计算机执行,包括了由代表了以一结构化型式中的数据之电子信号的计算机处理单元所操纵。此操纵转换该数据或将其维持在该计算机之内存系统中的位置处,其可重新配置或另外以本领域技术人员所熟知的方式来改变该计算机之运作。该数据所维持的数据结构为该内存之实体位置,其具有由该数据格式所定义的特定特性。但是,本发明原理以上述文字来说明,其并不代表为一种限制,本领域技术人员将可了解到以下所述的多种步骤及操作亦可实施在硬件当中。

[0037] 本发明的内容推送方法及内容推送装置可设置在任何的电子设备中,用于对多媒体内容进行过滤以及推送操作。该电子设备包括但不限于可穿戴设备、头戴设备、医疗健康平台、个人计算机、服务器计算机、手持式或膝上型设备、移动设备(比如移动电话、个人数字助理(PDA)、媒体播放器等等)、多处理器系统、消费型电子设备、小型计算机、大型计算机、包括上述任意系统或设备的分布式计算环境,等等。该电子设备优选为进行多媒体内容推送的多媒体推送服务器,其可对推送的多媒体内容进行有效的去重操作。

[0038] 请参照图1,图1为本发明的内容推送方法的第一实施例的流程图。本实施例的内容推送方法可使用上述的电子设备进行实施,本实施例的内容推送方法包括:

[0039] 步骤S101,接收内容推送请求,并根据内容推送请求获取待推送内容;其中内容推送请求包括用户标识;

[0040] 步骤S102,根据用户标识,获取用户标识对应的内容排重特征表;其中内容排重特征表记录了用户标识对应的已推送内容的特征值;

[0041] 步骤S103,使用预设内容过滤器,计算待推送内容的特征值;

[0042] 步骤S104,根据用户标识对应的已推送内容的特征值,对待推送内容进行推送操作。

[0043] 下面详细说明本实施例的内容推送方法的各步骤的具体流程。

[0044] 在步骤S101,内容推送装置,如多媒体内容推送服务器接收到其他终端的内容推送请求。该内容推送请求为该终端需要从内容推送装置获取符合其要求的待推送内容的请求,由于不同的用户会有需要推送不同待推送内容的内容推送请求,因此该内容推送请求应包括可以区分不同终端或用户的用户标识,以保证可获取较为准确的待推送内容。

[0045] 随后内容推送装置根据获取的内容推送请求中的要求,获取对应的待推送内容。该待推送内容可为多媒体文字内容,多媒体音频内容或多媒体视频内容等。如推送小说、推送歌曲或推送在线短视频等。

[0046] 在步骤S102中,内容推送装置根据步骤S101获取的内容推送请求对应的用户标识,获取该用户标识对应的内容排重特征表。

[0047] 该内容排重特征表包括多个特征位,每个特征位可为已推送特征值或未推送特征值(如已推送特征值为1,未推送特征值为0等),即内容排重特征表可由多个二级制数值组成的数列(二进制向量),如001001110101……等等。这样内容排重特征表的所有特征位可用来记录用户标识对应的已推送内容的特征值,内容排重特征表的特征位的数量可根据用户的需要进行设定。

[0048] 在步骤S103中,内容推送装置使用预设内容过滤器,计算步骤S101获取的待推送内容的特征值。

[0049] 预设内容过滤器为包含多个哈希函数的布隆过滤器。这里的哈希函数用于将待推送内容的标识(如视频vid标识)等元素映射为二进制向量中的一个位数值,多个哈希函数则可将待推送内容的标识映射为二进制向量中的多个位数值。这里的布隆过滤器包含多个哈希函数以及一个二进制向量,用于检测待推送内容的标识是否在一个二进制向量组成的集合中。

[0050] 具体的,内容推送装置使用布隆过滤器的多个哈希函数计算待推送内容的特征值在内容排重特征表的特征位,即二进制向量中的特征位。如布隆过滤器包括8个哈希函数,则每个待推送内容通过8个哈希函数可分别计算出8个特征位,即每个待推送内容均使用内容排重特征表中的8个特征位。

[0051] 在步骤S104中,当步骤S102获取的内容排重特征表中的用户标识对应的已推送内容的特征值不包括步骤S103中获取的待推送内容的特征值时,即步骤S103中计算的待推送内容的任意一个特征值在内容排重特征表的特征位为未推送特征值,说明该待推送内容没有被推送过,则对该待推送内容进行推送操作。

[0052] 当步骤S102获取的内容排重特征表中的用户标识对应的已推送内容的特征值包括步骤S103中获取的待推送内容的特征值时,即步骤S103计算的待推送内容的所有特征值在内容排重特征表的特征位为已推送特征值,说明该待推送内容已被推送过,则重新获取待推送内容,并返回步骤S103计算重新获取的待推送内容的特征值。

[0053] 这样即完成了本实施例的内容推送方法的待推送内容的过滤推送过程。

[0054] 本实施例的内容推送方法通过内容排重特征表记录已推送内容的特征值,从而可有效快捷的对已推送内容进行识别,有效的对推送内容进行有效去重。

[0055] 请参照图2,图2为本发明的内容推送方法的第二实施例的流程图。本实施例的内容推送方法可使用上述的电子设备进行实施,在第一实施例的基础上,本实施例的内容推送方法可实现对内容排重特征表进行有效的更新,从而有效的提高内容推送的有效性。该内容排重特征表的更新方法包括:

[0056] 步骤S201,获取用户标识对应的已推送内容以及用户标识对应的内容排重特征表;

[0057] 步骤S202,使用预设内容过滤器,计算用户标识对应的已推送内容的特征值;

[0058] 步骤S203,使用用户标识对应的已推送内容的特征值,对用户标识对应的内容排重特征表进行更新操作。

[0059] 下面详细说明本实施例的内容排重特征表的更新方法的各步骤的具体流程。

[0060] 在进行内容推送之后,内容推送装置需要对内容排重特征表进行更新,从而保证后续内容推送的有效性,具体的:

[0061] 在步骤S201中,内容推送装置获取用户标识对应的已推送内容以及用户标识对应的内容排重特征表,以便使用已推送内容对该内容排重特征表进行更新操作。

[0062] 在步骤S202中,内容推送装置使用预设内容过滤器,计算用户标识对应的已推送内容的特征值。预设内容过滤器为包含多个哈希函数的布隆过滤器。

[0063] 具体的,内容推送装置使用布隆过滤器的多个哈希函数计算待推送内容的特征值在内容排重特征表的特征位,即二进制向量中的特征位。

[0064] 在步骤S203中,内容推送装置使用步骤S202获取的用户标识对应的已推送内容的

特征值,对步骤S201获取的用户标识对应的内容排重特征表进行更新操作。

[0065] 具体的,内容推送装置将已推送内容的特征值在内容排重特征表的特征位设定为已推送特征值,即如该特征值在内容排重特征表的特征位为未推送特征值,则将特征位改为已推送特征值;如该特征值在内容排重特征表的特征位为已推送特征值,则不需要对该特征位进行修改。

[0066] 这样即完成了本实施例的内容推送后的内容排重特征表的更新操作。

[0067] 在第一实施例的基础上,本实施例的内容排重特征表的更新方法可通过已推送内容的特征值对内容排重特征表进行实时更新,进一步提高了内容推送的有效性。

[0068] 请参照图3,图3为本发明的内容推送方法的第三实施例的流程图。本实施例的内容推送方法可使用上述的电子设备进行实施,在第一实施例的基础上,本实施例的内容推送方法可实现对内容排重特征表进行有效的更新,从而有效的提高内容推送的有效性。

[0069] 由于随着多个已推送内容的特征值在内容排重特征表的二进制向量中的叠加,会使得该二进制向量中的越来越多的特征位为已推送特征值,由于任何推送内容均占用二进制向量中的n个特征位(如使用n个哈希函数),如二进制向量中的越来越多的特征位为已推送特征值,可能会导致错误的将某个未推送内容判断为已推送内容。

[0070] 因此这里需要对内容排重特征表定期进行复位操作,本实施例的方案可有效的保证内容排重特征表的复位操作,且又能保证内容排重特征表对已推送内容的有效过滤。

[0071] 在本实施例中,内容排重特征表包括一个内容排重特征主表以及至少一个内容排重特征次表。本实施例中对内容排重特征的更新操作的步骤包括:

[0072] 步骤S301,使用用户标识对应的已推送内容的特征值,对用户标识对应的内容排重特征主表进行更新操作;

[0073] 步骤S302,当更新后的内容排重特征主表满足预设条件,对最晚更新时间最早的内容排重特征次表进行复位操作;

[0074] 步骤S303,将更新后的内容排重特征主表设置为内容排重特征次表,将复位操作后的内容排重特征次表设置为内容排重特征主表。

[0075] 下面详细说明本实施例的内容排重特征表的更新方法的各步骤的具体流程。

[0076] 在步骤S301中,内容推送装置使用用户标识对应的已推送内容的特征值,对用户标识对应的内容排重特征主表进行更新操作。

[0077] 具体的,内容推送装置将已推送内容的特征值在内容排重特征主表的特征位设定为已推送特征值,即如该特征值在内容排重特征主表的特征位为未推送特征值,则将特征位改为已推送特征值;如该特征值在内容排重特征主表的特征位为已推送特征值,则不需要对该特征位进行修改。

[0078] 在步骤S302中,内容推送装置判断更新后的内容排重特征主表是否满足预设条件,即更新后的内容排重特征主表的二进制向量中设定数量的特征位是否为己推送特征值,这个设定数量可为40%至60%。

[0079] 如更新后的内容排重特征主表中设定数量的特征位为己推送特征值,则对最晚更新时间最早的内容排重特征次表进行复位操作,以便停止使用内容排重特征主表继续记录已推送内容。

[0080] 这里如有多个内容排重特征次表,这里的每个内容排重特征次表均默认为对应的

二进制向量中的设定数量的特征位已经为已推送特征值了。由于最晚更新时间最早的内容排重特征次表满足预设条件的较早,即该内容排重特征次表对应的已推送内容的推送时间较早,因此对该内容排重特征次表进行复位操作对内容排重特征表的影响最小。

[0081] 在步骤S303中,内容推送装置将步骤S301中进行更新操作的内容排重特征主表设置为内容排重特征次表,以避免继续记录已推送内容;将步骤S302中复位操作后的内容排重特征次表设置为内容排重特征主表,以便继续记录已推送内容的特征值。

[0082] 这样即完成了本实施例的内容推送后的内容排重特征表的更新操作。

[0083] 此外,由于每个用户均需要使用用户标识占用一个单独内容排重特征表,因此随着用户的增加,内容排重特征表占用的资源也会越来越多,因此当具有内容排重特征表的用户数量大于设定值时,内容推送装置对使用频率最低的内容排重特征表进行删除或转存操作。这里的删除操作是指直接将该用户的内容排重特征表删除,这里的转存操作是指将该用户的内容排重特征表作为冷数据进行存储操作。当然这里也可将最近未访问时间大于设定值的内容排重特征表进行删除操作或转存操作,从而有效的节省访问资源。

[0084] 在第一实施例的基础上,本实施例的内容排重特征表的更新以及删除方法即可对内容排重特征表进行实时更新,又可在保证对已推送内容有效过滤的基础上,对内容排重特征表进行实时复位;即提高了内容推送的有效性,又节省了内容推送信息的存储访问资源。

[0085] 本发明还提供一种内容推送装置,请参照图4,图4为本发明的内容推送装置的第一实施例的结构示意图。本实施例的内容推送装置可使用上述的内容推送方法的第一实施例进行实施,该内容推送装置40包括第一待推送内容获取模块41、排重特征表获取模块42、待推送内容特征值计算模块43、内容推送模块44以及第二待推送内容获取模块45。

[0086] 第一待推送内容获取模块41用于接收内容推送请求,并根据内容推送请求获取待推送内容;其中内容推送请求包括用户标识;排重特征表获取模块42用于根据用户标识,获取用户标识对应的内容排重特征表,其中内容排重特征表记录了用户标识对应的已推送内容的特征值;待推送内容特征值计算模块43用于使用预设内容过滤器,计算待推送内容的特征值;内容推送模块44用于当用户标识对应的已推送内容的特征值不包括所述待推送内容的特征值时,对待推送内容进行推送操作;第二待推送内容获取模块45用于当用户标识对应的已推送内容的特征值包括待推送内容的特征值时,则重新获取待推送内容。

[0087] 本实施例的内容推送装置40使用时,第一待推送内容获取模块41接收到其他终端的内容推送请求。该内容推送请求为该终端需要从内容推送装置获取符合其要求的待推送内容的请求,由于不同的用户会有需要推送不同待推送内容的内容推送请求,因此该内容推送请求应包括可以区分不同终端或用户的用户标识,以保证可获取较为准确的待推送内容。

[0088] 随后第一待推送内容获取模块41根据获取的内容推送请求中的要求,获取对应的待推送内容。该待推送内容可为多媒体文字内容,多媒体音频内容或多媒体视频内容等。如推送小说、推送歌曲或推送在线短视频等。

[0089] 然后排重特征表获取模块42根据第一待推送内容获取模块41获取的内容推送请求对应的用户标识,获取该用户标识对应的内容排重特征表。

[0090] 该内容排重特征表包括多个特征位,每个特征位可为已推送特征值或未推送特征

值(如已推送特征值为1,未推送特征值为0等),即内容排重特征表可由多个二级制数值组成的数列,如001001110101……等等。这样内容排重特征表的所有特征位可用来记录用户标识对应的已推送内容的特征值,内容排重特征表的特征位的数量可根据用户的需要进行设定。

[0091] 随后待推送内容特征值计算模块43使用预设内容过滤器,计算第一待推送内容获取模块获取的待推送内容的特征值。

[0092] 预设内容过滤器为包含多个哈希函数的布隆过滤器。这里的哈希函数用于将待推送内容的标识(如视频vid标识)等元素映射为二进制向量中的一个位数值,多个哈希函数则可将待推送内容的标识映射为二进制向量中的多个位数值。这里的布隆过滤器包含多个哈希函数以及一个二进制向量,用于检测待推送内容的标识是否在一个二级制向量组成的集合中。

[0093] 具体的,待推送内容特征值计算模块43使用布隆过滤器的多个哈希函数计算待推送内容的特征值在内容排重特征表的特征位,即二进制向量中的特征位。如布隆过滤器包括8个哈希函数,则每个待推送内容通过8个哈希函数可分别计算出8个特征位,即每个待推送内容均使用内容排重特征表中的8个特征位。

[0094] 最后,当排重特征表获取模块42获取的内容排重特征表中的用户标识对应的已推送内容的特征值不包括待推送内容特征值计算模块43获取的待推送内容的特征值时,即待推送内容特征值计算模块43计算的待推送内容的任意一个特征值在内容排重特征表的特征位为未推送特征值,说明该待推送内容没有被推送过,则内容推送模块44对该待推送内容进行推送操作。

[0095] 当排重特征表获取模块42获取的内容排重特征表中的用户标识对应的已推送内容的特征值包括待推送内容特征值计算模块43获取的待推送内容的特征值时,即待推送内容特征值计算模块43计算的待推送内容的所有特征值在内容排重特征表的特征位为已推送特征值,说明该待推送内容已被推送过,则第二待推送内容获取模块45重新获取待推送内容,并返回待推送内容特征值计算模块43计算重新获取的待推送内容的特征值。

[0096] 这样即完成了本实施例的内容推送装置40的待推送内容的过滤推送过程。

[0097] 本实施例的内容推送装置通过内容排重特征表记录已推送内容的特征值,从而可有效快捷的对已推送内容进行识别,有效的对推送内容进行有效去重。

[0098] 请参照图5,图5为本发明的内容推送装置的第二实施例的结构示意图。本实施例的内容推送装置可使用上述的内容推送方法的第二实施例进行实施,本实施例的内容推送装置可实现对内容排重特征表进行有效的更新,从而有效的提高内容推送的有效性。该内容推送装置50还包括已推送内容获取模块51、已推送内容特征值计算模块52以及排重特征表更新模块53。

[0099] 已推送内容获取模块51用于获取用户标识对应的已推送内容以及用户标识对应的内容排重特征表;已推送内容特征值计算模块52用于使用预设内容过滤器,计算用户标识对应的已推送内容的特征值;排重特征表更新模块53用于使用用户标识对应的已推送内容的特征值,对用户标识对应的内容排重特征表进行更新操作。

[0100] 在进行内容推送之后,内容推送装置50需要对内容排重特征表进行更新,从而保证后续内容推送的有效性,具体的:

[0101] 首先已推送内容获取模块51获取用户标识对应的已推送内容以及用户标识对应的内容排重特征表,以便使用已推送内容对该内容排重特征表进行更新操作。

[0102] 随后已推送内容特征值计算模块52使用预设内容过滤器,计算用户标识对应的已推送内容的特征值。预设内容过滤器为包含多个哈希函数的布隆过滤器。

[0103] 具体的,已推送内容特征值计算模块52使用布隆过滤器的多个哈希函数计算待推送内容的特征值在内容排重特征表的特征位,即二进制向量中的特征位。

[0104] 最后排重特征表更新模块53使用已推送内容特征值计算模块52获取的用户标识对应的已推送内容的特征值,对已推送内容获取模块51获取的用户标识对应的内容排重特征表进行更新操作。

[0105] 具体的,排重特征表更新模块53将已推送内容的特征值在内容排重特征表的特征位设定为已推送特征值,即如该特征值在内容排重特征表的特征位为未推送特征值,则将特征位改为已推送特征值;如该特征值在内容排重特征表的特征位为已推送特征值,则不需要对该特征位进行修改。

[0106] 这样即完成了本实施例的内容推送后的内容排重特征表的更新操作。

[0107] 在第一实施例的基础上,本实施例的内容推送装置可通过已推送内容的特征值对内容排重特征表进行实时更新,进一步提高了内容推送的有效性。

[0108] 请参照图6,图6为本发明的内容推送装置的第三实施例的结构示意图。本实施例的内容推送装置可使用上述的内容推送方法的第三实施例进行实施,在第一实施例的基础上,本实施例的内容推送装置可实现对内容排重特征表进行有效的更新,从而有效的提高内容推送的有效性。

[0109] 本实施例的内容推送装置的排重特征表更新模块63包括排重特征主表更新单元631、排重特征次表复位单元632以及排重特征主表设置单元633。

[0110] 排重特征主表更新单元631用于使用用户标识对应的已推送内容的特征值,对用户标识对应的内容排重特征主表进行更新操作;排重特征次表复位单元632用于当更新后的内容排重特征主表满足预设条件,对最晚更新时间最早的内容排重特征次表进行复位操作;排重特征主表设置单元633用于将更新后的内容排重特征主表设置为内容排重特征次表,将复位操作后的内容排重特征次表设置为内容排重特征主表。

[0111] 在本实施例中,内容排重特征表包括一个内容排重特征主表以及至少一个内容排重特征次表。本实施例中对内容排重特征的更新操作的步骤包括:

[0112] 排重特征主表更新单元631使用用户标识对应的已推送内容的特征值,对用户标识对应的内容排重特征主表进行更新操作。

[0113] 具体的,排重特征主表更新单元631将已推送内容的特征值在内容排重特征主表的特征位设定为已推送特征值,即如该特征值在内容排重特征主表的特征位为未推送特征值,则将特征位改为已推送特征值;如该特征值在内容排重特征主表的特征位为已推送特征值,则不需要对该特征位进行修改。

[0114] 排重特征次表复位单元632判断更新后的内容排重特征主表是否满足预设条件,即更新后的内容排重特征主表的二进制向量中设定数量的特征位是否为已推送特征值,这个设定数量可为40%至60%。

[0115] 如更新后的内容排重特征主表中设定数量的特征位为已推送特征值,则对最晚更

新时间最早的内容排重特征次表进行复位操作,以便停止使用内容排重特征主表继续记录已推送内容。

[0116] 排重特征主表设置单元633将进行更新操作的内容排重特征主表设置为内容排重特征次表,以避免继续记录已推送内容;将复位操作后的内容排重特征次表设置为内容排重特征主表,以便继续记录已推送内容的特征值。

[0117] 这样即完成了本实施例的内容推送后的内容排重特征表的更新操作。

[0118] 此外,由于每个用户均需要使用用户标识占用一个单独内容排重特征表,因此随着用户的增加,内容排重特征表占用的资源也会越来越多,因此内容推送装置还包括内容删除或转存模块。

[0119] 当具有内容排重特征表的用户数量大于设定值时,内容删除或转存模块对使用频率最低的内容排重特征表进行删除或转存操作。这里的删除操作是指直接将该用户的内容排重特征表删除,这里的转存操作是指将该用户的内容排重特征表作为冷数据进行存储操作。当然这里内容删除或转存模块也可将最近未访问时间大于设定值的内容排重特征表进行删除操作或转存操作,从而有效的节省访问资源。

[0120] 在第一实施例的基础上,本实施例的内容推送装置即可对内容排重特征表进行实时更新,又可在保证对已推送内容有效过滤的基础上,对内容排重特征表进行实时复位;即提高了内容推送的有效性,又节省了内容推送信息的存储访问资源。

[0121] 下面通过一具体实施例说明本发明的内容推送方法及内容推送装置的具体工作原理。本具体实施例以进行短视频推送为例说明多媒体内容的过滤以及推送流程。

[0122] 请参照图7a和图7b,图7a为本发明的内容推送方法及内容推送装置的具体实施例对应的内容推送服务器的结构示意图,图7b为本发明的内容推送方法及内容推送装置的具体实施例的内容推送过程的流程图。该内容推送服务器70包括短视频内容数据库71、视频列表服务器72、视频排重服务器73以及分布式存储服务器74。

[0123] 本实施例中的视频排重服务器73中设置有与浏览短视频内容的用户75的用户标识数量对应的内容排重特征表,以及对已推送内容进行特征值计算的布隆过滤器。本实施例的内容推送服务器70的内容推送过程包括:

[0124] 步骤S701、当内容推送服务器70接收到用户75的内容推送请求时,视频列表服务器72会根据内容推送请求的需求(如请求搞笑短视频等)从短视频内容数据库71中获取一批待推送短视频内容。

[0125] 步骤S702、视频列表服务器72获取这一批待推送短视频内容的视频标识码(如视频vid码),并把待推送短视频内容的视频标识码发送至视频排重服务器73。

[0126] 步骤S703、视频排重服务器73获取内容推送请求对应的用户标识,并获取该用户标识对应的内容排重特征表。

[0127] 步骤S704、视频排重服务器73使用布隆过滤器计算所有待推送短视频内容的视频标识码对应的特征值;并在内容排重特征表中查找是否具有上述特征值,如内容排重特征表中不包括上述待推送短视频内容的视频标识码的特征值,则对待推送短视频内容进行推送操作;如内容排重特征表中包括上述待推送短视频内容的视频标识码的特征值,则不对该待推送短视频内容进行推送。

[0128] 具体的,本实施例中,视频排重服务器73使用的布隆过滤器包括8个不同的哈希函

数,内容排重特征表包括一个内容排重特征主表以及一个内容排重特征次表,内容排重特征主表和内容排重特征次表均为占用8kb的空间的二进制向量,即内容排重特征主表和内容排重特征次表均包括由已推送特征值(如1)以及未推送特征值(如0)组成的多个特征位。

[0129] 这里内容排重特征主表和内容排重特征次表占用的空间以及布隆过滤器使用的哈希函数的数量可根据用户75的要求进行配置。

[0130] 在内容排重特征表中查找待推送短视频内容的视频标识码对应的特征值的过程包括:

[0131] 将待推送短视频内容的视频标识码传入到布隆过滤器的8个哈希函数中,8个哈希函数分别会返回一个该视频标识码的特征值,即已推送特征值在内容排重特征表中的位置(内容排重特征表的初始状态所有的特征位均为未推送特征值0),如8个哈希函数返回的特征值在内容排重特征表(内容排重特征主表以及内容排重特征次表)中均为已推送特征值1,则说明该视频标识码对应的视频已推送过,则不对该待推送短视频内容进行推送;如任意1个哈希函数返回的特征值在内容排重特征表中为未推送特征值0,则说明该视频标识码对应的视频未被推送过,则对该待推送短视频内容进行推送操作。

[0132] 步骤S705、视频排重服务器73对筛选后的待推送短视频内容进行推送操作,且使用上述推送的短视频内容的特征值对该用户标识对应的内容排重特征表进行更新操作,即将推送的短视频内容的特征值在该用户标识对应的内容排重特征表中的位置设置为已推送特征值1。

[0133] 具体的,对内容排重特征表的内容排重特征主表和内容排重特征次表更新时,可首先使用推送的短视频内容的特征值对内容排重特征主表进行更新,即将推送的短视频内容的特征值在内容排重特征主表中的位置设置为已推送特征值1。当内容排重特征主表的二进制向量50%以上的特征位均为已推送特征值1时,则认为内容排重特征主表已写满。

[0134] 这时将内容排重特征次表进行复位操作,即将内容排重特征次表的二进制向量的所有特征位均设置为未推送特征值0。将更新后的内容排重特征主表设置为内容排重特征次表,将复位后的内容排重特征次表设置为内容排重特征主表。

[0135] 采用上述的内容排重特征表来记录已推送短视频内容的情况,一个用户均需要占用16kB的内存,则视频排重服务器中1GB的内存可支撑5万多活跃用户同时进行短视频推送操作。

[0136] 为了防止内容排重特征表占用内存过大,这里还可设置将超过3*24小时没有访问的内容排重特征表进行转存操作或删除操作。当内容排重特征表的总数量大于15w时,将最近访问最少的内容排重特征表进行转存操作或删除操作。

[0137] 上述从内存中转存出来的内容排重特征表可存储在分布式存储服务器中,当用户重新进行数据访问时,可从分布式存储服务器中重新恢复。

[0138] 为了防止视频排重服务器73宕机导致内存中的内容排重特征表丢失,这里可以设置设定时间间隔(如每隔10分钟)将内存中的内容在分布式存储服务器74中进行存储操作;同时为了防止分布式存储服务器74中的占用空间过大,可把超过30*24小时没有访问的内容排重特征表进行清理操作。

[0139] 这样即完成了本具体实施例的内容推送方法及内容推送装置的推送内容过滤以及推送过程。

[0140] 本发明的内容推送方法及装置通过内容排重特征表记录已推送内容的特征值,从而可有效快捷的对已推送内容进行识别,有效的对推送内容进行有效去重;解决了现有的内容推送方法及内容推送装置的不能对推送内容进行有效去重的技术问题。

[0141] 如本申请所使用的术语“组件”、“模块”、“系统”、“接口”、“进程”等等一般地旨在指计算机相关实体:硬件、硬件和软件的组合、软件或执行中的软件。例如,组件可以是但不限于是运行在处理器上的进程、处理器、对象、可执行应用、执行的线程、程序和/或计算机。通过图示,运行在控制器上的应用和该控制器二者都可以是组件。一个或多个组件可以有在于执行的进程和/或线程内,并且组件可以位于一个计算机上和/或分布在两个或更多计算机之间。

[0142] 图8和随后的讨论提供了对实现本发明所述的内容推送装置所在的电子设备的工作环境的简短、概括的描述。图8的工作环境仅仅是适当的工作环境的一个实例并且不旨在建议关于工作环境的用途或功能的范围的任何限制。实例电子设备812包括但不限于可穿戴设备、头戴设备、医疗健康平台、个人计算机、服务器计算机、手持式或膝上型设备、移动设备(比如移动电话、个人数字助理(PDA)、媒体播放器等等)、多处理器系统、消费型电子设备、小型计算机、大型计算机、包括上述任意系统或设备的分布式计算环境,等等。

[0143] 尽管没有要求,但是在“计算机可读指令”被一个或多个电子设备执行的通用背景下描述实施例。计算机可读指令可以经由计算机可读介质来分布(下文讨论)。计算机可读指令可以实现为程序模块,比如执行特定任务或实现特定抽象数据类型的功能、对象、应用编程接口(API)、数据结构等等。典型地,该计算机可读指令的功能可以在各种环境中随意组合或分布。

[0144] 图8图示了包括本发明的内容推送装置中的一个或多个实施例的电子设备812的实例。在一种配置中,电子设备812包括至少一个处理单元816和存储器818。根据电子设备的确切配置和类型,存储器818可以是易失性的(比如RAM)、非易失性的(比如ROM、闪存等)或二者的某种组合。该配置在图8中由虚线814图示。

[0145] 在其他实施例中,电子设备812可以包括附加特征和/或功能。例如,设备812还可以包括附加的存储装置(例如可移除和/或不可移除的),其包括但不限于磁存储装置、光存储装置等等。这种附加存储装置在图8中由存储装置820图示。在一个实施例中,用于实现本文所提供的一个或多个实施例的计算机可读指令可以在存储装置820中。存储装置820还可以存储用于实现操作系统、应用程序等的其他计算机可读指令。计算机可读指令可以载入存储器818中由例如处理单元816执行。

[0146] 本文所使用的术语“计算机可读介质”包括计算机存储介质。计算机存储介质包括以用于存储诸如计算机可读指令或其他数据之类的信息的任何方法或技术实现的易失性和非易失性、可移除和不可移除介质。存储器818和存储装置820是计算机存储介质的实例。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字通用盘(DVD)或其他光存储装置、盒式磁带、磁带、磁盘存储装置或其他磁存储设备、或可以用于存储期望信息并可以被电子设备812访问的任何其他介质。任意这样的计算机存储介质可以是电子设备812的一部分。

[0147] 电子设备812还可以包括允许电子设备812与其他设备通信的通信连接826。通信连接826可以包括但不限于调制解调器、网络接口卡(NIC)、集成网络接口、射频发射器/接

收器、红外端口、USB连接或用于将电子设备812连接到其他电子设备的其他接口。通信连接826可以包括有线连接或无线连接。通信连接826可以发射和/或接收通信媒体。

[0148] 术语“计算机可读介质”可以包括通信介质。通信介质典型地包含计算机可读指令或诸如载波或其他传输机构之类的“已调制数据信号”中的其他数据，并且包括任何信息递送介质。术语“已调制数据信号”可以包括这样的信号：该信号特性中的一个或多个按照将信息编码到信号中的方式来设置或改变。

[0149] 电子设备812可以包括输入设备824，比如键盘、鼠标、笔、语音输入设备、触摸输入设备、红外相机、视频输入设备和/或任何其他输入设备。设备812中也可以包括输出设备822，比如一个或多个显示器、扬声器、打印机和/或任意其他输出设备。输入设备824和输出设备822可以经由有线连接、无线连接或其任意组合连接到电子设备812。在一个实施例中，来自另一个电子设备的输入设备或输出设备可以被用作电子设备812的输入设备824或输出设备822。

[0150] 电子设备812的组件可以通过各种互连(比如总线)连接。这样的互连可以包括外围组件互连(PCI)(比如快速PCI)、通用串行总线(USB)、火线(IEEE1394)、光学总线结构等等。在另一个实施例中，电子设备812的组件可以通过网络互连。例如，存储器818可以由位于不同物理位置中的、通过网络互连的多个物理存储器单元构成。

[0151] 本领域技术人员将认识到，用于存储计算机可读指令的存储设备可以跨越网络分布。例如，可经由网络828访问的电子设备830可以存储用于实现本发明所提供的的一个或多个实施例的计算机可读指令。电子设备812可以访问电子设备830并且下载计算机可读指令的一部分或所有以供执行。可替代地，电子设备812可以按需要下载多条计算机可读指令，或者一些指令可以在电子设备812处执行并且一些指令可以在电子设备830处执行。

[0152] 本文提供了实施例的各种操作。在一个实施例中，所述的一个或多个操作可以构成一个或多个计算机可读介质上存储的计算机可读指令，其在被电子设备执行时将使得计算设备执行所述操作。描述一些或所有操作的顺序不应当被解释为暗示这些操作必需是顺序相关的。本领域技术人员将理解具有本说明书的益处的可替代的排序。而且，应当理解，不是所有操作必需在本文所提供的每个实施例中存在。

[0153] 而且，尽管已经相对于一个或多个实现方式示出并描述了本公开，但是本领域技术人员基于对本说明书和附图的阅读和理解将会想到等价变型和修改。本公开包括所有这样的修改和变型，并且仅由所附权利要求的范围限制。特别地关于由上述组件(例如元件、资源等)执行的各种功能，用于描述这样的组件的术语旨在对应于执行所述组件的指定功能(例如其在功能上是等价的)的任意组件(除非另外指示)，即使在结构上与执行本文所示的本公开的示范性实现方式中的功能的公开结构不等同。此外，尽管本公开的特定特征已经相对于若干实现方式中的仅一个被公开，但是这种特征可以与如可以对给定或特定应用而言是期望和有利的其他实现方式的一个或多个其他特征组合。而且，就术语“包括”、“具有”、“含有”或其变形被用在具体实施方式或权利要求中而言，这样的术语旨在以与术语“包含”相似的方式包括。

[0154] 本发明实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能

模块的形式实现并作为独立的产品销售或使用，也可以存储在一个计算机可读取存储介质中。上述提到的存储介质可以是只读存储器，磁盘或光盘等。上述的各装置或系统，可以执行相应方法实施例中的方法。

[0155] 综上所述，虽然本发明已以实施例揭露如上，实施例前的序号仅为描述方便而使用，对本发明各实施例的顺序不造成限制。并且，上述实施例并非用以限制本发明，本领域的普通技术人员，在不脱离本发明的精神和范围内，均可作各种更动与润饰，因此本发明的保护范围以权利要求界定的范围为准。

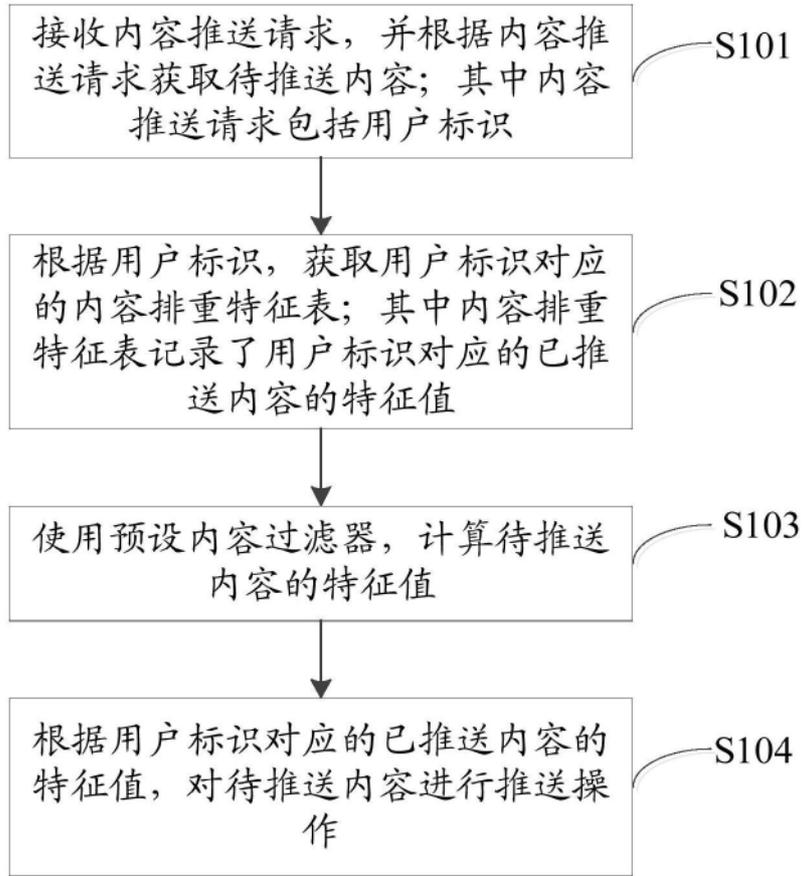


图1

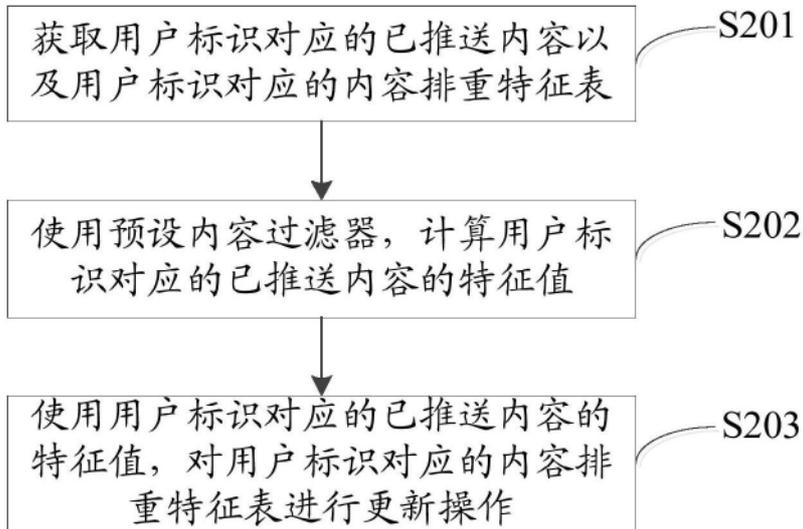


图2

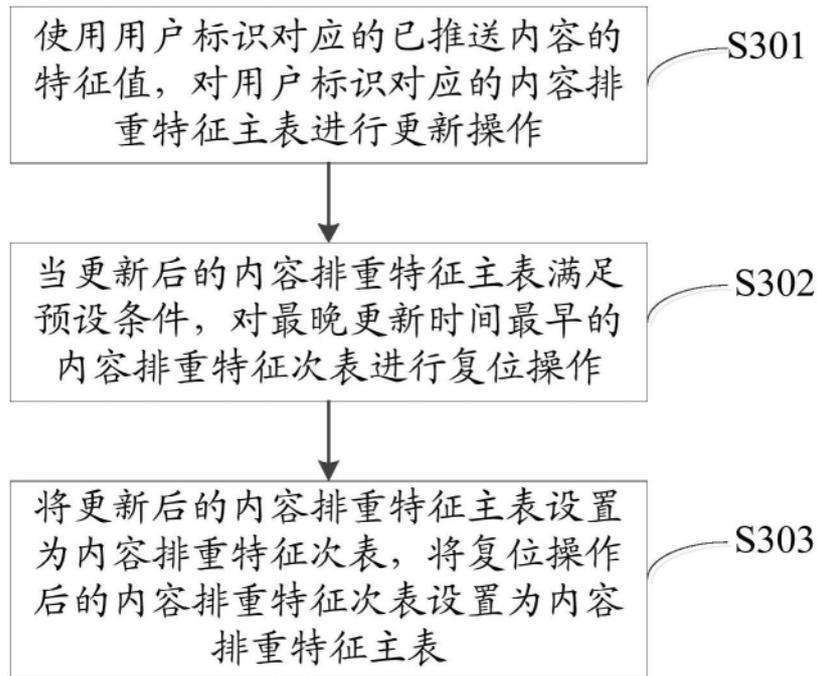


图3

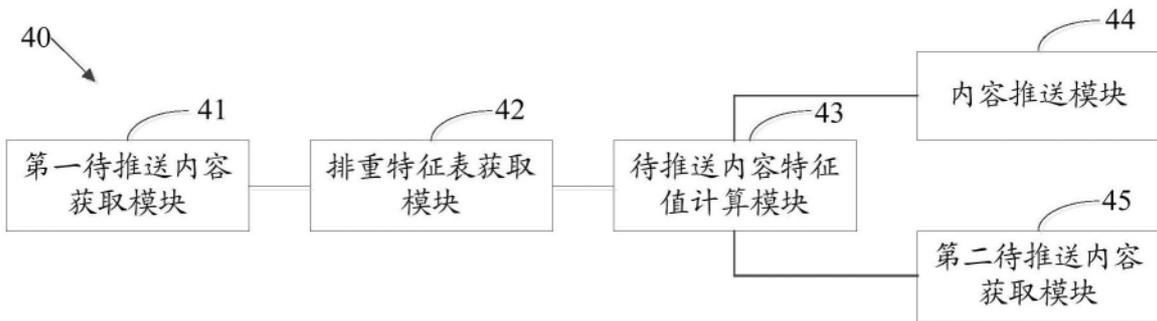


图4

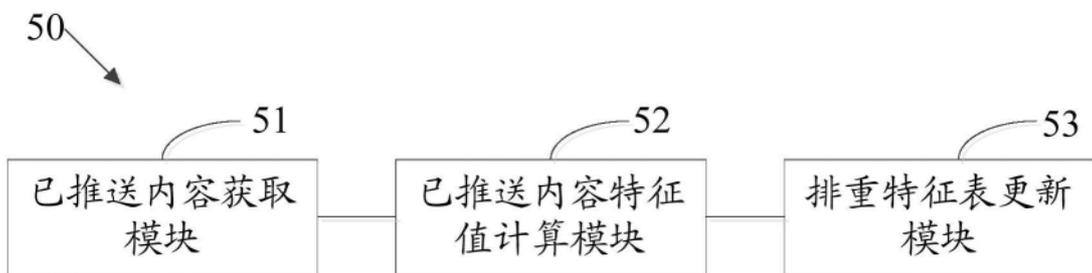


图5



图6

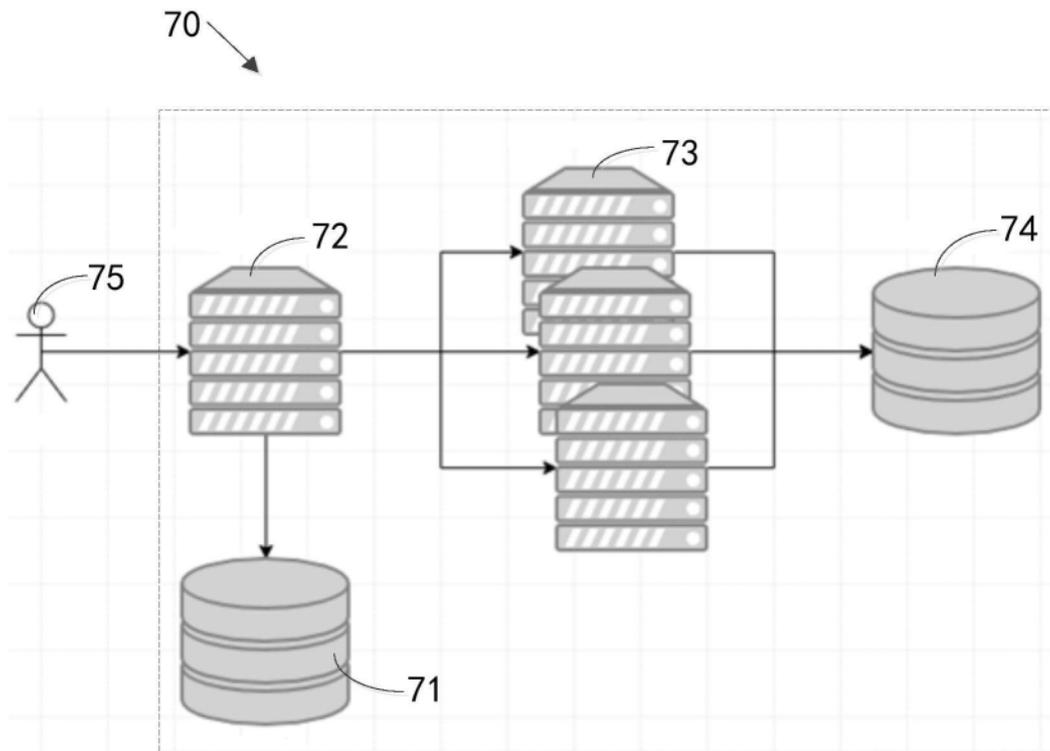


图7a

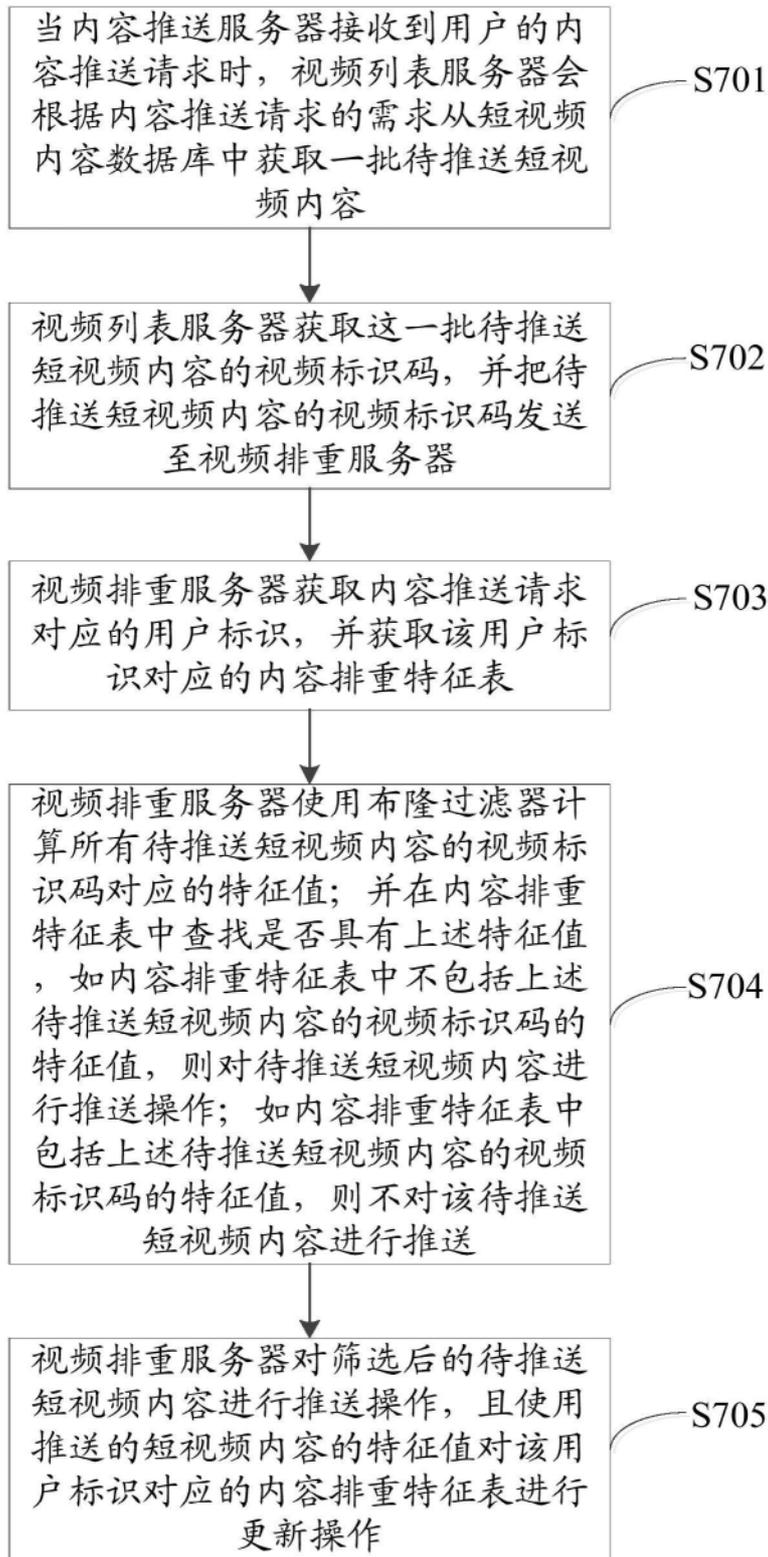


图7b

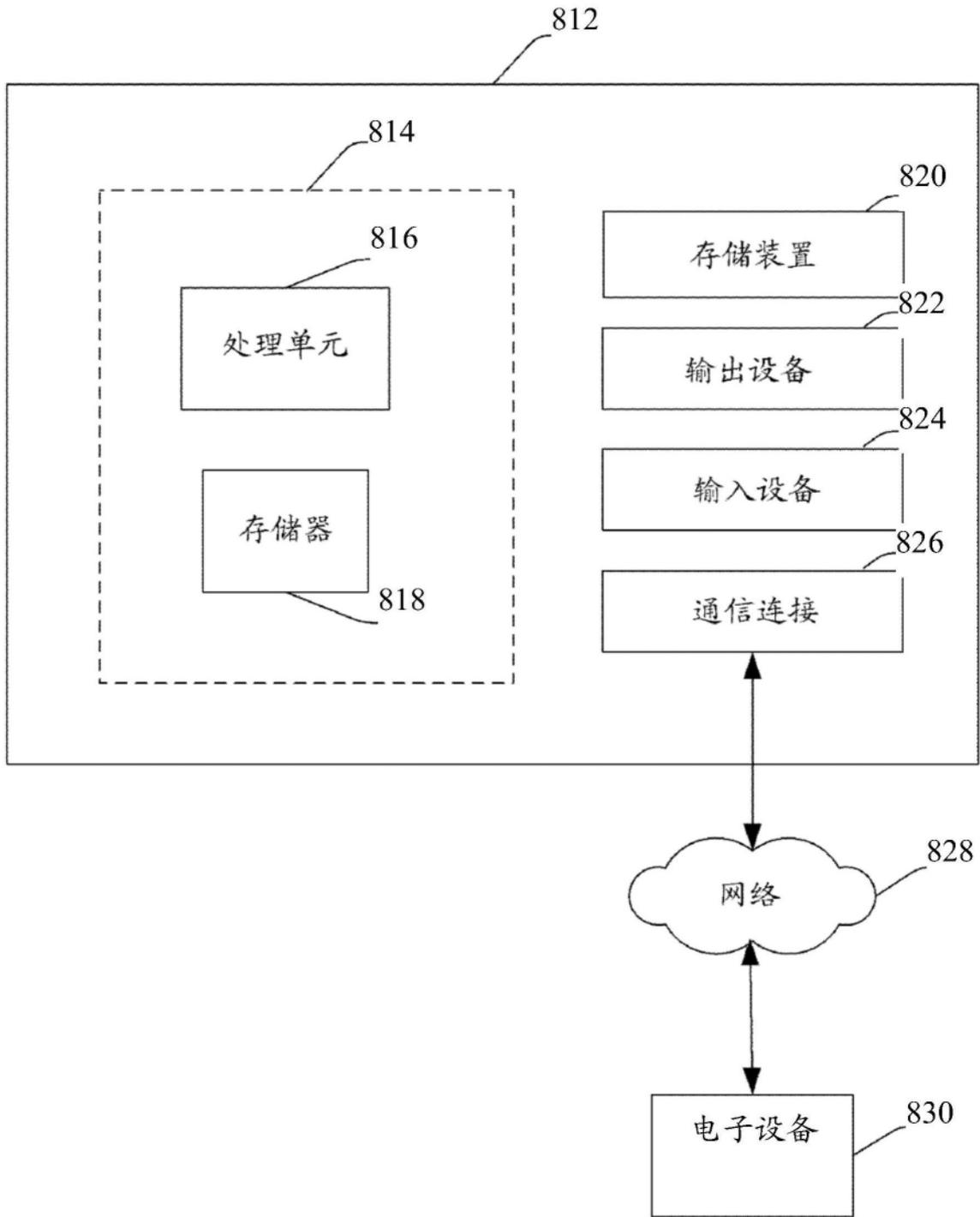


图8