说明书

发明名称: 一种针对媒体对象的推荐方法、装置、设备以及介质

[0001] 本申请要求于2024年12月05日提交中国专利局、申请号为2024117889825、申请名称为“一种针对媒体对象的数据处理方法、装置、设备以及介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

[0002] 本申请涉及互联网技术领域，尤其涉及一种针对媒体对象的推荐方法、装置、设备以及介质。

背景技术

[0003] 目前，冷启动推荐是指在推荐系统中对当前所获取到的冷启动媒体对象(例如，新引入的媒体对象，该新引入的媒体对象具体可以是当前所统计到的没有交互数据量和/或仅有较少交互数据量的媒体对象)进行推荐时的一种情况，由于该新引入的媒体对象(例如，由主播用户所新建的直播间)缺乏相应的历史交互行为数据，所以，推荐系统在通过推荐模型对媒体对象进行冷启动推荐的过程中，通常是基于该媒体对象的对象标签进行相似标签推荐。比如，对于媒体对象为A类风格的直播间1而言，在需要使用推荐模型进行直播推荐时，推荐系统会直接通过该直播间1的风格标签(即A类风格)，在海量直播间中查找携带该风格标签(即A类风格)的其他直播间(例如，直播间2)，进而可以向当前正在该直播间2中观看直播视频的业务对象(即观众用户)推送该直播间1，以实现对该直播间1的冷启动推荐。

[0004] 然而，发明人在实践中发现，在基于媒体对象的标签进行冷启动推荐时，由于直播间1中的直播视频(例如，直播视频1)缺乏实际的交互记录信息，因此，当通过相似标签匹配的方式查找到直播间2时，会直接基于该直播间2中的直播视频(例如，直播视频2)的交互记录信息中所记录的观众用户，无差别向这些观众对象推送该直播间1。然而，一旦该直播视频2的风格标签与该直播视频2中的实际直播内容存在差异，则会导致推送的直播间1中的直播内容，难以真正贴合这些观众用户的观看需求，以至于在一定程度上降低了对某些交互数据量少的媒体对象进行媒体推荐的准确度。

发明内容

[0005] 本申请实施例提供一种针对媒体对象的推荐方法、装置、设备以及介质，可以针对冷启动媒体对象(即交互数据量较少的媒体对象)，提升媒体推荐的准确度。

[0006] 本申请实施例一方面提供了一种针对媒体对象的推荐方法，包括：

[0007] 获取待推荐媒体对象在目标采样窗口中的目标音频数据，对目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量；

[0008] 获取用于表征N个历史媒体对象的N个历史文本表征向量，对目标文本表征向量和N个历史文本表征向量分别进行向量相似计算，得到N个向量相似度；每个历史媒体对象的交互数据量均高于待推荐媒体对象的交互数据量；N为大于1的正整数；N个历史文本表征向量中的每个历史文本表征向量是通过对每个历史音频数据对应的历史语音识别文本进行文本向量表征处理后所记录得到的，每个历史音频数据是指每个历史媒体对象在记录到的历史采样窗口中的音频数据；历史采样窗口的窗口时长与目标采样窗口的窗口时长保持一致；

[0009] 在通过N个向量相似度从N个历史媒体对象中筛选出K个历史媒体对象时，建立K个历史媒体对象与待推荐媒体对象之间的相似关联关系；K为小于N的正整数；

[0010] 在获取到K个历史媒体对象的K个历史推荐特征时，基于相似关联关系和K个历史推荐特征，确定用于表征待推荐媒体对象的目标推荐特征，通过目标推荐特征对待推荐媒体对象进行媒体推荐。

[0011] 本申请实施例一方面提供了一种针对媒体对象的推荐装置，包括：

[0012] 音频数据语音识别模块，用于获取待推荐媒体对象在目标采样窗口中的目标音频数据，对目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量；

[0013] 向量相似度计算模块，用于获取用于表征N个历史媒体对象的N个历史文本表征向量，对目标文本表征向量和N个历史文本表征向量分别进行向量相似计算，得到N个向量相似度；每个历史媒体对象的交互数据量均高于待推荐媒体对象的交互数据量；N为大于1的正整数；N个历史文本表征向量中的每个历史文本表征向量是通过对每个历史音频数据对应的历史语音识别文本进行文本向量表征处理后所记录得到的，每个历史音频数据是指每个历史媒体对象在记录到的历史采样窗口中的音频数据；历史采样窗口的窗口时长与目标采样窗口的窗口时长保持一致；

[0014] 相似关联关系建立模块，用于在通过所述N个向量相似度从所述N个历史媒体对象中筛选出K个历史媒体对象时，建立所述K个历史媒体对象与所述待推荐媒体对象之间的相似关联关系；K为小于N的正整数；

[0015] 目标推荐特征确定模块，用于在获取到K个历史媒体对象的K个历史推荐特征时，基于相似关联关系和K个历史推荐特征，确定用于表征待推荐媒体对象的目标推荐特征，通过目标推荐特征对待推荐媒体对象进行媒体推荐。

[0016] 本申请实施例一方面提供了一种计算机设备，包括存储器和处理器，存储器与处理器相连，存储器用于存储计算机程序，处理器用于调用计算机程序，以使得该计算机设备执行本申请实施例中上述一方面提供的方法。

[0017] 本申请实施例一方面提供了一种计算机可读存储介质，计算机可读存储介质中存储有计算机程序，计算机程序适于由处理器加载并执行，以使得具有处理器的计算机设备执行本申请实施例中上述一方面提供的方法。

[0018] 本申请实施例一个方面提供了一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，使得该计算机设备执行上述一方面提供的方法。

[0019] 本申请实施例所涉及的计算机设备在获取到待推荐媒体对象(例如，当前直播间)在目标采样窗口中的目标音频数据时，可以按照语音识别的方式，对该目标音频数据进行语音识别，以得到该目标音频数据对应的目标语音识别文本，应当理解，这里的待推荐媒体对象(例如，当前直播间)为相对于N个历史媒体对象(例如，具有较高交互数据量的N个历史直播间)的交互数据量而言，具备较少交互数据量的媒体对象，即这N个历史媒体对象中的每个历史媒体对象的交互数据量均高于该待推荐媒体对象的交互数据量。因此，本申请实施例在通过文本向量化表征处理的方式，对该目标语音识别文本进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量的情况下，可以进一步获取用于表征这N个历史媒体对象的N个历史文本表征向量，进而对目标文本表征向量与N个历史文本表征向量中的每个历史文本表征向量分别进行向量相似计算，以计算得到N个向量相似度。其中，应当理解，对于该N个历史媒体对象的任意一个历史媒体对象的历史文本表征向量而言，均是通过对相应历史媒体对象的历史语言识别文本进行文本向量表征处理后所记录得到的，即一个历史文本表征向量是通过对一个历史音频数据对应的历史语音识别文本进行文本向量表征处理后所记录得到的，且一个历史音频数据是指一个历史媒体对象在记录到的历史采样窗口中的音频数据；这里的历史采样窗口的窗口时长与所述目标采样窗口的窗口时长保持一致。进一步地，本申请实施例所涉及的计算机设备可以按照向量相似检索的方式，通过上述N个向量相似度，从上述N个历史媒体对象中筛选出K个历史媒体对象(即筛选出的这K个历史媒体对象均为与待推荐媒体对象具有较高向量相似度的媒体对象，例如，具备较高交互数据量的K个历史直播间)，这样，本申请实施例就可以通过筛选出的具有较高向量相似度的这K个历史媒体对象来表征待推荐媒体对象，从而能够直观(即显

式)建立该待推荐媒体对象与这K个历史媒体对象之间的相似关联关系。基于此，本申请实施例可以在获取到这K个历史媒体对象的K个历史推荐特征时，进一步基于该相似关联关系准确确定出该待推荐媒体对象(例如，当前直播间)的目标推荐特征，这样，在通过目标推荐特征对待推荐媒体对象进行媒体推荐时，不仅可以提升媒体推荐的准确度，还可以从根源上解决现有技术中因具有较少交互数据量的该待推荐媒体对象(即冷启动媒体对象，例如，冷启动直播间)的缺乏实际的交互记录信息，而影响媒体推荐的准确度的技术问题，换言之，本申请实施例还可以在待推荐媒体对象为冷启动待推荐媒体时，通过向量相似检索的方式所确定出的具有较高向量相似度的K个历史媒体对象(例如，具备较高交互数据量的K个历史直播间)来显示表征该冷启动媒体对象，从而可以实现对该冷启动媒体对象进行加热推荐的效果。

附图说明

[0020] 图1是本申请实施例提供的一种网络架构的结构示意图；

[0021] 图2是本申请实施例提供的一种数据交互场景示意图；

[0022] 图3是本申请实施例提供的一种针对媒体对象的推荐方法的流程示意图；

[0023] 图4是本申请实施例提供的一种历史媒体对象的交互数据曲线的示意图；

[0024] 图5是本申请实施例提供的一种相似关联关系示意图；

[0025] 图6是本申请实施例提供的一种在业务处理平台上对直播类媒体对象进行媒体推荐的场景示意图；

[0026] 图7是本申请实施例提供的一种在直播间进行直播的场景示意图；

[0027] 图8是本申请实施例提供的一种在业务处理平台上对点播类媒体对象进行媒体推荐的场景示意图；

[0028] 图9是本申请实施例提供的一种基于点播标识进行视频播放的场景示意图；

[0029] 图10是本申请实施例提供的一种业务处理平台的推荐显示列表的示意图；

[0030] 图11是本申请实施例提供的一种目标音频无效处理方法流程图；

[0031] 图12是本申请实施例提供的另一种目标音频无效处理方法流程图；

[0032] 图13是本申请实施例提供的一种针对媒体对象的目标业务对象预估流程示意图；

[0033] 图14是本申请实施例提供的一种针对直播间的目标业务对象预估流程示意图；

[0034] 图15是本申请实施例提供的另一种针对媒体对象的推荐方法的流程示意图；

[0035] 图16是本申请实施例提供的一种基于待处理直播数据播放时间轴的动态采样示意图；

[0036] 图17是本申请实施例提供的一种针对媒体对象的推荐装置的结构示意图；

[0037] 图18是本申请实施例提供的一种计算机设备的结构示意图。

具体实施方式

[0038] 下面将结合本申请实施例的相关技术用语的概念进行说明:

[0039] 一、媒体对象，本申请实施例所涉及的媒体对象包括待处理媒体对象(即待推荐媒体对象)和已推荐的历史媒体对象，这里的待处理媒体对象(即待推荐媒体对象)和历史媒体对象为同一类型的媒体对象。这里的媒体对象是指在应用客户端中所展示的具备实时视频处理的第一媒体对象(例如，虚拟的直播间、会议室以及教室等)或具备视频进度灵活调整的第二媒体对象(例如，点播视频)，该第一媒体对象内的视频可以包括但不限于：直播视频、会议视频和教育视频。应当理解，对于媒体对象为待推荐媒体对象而言，这里的第一媒体对象具体可以包括在实时直播场景下的待推荐直播类媒体对象，同理，第二媒体对象具体可以包括对该实时直播场景进行直播录制所得到当前能够进行视频回放的待推荐点播类媒体对象。

[0040] 其中，直播视频具体可以包括在实时直播场景下呈现在直播客户端中的某个直播间(即上述第一媒体对象)内的实时直播视频和/或在直播回放场景下呈现在该直播客户端中的直播间(即上述第二媒体对象)内的录播视频，点播视频具体可以包括在视频点播场景下呈现在视频播放客户端中的长视频以及短视频。其中，长视频是指视频时长大于视频时长阈值的视频，短视频是指视频时长小于该视频时长阈值的视频。

[0041] 其中，可以理解的是，本申请实施例可以将在推荐系统中需要进行加热推荐的具有较小交互数据量的媒体对象统称为待处理媒体对象或者待推荐媒体对象(例如，待推荐媒体对象可以是指推荐系统所统计都的一些与用户存在较小交互数据量的媒体对象)，并将在推荐系统中已进行媒体推荐的具有较大交互数据量的媒体对象统称为历史媒体对象(例如，历史媒体对象可以是指推荐系统所统计到的一些与用户存在较大交互数据量的媒体对象)。具体地，比如，在实时直播场景下，可以将当前需要进行加热推荐的某个冷启动直播间或者满足长尾分布的尾部直播间(即长尾直播间)称之为待推荐媒体对象，并可以将一些已经直播完成或者已完成直播推荐的热门直播间或者满足长尾分布的头部直播间称之为历史媒体对象。换言之，在实时直播场景下，本申请实施例所涉及的待推荐媒体对象具体可以包括冷启动直播间(例如，新引入的直播间)和满足长尾分布的尾部直播间(即长尾直播间，例如，一些非新引入的、但具有较小交互数据的直播间)，且本申请实施例所涉及的历史媒体对象具体可以包括一些具有较大交互数据量的热门直播间和满足长尾分布的头部直播间。应当理解，这里的交互数据量(即点击浏览量)是指推荐系统所累计统计到的一个或者多个用户对相应媒体对象(或者相应媒体对象的多媒体数据)执行点击和/或浏览等交互操作的数据量，在本申请实施例中，交互数据量的大小是指相对于推理系统所配置的预设的交互数据量阈值而确定的。比如，在一种可实现方式中，推荐系统所涉及的计算机设备会将当前统计到的交互数据量小于交互数据量阈值的直播间视为具有较小交互数据量的直播间(即低交互直播间)，而将当前统计到的交互数据量大于或者等于交互数据量阈值的直播间视为具有较大交互数据量的直播间(即高交互直播间)。

[0042] 在本申请实施例中，交互数据量越大，则说明该交互数据量所对应的媒体对象的热度越高，反之，交互数据量越小，则说明该交互数据量所对应的媒体对象的热度越低，例如，对于冷启动直播间或者长尾直播间等媒体对象而言，推荐系统所累积统计到的这些媒体对象的交互数据量通常是小于交互数据量阈值的，由于这些媒体对象具有较小的交互数据量，故而缺乏用于建模物品相对于用户的偏好的历史行为交互数据信息的，因此，若直接使用这些媒体对象自身所具备的物品的属性特征，则会导致在推荐系统中延迟对这些低热度媒体对象中的物品的分发。因此，本申请实施例为提升对这些低热度媒体对象中的物品的分发的时效性和准确性，提出可以在检索到与该具有较小交互数据量的媒体对象(即待推荐媒体对象)相匹配的多个具有较高语义相似、且具有较大交互数据量的其他已推荐的历史媒体对象时，能够通过这些具有较高语义相似、且具有较大交互数据量的其他已推荐的历史媒体对象的推荐特征(即历史推荐特征)来等效表示该具有较小交互数据量的媒体对象(即待推荐媒体对象)的推荐特征(即目标推荐特征)，以实现对该具有较小交互数据量的媒体对象(即待推荐媒体对象)的加热推荐。

[0043] 二、采样窗口，本申请实施例所涉及的采样窗口包括目标采样窗口和历史采样窗口。这里的目标采样窗口和历史采样窗口均是指在视频采集过程中，用于定义视频信号采样参数的窗口。该视频信号采样参数具体包括采样频率、分辨率、帧率等。

[0044] 其中，可以理解的是，本申请实施例可以在某个业务场景(例如，实时直播场景)下，将按照采样频率实时从当前采样时间点(例如，当前用于反映直播间的最新直播进度的时间点)开始，往回截取到的一定时长(即预设的窗口时长)的待推荐媒体对象的多媒体数据(例如，待推荐直播数据)的采样窗口统称为目标采样窗口，并将按照同一采样频率所历史截取到的一定时长的历史媒体对象的多媒体数据的采样窗口统称为历史采样窗口。具体地，在实时直播场景下，可以将随着某个直播间的直播进度的进行而用于动态截取到一定时长的直播数据的采样窗口称之为目标采样窗口，即本申请实施例在按照动态采样方式通过不同直播进度而动态确定出不同采样结束时间点时，可以得到不同采样结束时间点所对应的采样窗口，从而可以将从这些不同采样结束时间点所对应的采样窗口中所确定的最新的采样窗口，动态确定为目标采样窗口，进而可以通过该目标采样窗口动态采样得到该待推荐媒体对象内的不同音频数据，以便于后续可以通过噪音去除的方式从当前采样得到的音频数据去除冗余的噪音数据，以得到与说话人相关的目标音频数据，并可

以在直播完成时，将当前已经完成直播且已经完成采样的所有采样窗口中的最后一个采样窗口称之为历史采样窗口。

[0045] 应当理解，这里的目标采样窗口是指针对当前需要加热推荐的同一媒体对象(即同一待推荐媒体对象)而言，从一系列的采样窗口中所选取的最新的采样窗口，通过该最新的采样窗口可以从自身具有较大媒体信息量的待推荐媒体对象(例如，直播间或者点播视频等)中，采集得到一段时长为预设的窗口时长的视频片段和/或音频片段，进而可以从采集到的视频片段或者音频片段中提取得到位于该目标采样窗口中的目标音频数据(例如，与说话人的语音相关的音频数据)。

[0046] 三、音频数据，本申请实施例所涉及的音频数据包括目标音频数据和历史音频数据。该目标音频数据和历史音频数据均是指通过模数转换，将相应的模拟声音信号编码转换为数字化形式所得到的音频数据(比如，可以是以二进制表示的音频编码数据)。其中，可以理解的是，本申请实施例在通过目标采样窗口采集到待推荐媒体对象的多媒体数据(例如，采集得到视频片段或者音频片段等待推荐直播数据)时，可以将从该待推荐媒体对象的多媒体数据中所提取到的与说话人的语音相关的音频数据称之为目标音频数据，此时，目标音频数据可以是通过对位于该目标采样窗口中的视频片段或者音频片段中的音频数据进行去噪处理所得到的。同理，本申请实施例在通过历史采样窗口采集到历史媒体对象的多媒体数据(例如，历史直播数据)时，可以将从该历史媒体对象的多媒体数据中所提取到的音频数据称之为历史音频数据。

[0047] 可选地，在一种可实现方式中，比如，在实时直播场景下，还可以将通过目标采样窗口从正在进行的当前直播间中所截取的视频片段或者音频片段中的音频数据称之为目标音频数据，并可以将通过历史采样窗口从当前已经完成直播的历史直播间所截取的视频片段或者音频片段中的音频数据称之为历史音频数据。

[0048] 四、相似检索，本申请实施例所涉及的相似检索是指基于不同检索对象的向量相似度进行向量相似检索的方式，在本申请实施例中，对于给定的某个检索对象(即上述待推荐媒体对象)而言，推荐系统可以按照该向量相似检索的方式在预设的检索对象集合(即由N个历史媒体对象所构成的集合)中查找(即检索)出与该检索对象(即上述待推荐媒体对象)相似或者相匹配的其他检索对象(即上述K个历史媒体对象)。

[0049] 其中，可以理解的是，本申请实施例在按照该向量相似检索的方式进行相似检索的过程中，需要先将目标文本表征向量与检索对象集合中的各个历史媒体对象的历史文本表征向量进行相似度计算，以得到该目标文本表征向量与每个历史文本表征向量之间的向量相似度，进而可以基于得到的这些向量相似度，准确从该检索对象集合中检索出与待推荐媒体对象具有较高向量相似度的历史媒体对象。具体地，比如，在实时直播场景下，可以将当前直播间(即待推荐媒体对象)的目标文本表征向量与各个历史直播间(即各个历史媒体对象)的历史文本表征向量进行向量相似度计算，以得到当前直播间(即待推荐媒体对象)与各个历史直播间(即各个历史媒体对象)之间的向量相似度，此时，推荐系统可以进一步按照计算得到这些向量相似度由高到低的排序方式，对这些历史媒体对象(例如，N个历史直播间)进行排序处理，进而可以将排序处理后的这些历史媒体对象所构成的排序列表统称为历史媒体对象列表，从而可以在该历史媒体对象列表中筛选排序靠前的前K(即K＝20)个历史媒体对象(例如，K个历史直播间)作为与待推荐媒体对象相关联的历史媒体对象(例如，与该当前直播间具有较高向量相似度的历史直播间)。

[0050] 五、推荐特征，本申请实施例所涉及的推荐特征包括目标推荐特征和历史推荐特征。这里的推荐特征是指在推荐系统中用于描述用户和/或物品的属性数据的特征数据，这些特征数据可以帮助推荐系统理解业务对象(例如，用户U1)对某些物品(例如，在某些直播间中的与讲解某个主题下的某些物品相关的直播视频)的偏好，从而在向该业务对象(例如，用户U1)推荐相关的物品时，可以将该物品所在的该直播间中的该直播片段推荐给用户U1，以增强推荐的准确度，从而可以帮助用户U1找到贴合自己偏好的物品。

[0051] 其中，可以理解的是，本申请实施例将可以描述待推荐媒体对象所对应的属性数据的特征数据称为目标推荐特征，并将可以描述历史媒体对象所对应的属性数据的特征数据称为历史推荐特征。具体地，比如，在实时直播场景下，本申请实施例可以通过推荐模型从历史媒体对象的历史行为交互信息中提取到该

历史媒体数据所对应的属性数据，并进一步对该属性数据进行向量化的特征计算，进而可以将特征计算所得到的特征数据统称为针对历史媒体对象的历史推荐特征，这样，推荐系统在通过推荐模型计算得到K个历史媒体对象中的每个历史媒体对象的历史推荐特征的情况下，可以进一步通过在这K个历史媒体对象与该待推荐媒体对象之间建立好的相似关联关系，直接采用池化处理的方式对这K个历史推荐特征进行池化处理，以通过池化处理所得到的池化特征来等效得到用于描述该待推荐媒体对象所对应的属性数据的特征数据(即目标推荐特征)。

[0052] 在本申请实施例中，池化处理至少可以包括平均池化处理、最大池化处理以及随机池化处理。这里的池化处理为卷积神经网络中的一个特殊操作，主要用于在一定的区域内提取出该区域的关键信息。比如，本申请实施例可以通过池化处理的方式在K个历史推荐特征所构成的区域中，提取出能够充分表征该区域的关键信息作为池化特征(即池化处理特征)，进而可以利用充分提取到的池化特征(即池化处理特征)来等效得到用于表征该待推荐媒体对象的目标推荐特征，以便于后续在推荐系统中按照协同过滤的方式，利用该目标推荐特征实现对该待处理媒体对象(即待推荐媒体对象)的媒体推荐(例如，冷启动推荐)。这里的卷积神经网络可以为推荐模型中的一个网络单元，也可以为独立于推荐模型的一个网络单元，本申请实施例将不对用于进行池化处理的卷积神经网络的具体部署方式进行限定，此外，本申请实施例也不对用于进行池化处理的具体实现方式进行限定。

[0053] 其中，可以理解的是，平均池化处理是指计算机设备可以对由K个历史推荐特征所构成的区域进行均值化操作的处理方式，这样，计算机设备可以进一步将均值化操作后所得到的均值(即池化处理特征)作为目标推荐特征。同理，最大池化处理是指计算机设备可以在由K个历史推荐特征所构成的区域中保留最大历史推荐特征(即特征值最大的历史推荐特征)的另一处理方式，这样，计算机设备可以进一步将在该区域中所保留的最大历史推荐特征(即另一种池化处理特征)作为目标推荐特征。以此类推，随机池化处理是指计算机设备可以在对由K个历史推荐特征所构成的区域中，按照为各个历史推荐特征分别配置的随机选择概率，对这K个历史推荐特征进行随机选取的又一处理方式。比如，在由K个历史推荐特征所构成的区域中，若历史推荐特征的特征值越大，则配置的随机选择概率则越大，反之，历史推荐特征的特征值越小，则配置的随机选择概率则越小。这样，本申请实施例在通过加权计算方式，通过为各个历史推荐特征所配置的随机选择概率分别对相应历史推荐特征进行加权计算之后，可以进一步对加权计算有相应随机选择概率的各个历史推荐特征进行随机选择，这样，计算机设备可以进一步将在该区域中所随机选择的某个加权计算有相应随机选择概率的历史推荐特征(即又一种池化处理特征)作为目标推荐特征。

[0054] 六、业务处理平台，本申请实施例所涉及的业务处理平台为部署了目标推荐模型的软件平台。业务处理平台是指一种支持企业或组织进行业务运营和管理活动的数字化平台。

[0055] 其中，可以理解的是，本申请实施例可以将推荐系统所提供的用于对需要进行加热推荐的媒体对象(即上述待推荐媒体对象)进行媒体推荐以及媒体展示的软件平台统称为业务处理平台。具体地，比如，在实时直播场景下，推荐系统在需要对某个直播间中的直播数据进行加热推荐时，可以基于用户–物品之间的匹配本质，将当前需要加热推荐的该直播间添加并展示在业务处理平台上，这样，推荐系统可以基于用户–物品之间的匹配本质，确定出与该待推荐媒体对象(即直播间)相匹配的用户，进而可以将与该待推荐对象相匹配的用户统称为目标业务对象，基于此，当该目标业务对象访问该业务处理平台时，推荐系统就可以为该目标业务对象推送包含该待推荐媒体对象的推荐显示列表，进而可以在该推荐显示列表中为该目标业务对象展示(例如，播放)该直播间中的直播数据，其中，本申请实施例所涉及的业务处理平台包括但不限于社交软件平台。

[0056] 七、业务对象，本申请实施例所涉及的业务对象包括访问该业务处理平台的所有业务对象，这些业务对象中包括由推荐系统按照用户与物品(即用户–物品)之间的匹配本质，从所有业务对象中所确定出的与该待推荐媒体对象相匹配的目标用户(即目标业务对象)。应当理解，业务对象是现实世界实体的软件抽象，可以代表具体业务领域中的人，比如，本申请实施例所涉及的业务对象在社交软件平台这一社交领域

下，可以表示访问业务处理平台的任一用户，还可以代表推荐模型经分析处理所确定出的用于接收推荐系统所推荐的某个媒体对象的目标用户(即目标业务对象)。

[0057] 其中，可以理解的是，本申请实施例将能够访问业务处理平台的所有用户称为业务对象，并将与待推荐媒体对象相匹配的在线用户称为目标业务对象。具体地，比如，在实时直播场景下，本申请实施例将前述加热推荐的直播间的直播数据所在业务处理平台上的所有用户称之为业务对象，并将与该直播间(具体可以是该直播间的直播数据)相匹配的在线用户称之为目标业务对象，这样，推荐系统可以调用目标推荐模型向目标业务对象推荐该直播间，以实现对该直播间中的直播数据的加热推荐。

[0058] 八、加热推荐，本申请实施例所涉及的加热推荐是指在检索到与具有较小交互数据量的媒体对象(即待推荐媒体对象)具有较高语义相似、且具有较大交互数据量的其他已推荐的历史媒体对象时，能够通过具有较高语义相似、且具有较大交互数据量的其他已推荐的历史媒体对象的推荐特征(即历史推荐特征)来等效表示该具有较小交互数据量的媒体对象(即待推荐媒体对象)的推荐特征(即目标推荐特征)的推荐手段。换言之，加热推荐是指通过本申请所涉及的推荐方法中的特定的算法和技术手段，能够将本身缺乏交互数据量和/或本身仅有较少交互数据量的媒体对象(例如，冷启动直播间或者长尾直播间)的内容，准确推荐给目标用户，以提高内容的曝光率和互动率。

[0059] 其中，冷启动直播间或者长尾直播间的内容具体可以包括在实时视频化和/或音频化的推荐场景中，通过目标采样窗口所动态采集到的主播在直播间中针对相应直播内容(例如，与可供用户触发交易的一个或者多个物品相关的直播主题)进行讲解的视频片段或者音频片段，这意味着对于同一媒体对象中的不同采样窗口下的这些直播主题而言，采用本申请实施例所提供的加热推荐方法，能够适应不同直播主题动态匹配到不同的目标用户，进而可以将不同采样窗口下的不同直播主题尽可能地推送给不同的目标用户，以提升与不同直播主题相关的视频片段或者音频片段的曝光率和互动率。

[0060] 其中，可以理解的是，本申请实施例可以在加热推荐的过程中，将与用户存在较小交互数据量的待推荐媒体对象推荐给目标用户，以提高待推荐媒体对象的曝光率和互动率。具体地，比如，在实时直播场景下，目标推荐模型通过等效计算得到的该待推荐媒体对象(例如，上述直播间)的目标推荐特征，智能匹配到与该直播间(例如，具体还可以是该直播间中用于讲解某个物品的直播视频)相匹配的目标业务对象(例如，目标用户)，以向该目标用户推荐该直播间，从而可以在目标业务对象的推荐显示列表中显示该被加热推荐的直播间，进而可以实现对该直播间的加热推荐。

[0061] 请参见图1，图1是本申请实施例提供的一种网络架构的结构示意图。如图1所示，该网络架构可以包括服务器10d和用户终端集群，该用户终端集群可以包括一个或者多个用户终端，这里不对用户终端的数量进行限制。如图1所示，该用户终端集群可以具体包括用户终端10a、用户终端10b、用户终端10c、用户终端20a、用户终端20b以及用户终端20c等。其中，服务器10d可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。用户终端10a、用户终端10b、用户终端10c、用户终端20a、用户终端20b以及用户终端20c等均可以包括：智能手机、平板电脑、笔记本电脑、掌上电脑、移动互联网设备(mobile internet device，MID)、可穿戴设备(例如智能手表、智能手环等)、智能语音交互设备、智能家电(例如智能电视等)以及车载设备等具有视频/图像播放功能的电子设备。如图1所示，用户终端10a、用户终端10b、用户终端10c、用户终端20a、用户终端20b以及用户终端20c等可以分别与服务器10d进行网络连接，以便于每个用户终端可以通过该网络连接与服务器10d之间进行数据交互。

[0062] 例如，在实时直播场景下，本申请实施例可以将用于进行直播数据采集的用户终端统称为第一本地设备，此时，待推荐媒体对象(例如待推荐直播类媒体对象)的多媒体数据(例如待推荐直播数据)就是由该第一本地设备(例如，用户终端10a)所采集到的，然后，第一本地设备(例如，用户终端10a)可以将采集到的待推荐媒体对象的多媒体数据发送给云服务器，此时，云服务器(例如，服务器10d)可以采用本申请实施例所提供的针对媒体对象的推荐方法，对接收到的该待推荐媒体对象的多媒体数据进行数据处理，以便于

后续可以通过向量相似检索的方式从具有较大交互数据量的大量历史媒体对象中筛选出与该待推荐媒体对象(例如待推荐直播类媒体对象)相匹配的历史媒体对象，进而可以通过筛选出的这些历史媒体对象的历史推荐特征来显示表征该待推荐媒体对象的目标推荐特征，进而可以通过目标推荐特征最终查找出与该待推荐媒体对象的多媒体数据相匹配的目标业务对象进行媒体推荐，即该云服务器(例如，服务器10d)可以进一步将上述待推荐媒体对象(具体可以是该待推荐媒体对象的多媒体数据)推送给上述目标业务对象，为便于进行区分，本申请实施例还可以在上述用户终端集群中，将该目标业务对象所使用的用户终端统称为第二本地设备，这样，第二本地设备在接收到由云服务器(例如，服务器10d)加热推荐的该待推荐媒体对象(具体可以是该待推荐媒体对象的多媒体数据)时，使得目标业务对象可以在该第二本地设备所呈现的业务处理平台上查看到该待推荐媒体对象的多媒体数据。换言之，本申请实施例所涉及的业务处理平台(例如社交软件平台)可以配置运行在另一本地设备(即第二本地设备，例如，用户终端20a)上，这样，第二本地设备(例如，用户终端20a)在接收到上述云服务器所加热推荐的待推荐媒体对象(例如待推荐直播类媒体对象)时，可以在业务处理平台上为该目标业务对象显示和播放该待推荐媒体对象的多媒体数据(例如待推荐直播数据)。

[0063] 在本申请实施例中，待推荐媒体对象的多媒体数据想要精准地推送给目标业务对象以实现对该待推荐媒体对象的加热推荐，首先需通过本地设备(即第一本地设备，例如，用户终端10a)上传上述待推荐媒体对象的多媒体数据。为便于理解，这里以上述待推荐媒体对象为直播类媒体对象，该待推荐媒体对象的多媒体数据为待推荐直播数据，历史媒体对象为历史直播类媒体对象为例，考虑到待推荐直播数据的直播进度在不断的动态变化，本申请实施例为了确保直播内容挖掘的可靠性和准确性，提出可以通过预设窗口长度的采样窗口(例如，目标采样窗口)对上述待推荐直播数据进行动态采样，以截取到位于目标采样窗口中的待推荐直播数据，并从截取到的待推荐直播数据中提取用于进行语音识别的音频数据作为目标音频数据。比如，本申请实施例可以将从待推荐直播数据中所提取到的说话人的音频数据(例如，主播对象的声音数据或者语音数据)作为目标音频数据。应当理解，若本申请实施例直接从待推荐直播数据(即待推荐媒体对象的多媒体数据)提取得到目标音频数据，则可以进一步通过语音转换理解组件对该目标音频数据进行语音识别处理，以得到目标语音识别文本。

[0064] 可选地，考虑到在实时直播场景下，说话人的音频数据(例如，主播对象的声音数据或者语音数据)容易受背景噪音、说话人口语化、说话内容重复、以及说话人不唯一(例如，存在其他说话人的音频数据)等因素的干扰，可能难以准确提取出能够有效进行语音识别的目标音频数据，或者导致当前提取到的目标语音数据变成难以进行语音识别的无效音频数据，基于此，本申请实施例提出还可以在未直接从待推荐直播数据中提取到目标音频数据时，进一步基于该待推荐直播数据所包含的多模态数据(例如，语音数据和视频数据等不同模态数据)之间的关联信息，快速从该待推荐直播数据中提取出视频数据，进而可以从视频数据中提取视觉呈现在该视频数据中的文字信息，以将提取到的文本信息等效作为目标语音识别文本，进而可以通过文本向量化表征处理的方式，对该目标语音识别文本进行文本向量表征处理，以得到该目标语音识别文本的文本特征，进而可以将该目标语音识别文本的文本特征作为用于表征待推荐媒体对象的目标文本表征向量。具体的，本申请实施例可以调用语义理解模型，基于引导提示模板数据对上述目标语音识别文本进行语义信息提取，得到目标语义理解信息，并将目标语义理解信息进行文本向量表征处理，以得到用于表征待推荐媒体对象(例如待推荐直播类媒体对象，比如，具体可以是当前通过目标采样窗口所动态采样到的该待推荐直播媒体对象的待推荐直播数据)的目标文本表征向量。

[0065] 其中，本申请实施例所涉及的针对媒体对象的推荐方法，具体包括获取用于表征N个历史直播类媒体对象的N个历史文本表征向量，并将上述用于表征上述待推荐直播类媒体对象的目标文本表征向量分别与N个历史直播类媒体对象的N个历史文本表征向量进行向量相似度计算，得到与上述N个历史直播类媒体对象相关联的N个向量相似度。进而基于上述计算得到的N个向量相似度从上述N个历史直播类媒体对象中筛选出K个历史直播类媒体对象，建立上述待推荐直播类媒体对象与K个历史直播类媒体对象的相似关联关系。进一步地，获取K个历史直播类媒体对象的K个历史推荐特征，对上述K个历史推荐特征进行池

化处理，得到用于表征待推荐直播类媒体对象的目标推荐表征，进而通过上述目标推荐特征实现对该待推荐直播类媒体对象的加热推荐。

[0066] 需要进行说明的是，本申请实施例中，用于采集待推荐直播类媒体对象(即上述待推荐媒体对象)的多媒体数据的本地设备(即第一本地设备，例如，用户终端10a)，与上述目标业务对象所使用的本地设备(即第二本地设备，例如，用户终端20a)为不同用户终端。应当理解，本申请实施例中所涉及的用于语音识别处理、视频文本识别、文本信息提取、文本向量表征处理以及相似度计算等的服务器为同一服务器。

[0067] 在一种可实现方式中，本申请实施例以待推荐媒体对象为直播类媒体对象(例如，当前直播间)，多媒体数据为待推荐直播数据，历史媒体对象为历史直播类媒体对象(例如，历史直播间)为例，阐述第一本地设备、云服务器以及第二本地设备之间的数据交互过程。比如，当主播对象在使用本地设备(即上述第一本地设备，例如，用户终端10a)开始进行直播时，该第一本地设备(用户终端10)可以源源不断的采集得到该直播类媒体对象的待推荐直播数据，并可以按照直播推流的方式将不断采集到的待推荐直播数据发送给云服务器(例如，上述图1所对应实施例中的服务器10d)。这样，云服务器(例如，服务器10d)可以基于预设的采集窗口(即目标采样窗口)对当前所接收到的待推荐直播数据进行动态采集，以便于后续可以从位于该目标采样窗口内的待推荐直播数据中提取得到目标音频数据。进一步地，云服务器可以对上述目标音频数据进行语音识别处理，以得到语音识别处理结果。若语音识别处理结果指示识别到上述目标音频数据中包含说话人的音频数据(例如，主播对象的声音数据，即说话人的说话内容)，则可以输出对该识别到的说话人的音频数据进行文本转换所得到的目标语音识别文本。若在上述语音识别处理结果指示识别到目标音频数据中不包含说话人的音频数据，则需要进一步从上述待推荐直播数据中提取视频数据，以识别视频数据中的文本信息，并将识别得到的文本信息等效作为目标语音识别文本。进一步地，为确保语义理解与语义提取的准确性和可靠性，本申请实施例提出可以调用语音转换理解组件的语义理解模型，基于引导提示模板数据对目标语音识别文本进行信息提取，得到目标语义理解信息，从而对该目标语义理解信息进行文本向量表征处理，以得到用于表征上述待推荐直播类媒体对象的目标文本表征向量，即这里的目标文本表征向量是指对目标语音识别文本进行语义理解后所得到的文本特征。应当理解，云服务器(例如，服务器10d)在获取到N个历史直播类媒体对象的N个历史文本表征向量时，将目标文本表征向量分别与N个历史文本表征向量进行向量相似度计算，得到N个向量相似度。再基于N个向量相似度从N个历史直播类媒体对象中筛选出K个历史直播类媒体对象，以直接建立该待推荐直播类媒体对象与K个历史直播类媒体对象之间的相似关联关系，应当理解，本申请实施例通过向量相似检索的方式，能够充分挖掘出这K个历史直播类媒体对象中的每个历史直播类媒体对象与该待推荐直播类媒体对象之间所存在的显示关联关系(例如，每个历史直播类媒体对象均与该待推荐直播类媒体对象高度相似)，进而可以基于挖掘出的显示关联关系直接建立每个历史直播类媒体对象与该待推荐直播类媒体对象相似关联关系，以便于后续可以直接通过这K个历史直播类媒体对象的K歌历史推荐特征，等效计算得到用于推荐该待推荐直播类对象的目标推荐特征。进一步地，获取上述K个历史直播类媒体对象的K个历史推荐特征，将获取到的K个历史推荐特征进行池化处理，从而得到用于表征上述待推荐直播类媒体对象的目标推荐特征。进而使上述目标推荐模型通过上述目标推荐特征确定与待推荐直播类媒体对象相匹配的目标业务对象，以推荐待推荐直播类媒体对象。这样，在目标业务对象为通过第二本地设备(例如，用户终端20a)访问社交软件平台的在线用户时，可以通过该社交软件平台查看到由云服务器所加热推荐的待推荐直播类媒体对象，比如，第二本地设备可以用于通过该社交软件平台播放该待推荐直播类媒体对象的待推荐直播数据。

[0068] 应当理解，待推荐媒体对象与历史媒体对象为同一类型数据，例如若待推荐媒体对象的多媒体数据为直播数据，则历史媒体对象的多媒体数据也应当为直播数据，若待推荐媒体对象的多媒体数据为点播数据，则历史媒体对象的多媒体数据也应当为点播数据，还有更多情况，这里将不再进行赘述。

[0069] 其中，可以理解的是，在本申请实施例中，待推荐媒体对象(例如，待推荐直播类媒体对象)对应的业务终端和目标匹配用户所对应的业务终端可以包括：智能手机、平板电脑、笔记本电脑、桌上型电脑等智能终端。

[0070] 其中，在本申请实施例中，对待推荐媒体对象的多媒体数据进行处理的计算机设备可以是服务器(例如，上述云服务器)，也可以是上述用户终端集群中的任一用户终端(例如，上述第一本地设备)，其中，服务器可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。

[0071] 进一步地，如图2所示，图2是本申请实施例提供的一种数据交互场景示意图。图2中所示的用户终端10a可以用于采集某个需要进行加热推荐的媒体对象(即待推荐媒体对象，该待推荐媒体对象可以为图2所示的待处理媒体对象，例如，图2所示的主播A所在的直播间)的多媒体数据(例如，显示在图2所示的主播A直播界面中的直播数据)，以将采集到的多媒体数据发送给图2所述的服务器10d。其中，主播A直播界面是指主播对象(例如，主播A)所在的直播间的显示界面。

[0072] 应当理解，在实时直播场景下，考虑到随着该主播A所在的直播间的直播进度的进行，在该主播A所在的直播间中所采集到的多媒体数据(即直播数据)的数据信息量可能会相对比较庞大。

[0073] 因此，在一种可实现方式中，为了能够帮助服务器10d快速实现对该直播间中的多媒体数据(即在直播间中以视频和/或音频所开展的直播数据)的深度理解，本申请实施例提出用户终端10a在采集得到多媒体数据(即直播数据)时，可以通过预设的采样窗口(即目标采样窗口)，从该直播间的多媒体数据中截取的一段直播视频(即截取得到视频片段)，作为目标采样窗口中的待推荐直播数据片段，以将该待推荐直播数据片段统称为动态采集到的待推荐多媒体数据，进而可以将由该用户终端10a所动态采集到的各个待推荐直播数据片段(即待推荐多媒体数据)遍历发送给服务器10d，以使服务器10d可以通过关键帧抽取、OCR识别、自动语音识别等方式，从接收到的各个待推荐直播数据片段(即待推荐多媒体数据)中挖掘得到该直播间的文本信息，以便于能够通过当前最新挖掘到的待推荐直播数据片段(即待推荐多媒体数据)的文本信息，动态确定出用于表征该直播间的目标文本表征向量。

[0074] 比如，用户终端10a在采集得到该直播间的多媒体数据(即直播数据)时，可以基于该直播间的直播进度(即媒体数据关联时长)，首先获取与该直播进度(即媒体数据关联时长)相关的目标动态采样时间点、窗口时长和采样间隔时长。应当理解，目标动态采样时间点是指随着待推荐媒体对象(即直播类媒体对象，例如，直播间)直播进度(即媒体数据关联时长)的变化而变化的时间点。比如，本申请实施例可以将直播进度(即媒体数据关联时长，即上述待推荐直播数据的开始时长)为窗口时长时的时间点(即第一个动态采样时间点)称之为第一动态采样时间点，同理，随着直播进度的变化，本申请实施例还可以将新的直播进度(即新的媒体数据关联时长)为该窗口时长时的另一时间点(即第二个动态采样时间点)称之为第二动态采样时间点，这里的第二动态采样时间点为第一动态采样时间点的下一时间点，且第一动态采样时间点与第二动态采样时间点之间的时长为上述采样间隔时长，以此类推，可以将第二动态采样时间点的下一时间点作为第三动态采样时间点。这样，用户终端10a可以先后将随着直播进度的变化而确定出的不同动态采样时间点(例如，第一动态采样时间点、第二动态采样时间点以及第三动态采样时间点)统称为目标动态采样时间点，并将以不同动态采样时间点为数据采样结束时间点时的采样窗口统称为目标采样窗口，进而可以将通过目标采样窗口所截取到的待推荐直播数据片段作为待推荐直播数据，以发送给服务器10d。

[0075] 具体的，对于第一动态采样时间点而言，可以将以第一动态采样时间点为数据采样结束时间点的采样窗口称之为第一采样窗口，同理，也可以将以第二动态采样时间点为数据采样结束时间点的另一采样窗口称之为第二采样窗口。其中，在与该第一动态采样时间点间隔的下一时间点之间的时长为采样间隔时长时，确定出第二动态采样时间点，进而用第二动态采样时间点替换第一动态采样时间点作为更新后的目标动态采样时间点，进而可以将第二采样窗口作为目标采样窗口，以截取得到新的待推荐直播数据(即新的待推荐直播数据片段)，从而达到动态采样的效果，这里将不再赘述。需要注意的是，上述采样间隔时长应小于或者等于上述窗口时长，以使整个直播间的多媒体数据(即直播数据)都能够被采集充分，且上述第一采样窗口的窗口时长与第二采样窗口的窗口时长相同。

[0076] 可选地，在另一种可实现方式中，为避免用户终端10a所采集到的数据信息量可能会比较庞大而造成额外的存储空间的占用，本申请实施例还可以在采集得到该直播间的多媒体数据(例如，直播数据)时，直接将该直播间的多媒体数据(例如，直播数据)以数据推流的方式发送给服务器10d，使得服务器10d可以利用服务器强大的计算能力，按照上述用户终端10a确定目标动态采样时间点的方式，在服务器侧动态确定出不同动态采样时间点，以将以不同动态采样时间点为数据采样结束时间点时的采样窗口统称为目标采样窗口，进而可以将通过目标采样窗口所截取到的待推荐直播数据片段作为待推荐直播数据，以便于后续可以通过自动语音识别和大语言模型集成的方式，在推荐系统中实现对该交互数据量较小的直播间的实时加热推荐，以及对该直播间中的直播数据进行数据分发的时效性和精准性。换言之，本申请在以直播为代表的实时视频化/音频化的推荐场景中，可以通过自动语音识别和大语言模型集成的方式，对动态采集到的该直播间(例如，)中的局部直播视频片段(即通过目标采样窗口所截取到的直播视频片段，例如，上述待推荐直播数据片段)进行更为时效性的内容理解，以便于后续可以从相较于该直播间(例如，冷门直播间)的交互数据量更大的N个历史直播间(例如，热门直播间)中，确定出与该直播间(具体可以是该目标采样窗口内的局部直播视频片段)高度相似的K个历史直播间(具体是通过历史采样窗口所截取到的局部直播视频片段)，以建立显示关联关系，进而可以通过推荐系统所记录到的历史对这K个历史直播间(即因具有较大交互数据量而存在交互特征充足的热门直播间)进行推荐时的K个历史推荐特征，来等效计算出用于对该直播间进行推荐的目标推荐特征，以通过该目标推荐特征来弥补该直播间(例如，冷门直播间)因与用户的交互数据量较小而存在交互特征稀疏的问题。

[0077] 如图2所示，服务器10d在接收到用户终端10a发送的待处理媒体对象(即待推荐媒体对象)的多媒体数据(例如，待推荐直播数据片段)时，通过步骤S11获取多媒体数据中的目标音频数据，也就是本申请实施例在通过目标采样窗口预先采集到或者实时采集到的待推荐直播数据片段时，可以由该服务器10d进一步将从该待推荐直播数据片段中提取到的音频数据，作为目标音频数据。进一步地，服务器10d获取用于对目标音频数据进行语音转换处理的语音转换理解组件和相似检索组件。进而，执行步骤S12，以通过语音识别器进行语音识别处理，也就是使用语音转换理解组件中的语音识别器对目标音频数据进行语音识别处理，得到语音识别处理结果。若语音识别处理结果为识别到目标音频数据的目标语音识别文本(即该语音识别处理结果为识别得到目标音频数据中的说话人的声音数据对应的语音文本，则将识别得到的语音文本作为目标语音识别文本)，则执行步骤S13，以通过语义理解模型进行文本语义理解处理，详细地，也就是说通过上述语音转换理解组件中的语义理解模型基于引导提示模板数据进行文本语义理解处理，得到与目标语音识别文本相关联的目标语义理解信息。进而，执行步骤S14，以通过文本表征器进行文本向量表征处理，也就是使用上述相似检索组件中的文本表征器对目标语义理解信息进行文本向量表征处理(即对目标语义理解信息进行向量化处理，以得到该目标语音视频文本的文本特征)，以得到用于表征上述待推荐媒体对象的目标文本表征向量。

[0078] 其中，具体地，若待推荐媒体对象的类型为视频点播类型，则将与视频点播类型相关联的提示模版数据作为引导提示模板数据。且若语音识别处理结果为未识别到目标音频数据的目标语音识别文本(即该语音识别处理结果为未识别得到目标音频数据中的说话人的声音数据对应的语音文本)，则将该目标音频数据标记为无效音频数据。进而，获取通过目标采样窗口中所采集到的待推荐媒体对象的多媒体数据(例如，待推荐直播数据片段)中的视频数据，以识别该视频数据中的画面(即视频帧)中的文字，将其作为目标语音识别文本。

[0079] 进一步地，服务器10d执行步骤S21，以获取历史文本表征向量。具体地，服务器10d可以从预设的历史媒体对象数据库中获取N个历史媒体对象(例如，N个历史直播类媒体对象)的N个历史文本表征向量，进而通过步骤S31计算向量相似度，也就是分别计算目标文本表征向量与这N个历史文本表征向量中的每个历史文本表征向量的向量相似度，以得到N个向量相似度。如图2所示，服务器10d执行步骤S32，以从N个历史媒体对象(例如，N个历史直播类媒体对象)中筛选向量相似度较大的前TopK个历史媒体对象(例如，K个历史直播类媒体对象)，具体地，也就是在上述N个向量相似度中筛选出相似度靠前的K历史文本表征向量，从而可以筛选出这K个历史文本表征向量分别对应的历史直播类媒体对象，以得到K个历史直

播类媒体对象，进而可以建立待推荐直播类媒体对象(即上述待推荐媒体对象)与这K个历史直播类媒体对象之间的相似关联关系。进一步地，服务器10d执行步骤S33，以获取K个历史媒体对象的历史推荐特征，进而执行步骤S34，基于K个历史推荐特征计算目标推荐特征，需要注意的是，目标推荐特征为该服务器10d所提供的推荐系统用于对该待推荐直播类媒体对象进行媒体推荐(例如，加热推荐)时所需要用到的推荐特征。进一步地，服务器10d可以执行步骤S35，基于目标推荐特征向目标业务对象推荐待推荐媒体对象，也就是服务器10d可以在计算得到可对该待推荐媒体对象进行加热推荐的目标推荐特征时，可以进一步由该推荐系统调用目标推荐模型，使得该目标推荐模型可以基于目标推荐特征从大量用户中定向匹配出与待推荐直播类媒体对象(即待推荐媒体对象)相匹配的用户(即目标业务对象)，以向确定出的目标业务对象推荐该待推荐直播类媒体对象，这样，不仅可以提升对该待推荐媒体对象的加热推荐，还可以提升加热推荐的精准度。

[0080] 其中，上述N个历史直播类媒体对象为在所有历史直播数据中基于历史媒体对象数据库对应的交互数据曲线分布，所抽取的达到交互数据量达到阈值的媒体对象，进而可以从该历史媒体对象数据库所对应的历史文本向量表征数据库中，获取这N个历史直播类媒体对象的历史文本表征向量，以得到N个历史文本表征向量。应当理解，这N个历史文本表征向量由这N个历史直播类媒体对象的历史直播数据经处理得到的，且服务器10d对这N个历史直播类媒体对象的历史直播数据进行处理的具体处理过程与该服务器10d对上述待推荐直播类媒体对象的待推荐直播数据进行处理的具体处理过程相同，这里不进行赘述。需要注意的是，任意一个历史文本表征向量对应的历史音频数据是指相应历史媒体对象在服务器10d所记录到(例如，最后一次记录到的)的历史采样窗口中的音频数据，即服务器10d可以基于历史动态采样时间点为在历史直播数据播放结束时所记录到的结束历史动态采样时间点、以及窗口时长为历史窗口时长的历史采样窗口中，获取到相应历史媒体对象的目标音频数据。其中，历史窗口时长与上述目标采样窗口的窗口时长相同。

[0081] 需要注意的是，想要使目标推荐模型基于目标推荐特征向与待推荐直播类媒体对象相匹配的目标业务对象、推荐该待推荐直播类媒体对象，首先需要获取上述N个历史直播类媒体对象的N个历史交互行为信息，并将上述N个历史交互行为信息作为训练初始推荐模型的训练样本。进一步地，服务器10d获取上述训练样本的推荐指标，并通过上述推荐指标和训练样本对初始推荐模型进行模型训练，得到目标推荐模型。

[0082] 进一步地，在将目标推荐模型部署在用户终端20a的业务处理平台(例如社交软件平台)中时，可以调用目标推荐模型向与上述待推荐直播类媒体对象(例如，图2所示的主播A所在的直播间)的待推荐媒体数据(例如，图2所示的主播A显示界面所对应的待推荐直播数据)相匹配的目标业务对象推荐该待推荐直播类媒体对象，上述目标业务对象为在用户终端20a上访问该社交软件平台的在线用户。进而，服务器10d将上述待推荐直播数据发送给用户终端20a，以使用户终端20a在接收到该待推荐直播类媒体对象的待推荐直播数据时，在推荐显示列表W中显示待推荐直播类媒体对象。

[0083] 具体地，在将推荐系统中的目标推荐模型部署至业务处理平台中(例如社交软件平台)中时，服务器10d可以获取与该业务处理平台(例如社交软件平台)相关联的业务对象的对象关联数据信息，并将对象关联数据信息输入目标推荐模型中，以通过该目标推荐模型进行信息向量表征提取，从而得到业务对象的对象向量表征特征，这样，目标推荐模型可以在业务对象的对象向量表征特征中，筛选与目标推荐特征相匹配的对象向量表征特征，并将筛选出的对象向量表征特征所对应的业务对象作为与待推荐直播类媒体对象相匹配的目标业务对象。进一步地，目标推荐模型可以将待推荐直播类媒体对象作为用于推荐给目标业务对象的目标推荐对象，并将目标推荐对象推送给目标业务对象所对应的终端设备(例如，图2所示的用户终端20a)，以使终端设备(例如，图2所示的用户终端20a)在该业务处理平台的推荐显示列表(例如，图2所示的推荐显示列表W)中显示目标推荐对象。

[0084] 由此可见，本申请实施例通过将待推荐媒体对象的目标文本表征向量与历史媒体对象的历史文本表征向量相关联，可以准确筛选出与待推荐媒体对象相关联的K个历史媒体对象。进一步地，本申请实施例

在获取到这K个历史媒体对象的历史推荐特征时，可以直接基于这K个历史媒体对象的历史推荐特征，处理得到(即等效计算得到)待推荐媒体对象的目标推荐特征，以便于可以使用(即调用)目标推荐模型，通过该目标推荐特征向与待推荐媒体对象相匹配的目标业务对象推荐该待推荐媒体对象。

[0085] 进一步地，如图3所示，图3是本申请实施例提供的一种针对媒体对象的推荐方法的流程示意图。该方法可以由服务器执行，也可以由上述第一本地设备与服务器交互执行。比如，在第一本地设备(例如，上述图1所对应的用户终端10a)采集到待推荐媒体对象的多媒体数据时，将多媒体数据发送给服务器(例如，上述图1所对应实施例中的服务器10d)，使得服务器可以按照本申请实施例所提供的针对媒体对象的推荐方法，对接收到的多媒体数据进行数据处理，以最终定向匹配到与该待推荐媒体对象相匹配的目标业务对象。这样，服务器后续可以将待推荐媒体对象的多媒体数据准确发送给该目标业务对象，即可以在该目标业务对象所使用的第二本地设备(例如，上述图1所对应的用户终端20a)的业务处理平台(例如，社交软件平台)上，通过推荐显示列表显示待推荐媒体对象。应当理解，在本申请实施例中，该针对媒体对象的推荐方法至少可以包括以下步骤S101至步骤S104：

[0086] 步骤S101：获取待推荐媒体对象在目标采样窗口中的目标音频数据，对目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量。

[0087] 其中，目标音频数据为通过目标采样窗口所采集到的待推荐媒体对象的音频数据；目标文本表征向量为将与目标音频数据相关联的目标语音识别文本进行文本向量表征得到的。

[0088] 其中，服务器获取用于表征待推荐媒体对象的目标文本表征向量的具体过程可以描述为：获取待推荐媒体对象在目标采样窗口中的目标音频数据，进而获取用于对目标音频数据进行语音转换处理的语音转换理解组件和相似检索组件。其中，语音转换理解组件包括语音识别器和用于进行语义提取的语义理解模型，且相似检索组件包括文本表征器。进一步的，服务器通过语音识别器对目标音频数据进行语音识别处理(即语音识别器判断该目标音频数据中是否存在说话人的声音数据或者语音数据，如果存在，则可以将说话人的声音数据或者语音数据识别为目标语音识别文本)，得到语音识别处理结果。若语音识别处理结果指示识别到目标音频数据的目标语音识别文本，则通过语义理解模型对目标语音识别文本进行文本语义理解处理，得到目标语音识别文本的目标语义理解信息。通过文本表征器对目标语义理解信息进行向量表征，得到用于表征待推荐媒体对象的目标文本表征向量。

[0089] 可选地，在一种可实现方式中，若语音识别处理结果指示未识别到目标音频数据的目标语音识别文本，则通过目标采样窗口从待推荐媒体对象的多媒体数据中，获取待推荐媒体对象的视频数据，将获取到的视频数据作为待推荐媒体对象在目标采样窗口中的目标视频数据。对目标视频数据中的视频帧进行视频内容识别，得到用于表征目标视频数据的视频文本内容，以通过视频文本内容更新目标语音识别文本。

[0090] 可选地，在另一种实现方式中，若语音识别处理结果指示未识别到目标音频数据的目标语音识别文本，则将目标音频数据标记为无效音频数据。这里的无效音频数据是指通过目标采样窗口从该待推荐媒体对象的多媒体数据中所提取到的目标语音数据，本质上为判断出的不存在说话人的声音数据的音频数据。此时，服务器可以丢弃当前被标记为无效音频数据的目标语音数据，以避免服务器上的存储资源和计算资源的浪费。

[0091] 为便于理解，这里以待推荐媒体对象为直播类媒体对象，通过目标采样窗口所采样得到的待推荐媒体对象的多媒体数据(例如，待推荐直播数据片段)为待推荐直播数据为例，此时，历史媒体对象为历史直播类媒体对象，历史媒体对象的多媒体数据为历史直播数据，应当理解，考虑到在实时直播场景下，通过目标采样窗口所采集得到的上述待推荐直播数据中存在有音频数据(即存在与说话人的声音相关的音频数据)和无音频数据(即不存在与说话人的声音相关的音频数据)两种情况，因此，在从待推荐直播数据中提取到在该目标采样窗口中的音频数据(即目标音频数据)时，可以通过语音转换理解组件中的语音识别器，对目标音频数据进行语音识别处理，得到语音识别处理结果。进一步地，若上述语音识别处理结果表示上述待推荐直播数据中存在音频数据(即存在与说话人的声音相关的音频数据)，则可以对待推荐直播数据中

存在的与说话人的声音相关的音频数据进行语音转换，以将语音转换所得到的语音识别文本作为有该语音识别器所输出的与待推荐直播数据的目标音频数据相关联的目标语音识别文本。应当理解，若上述语音识别处理结果表示上述待推荐直播数据中未含有音频数据(即不存在与说话人的声音相关的音频数据)，则将上述目标音频数据标记为无效音频数据，后续将不再对该无效音频数据进行使用。进而，还可以获取上述待推荐直播数据的视频数据，作为目标视频数据，从而识别该目标视频数据中的文本信息，以将识别到的该目标视频数据中的文本信息作为由该语音识别器所输出的与上述待推荐直播数据相关联的视频识别文本(即视频文本内容)，从而可以将该视频识别文本(即视频文本内容)等效作为目标语音识别文本。

[0092] 其中，在语音识别器采用自动语音识别(Automatic Speech Recognition，ASR)算法进行语音识别的情况下，通过ASR算法进行语音识别，以得到目标语音识别文本的具体方式，具体参见以下公式(1)：

$$text_i = asr(i_c) \quad (1)$$

[0093] 其中，i表示待推荐媒体对象，例如，待加热的冷启动/长尾直播间，简称为当前直播间；c表示当前时间戳(即基于当前直播间的直播进度所最新确定出的目标动态采样时间点)下的目标采样窗口，因此，i_c用于表示通过目标采样窗口从当前直播间的多媒体数据中所采样得到在该目标采样窗口下的待推荐直播数据；asr表示自动语音识别算法，此时，该语音识别器可以通过该自动语音识别算法，从在该目标采样窗口内的待推荐直播数据中提取得到目标语音数据，以对该目标语音数据中所存在的与说话人的声音相关的音频数据进行语音识别，以识别得到用于表征当前直播间(即i)的ASR文本信息(即$text_i$)。

[0094] 进一步地，服务器通过语音转换理解组件中的用于进行语义提取的语义理解模型，对语音识别得到的目标语音识别文本进行语义理解处理，得到目标语义理解信息的具体过程，可以描述为：获取用于引导语义理解模型进行模型输出的引导提示模板数据，将引导提示模版数据和目标语音识别文本输入语义理解模型，通过语义理解模型按照引导提示模版数据中的数据内容，对目标语音识别文本进行语义信息提取，得到目标语音识别文本的目标语义提取信息，进一步的，通过目标语义提取信息确定目标语音识别文本的目标语义理解信息。

[0095] 应当理解，上述语义理解模型在进行语义提取时，需有模板对其进行引导，以输出得到目标语义理解信息。比如，在服务器(例如，上述图1所对应实施例中的服务器10d)所获取到的语义理解模型为大语言模型(Large Language Model，LLM)的情况下，该大语言模型可以基于当前所获取到的引导提示模版数据，具备理解能力和推理能力，比如，大语言模型可以通过引导提示模版实现对目标语音识别文本进行语义理解和推理分析的能力。例如，在待推荐媒体对象的对象类型为视频直播类型(即该待推荐媒体对象为上述待推荐直播类媒体对象)时，该引导提示模板数据具体可以包含与该视频直播类型相关联的提示模版数据，比如，直播类媒体对象风格(例如，主播风格)、直播主题、直播内容、社区氛围、直播间受众和潜在话题等，这里不做限定。

[0096] 进一步地，服务器将引导提示模板数据和目标语音识别文本输入上述大语言模型，由大语言模型根据引导提示模板数据对目标语音识别文本进行语义信息提取，得到与上述目标语音识别文本相关联的目标语义提取信息。进而，由该大语言模型输出包含主播风格、直播主题、直播内容、社区氛围、直播间受众和潜在话题等主题的目标语义理解信息，这里将不对该大语言模型的具体输出内容进行限定。

[0097] 其中，当上述语义理解模型为大语言模型时，对目标语音识别文本进行语义理解的过程如下，具体可以通过以下公式(2)确定目标语义理解信息：

$$semantic_i = llm(promt(text_i)) \quad (2)$$

[0098] 其中，$text_i$是指通过上述公式(1)所识别得到的目标语音识别文本；i表示待加热的冷启动/长尾直播间；prompt是指LLM文本生成模板(即上述引导提示模板数据)，llm是预训练的大语言模型；由于ASR文本信息(即目标语音识别文本)和生成的直播间多维度语义信息(即$semantic_i$)可以归类为自然语言，因此，本申请实施例可以应用LLM的推理能力，按照引导提示模版数据对目标语音识别文本进行推理分析，以便于可

以对推理分析所得到的推理分析结果进行多维度语义信息提取，进而可以将提取到的多维度语义信息统称为目标语音识别文本的目标语义提取信息。

[0099] 需要注意的是，若待推荐媒体对象的对象类型为视频直播类型，则将与视频直播类型相关联的提示模版数据作为引导提示模板数据。可选地，若待推荐媒体对象的对象类型为视频点播类型，则将与视频点播类型相关联的提示模版数据作为引导提示模板数据。

[0100] 应当理解，在待推荐媒体对象的多媒体数据为待推荐直播数据时，上述引导提示模板数据为与视频直播类型相关联的提示模版数据。同理，若上述待推荐媒体对象的多媒体数据为视频点播数据时，上述引导提示模板数据为与视频点播类型相关联的提示模版数据，例如，人物台词、人物风格、场景风格等主题，这里不做限定。

[0101] 进一步地，本申请实施例可以通过文本表征器对上述语义理解模型输出的目标语义理解信息进行向量表征(即文本向量表征处理)，进而，得到用于后续相似度检索处理的目标文本表征向量。

[0102] 步骤S102：获取用于表征N个历史媒体对象的N个历史文本表征向量，对目标文本表征向量和每个历史文本表征向量分别进行向量相似计算，得到N个向量相似度。

[0103] 其中，N个历史文本表征向量中的每个历史文本表征向量是通过语音转换理解组件和相似检索组件，对每个历史媒体对象的历史音频数据进行语音转换处理后所得到的；这里的N个向量相似度为目标文本表征向量与N个历史文本表征向量中的每个历史文本表征向量分别进行向量相似度计算所得到的。

[0104] 其中，上述目标文本表征向量是通过语音转换处理的语音转换理解组件和相似检索组件对上述目标音频数据进行语音转换处理后得到的。

[0105] 其中，服务器获取用于表征N个历史媒体对象的N个历史文本表征向量的具体过程包括：获取历史媒体对象数据库所对应的交互数据曲线分布。其中，交互数据曲线分布是基于历史媒体对象数据库中的各个历史媒体对象和各个历史媒体对象的交互数据量所确定的。进一步的，通过交互数据曲线分布，从历史媒体对象数据库中获取交互数据量达到交互数据阈值的N个历史媒体对象，进而从与N个历史媒体对象相关联的历史文本向量表征数据库中，获取N个历史媒体对象的N个历史文本表征向量。其中，N个历史文本表征向量中的各个历史文本表征向量是通过语音转换理解组件和相似检索组件，对各个历史媒体对象的历史音频数据进行语音转换处理后所得到的。此外，各个历史音频数据是指该N个历史媒体对象中的各个历史媒体对象在记录到的历史采样窗口中的音频数据。

[0106] 为便于理解，请参见图4，图4是本申请实施例提供的一种历史媒体对象的交互数据曲线的示意图。其中，服务器所获取到的历史数据库所具体可以包括历史媒体对象数据库和历史文本向量表征数据库。此时，服务器可以基于该历史媒体对象数据库中所记录到的所有历史媒体对象(例如，历史直播间)，统计这些历史媒体对象(例如，历史直播间)的交互数据量，进而可以按照统计到的这些交互数据量的大小，得到图4所示的交互数据曲线分布。这样，服务器可以基于为该历史媒体对象数据库所配置的交互数据阈值，将这些历史媒体对象划分为位于头部区域的媒体对象(例如，交互数据量达到交互数据阈值的媒体对象，比如，图4所示的交互数据量处于前10%的媒体对象)和位于尾部区域的媒体对象(例如，交互数据量未到达交互数据阈值的媒体对象，比如，图4所示的交互数据量处于后90%的媒体对象)。

[0107] 比如，如图4所述，位于头部区域的媒体对象可以为在该交互数据曲线分布中，交互数量处于前10%的历史媒体对象，进而可以将交互数量处于前10%的历史媒体对象统称为本申请实施例所涉及的需要参与到加热推荐这一推荐过程中的N个历史媒体对象，因上述待推荐媒体对象的多媒体数据的类型为直播数据，则历史媒体对象的多媒体数据的类型也应为直播数据。进而，服务器可以从上述历史文本向量表征数据库中获取上述筛选出的这N个历史媒体对象的N个历史文本表征向量，比如，具体可以是获取到这N个历史媒体对象在历史采样窗口下的N个历史直播数据，以便于后续可以快速筛选出这N个历史直播数据分别对应的历史文本表征向量，以最终得到N个历史文本表征向量。

[0108] 应当理解，上述N个历史直播类媒体对象的N个历史文本表征向量的生成过程，与上述待推荐媒体对象的目标文本表征向量生成过程相同。需要注意的是，上述目标音频数据为目标动态采样时间点随着待推荐直播数据的播放变化(即随着待推荐媒体对象的媒体数据关联时长的变化而变化)时，通过目标采样窗口中所动态采集到的时长为窗口时长的音频数据，即随着目标动态采样时间点的动态不断变化，该服务器所动态采集到的位于目标采样窗口内的音频数据也会不断变化，因此，当服务器在当前动态采样时间点对应的采样窗口(例如，第二采样窗口)内所采集到的说话人的音频数据，与该服务器在上一动态采样时间点对应的另一采样窗口(例如，第一采样窗口)内所采集到的说话人的音频数据为同一音频数据时，可以认为当前在第二采样窗口中所采集到的音频数据为与在第一采样窗口中所采集到的音频数据重复，故而，为避免重复采集到的音频数据(比如，可能因主播对象的说话内容重复而导致采集到的音频数据重复)，对该服务器的存储资源和计算资源的浪费现象，可以将当前在第二采样窗口中所采集到的音频数据统称为上述无效音频数据。

[0109] 其中，对于通过图4所示的交互数据曲线分布所确定出的N个历史媒体对象中的任一历史媒体对象而言，历史直播数据所对应的历史音频数据具体可以为通过历史采样窗口(例如，该历史采样窗口可以为记录到的历史动态采样时间点为历史直播数据结束时的时间点、且采样时长为窗口时长的采样窗口)所采集到的音频数据。进而，通过上述语音转换理解组件和相似检索组件对上述历史音频数据进行语音转换处理，从而得到历史文本表征向量。需要注意的是，目标采样窗口的窗口时长与历史采样窗口的窗口时长相同。

[0110] 其中，具体地，当上述语义理解模型为大语言模型时，分别对上述待推荐直播数据和历史直播数据进行语义提取的实现方式如下，具体可以通过以下公式(3)确定与上述目标语音识别文本相关联的目标语义理解信息，通过以下公式(4)确定与上述历史语义理解信息相关联的历史文本表征向量：

$semantic_i = llm(promt(text_{ci}))i \in C$ (3)

$semantic_j = llm(promt(text_{cj}))j \in H$ (4)

[0111] 其中，$text_{ci}$是指按照上述公式(1)中所识别得到的在目标时间窗口下的目标语音识别文本，同理，$text_{cj}$是指按照上述公式(1)中所记录到的历史识别得到的在历史时间窗口下的历史语音识别文本；其中，i表示待推荐媒体对象(例如，当前媒体对象i可以为待加热的当前直播间，比如，冷启动/长尾直播间)；j表示历史媒体对象(例如，历史高交互直播间，这里的历史高交互直播间是指位于上述头部区域的历史直播间)；ci表示当前时间戳(即基于当前直播间的直播进度所最新确定出的目标动态采样时间点)下的目标采样窗口，cj表示历史时间戳(即基于历史直播间的直播进度所最新确定出的历史动态采样时间点，例如，以历史直播间在直播结束时的时间点为结束历史动态采样时间点)下的历史采样窗口，考虑到历史采样窗口与目标采样窗口的窗口时长保持一致，故而存在ci＝cj＝c，即这里的c可以统称为采样窗口；prompt是指上述公式(2)中的LLM文本生成模板，即上述引导提示模板数据，llm是预训练的大语言模型，C表示当前需要加热的所有待推荐媒体对象的总和(即所有待加热的待推荐媒体对象所构成的待加热媒体对象列表)；H表示N个历史媒体对象中的所有历史媒体对象的总和(即N个历史媒体对象所构成的历史媒体对象列表)。

[0112] 进一步地，本申请实施例分别对待推荐媒体对象的目标语义理解信息和历史媒体对象的历史语义理解信息进行文本向量表征处理的具体实现方式如下，具体可以通过以下公式(5)确定与目标语义理解信息相关联的目标文本表征向量，通过以下公式(6)确定与历史语义理解信息相关联的历史文本表征向量，且可以通过公式(7)得到目标文本表征向量与历史文本表征向量的向量相似度：

$e_i = embedding(semantic_i)i \in C$ (5)

$e_j = embedding(semantic_j)j \in H$ (6)

$s_{i,j} = similarity(e_i, e_j)$ (7)

[0113] 其中，i表示所有待加热的直播间(即C)中的任一当前直播间(例如，冷启动/长尾直播间)；j表示N个历史高交互直播间(即H)中的任一历史高交互直播间；$semantic_i$是指上述公式(3)中的目标语义理解信息，$semantic_j$是指上述公式(4)中的历史语义理解信息；embedding()表示文本向量化算法；similarity()表示向量相

似度度量函数，比如，similarity(e_i,e_j)可以表示计算对目标文本表征向量(即e_i)和某个历史文本表征向量(即e_j)进行向量相似计算，以计算得到当前直播间与某个历史高交互直播间(统称为历史直播间)之间的向量相似度(即$s_{i,j}$)。

[0114] 进一步地，在通过采样窗口(例如，历史采样窗口)获取到上述N个历史直播间中的N个历史直播数据时，可以采用公式(4)和公式(6)计算得到这N个历史直播数据的N个历史文本表征向量，进而可以按照公式(7)，将与待推荐媒体对象(例如，待推荐直播类媒体对象，比如，当前直播间)相关联的目标文本表征向量分别与N个历史媒体对象(例如，N个历史直播类媒体对象，比如，N个历史直播间)的N个历史文本表征向量进行向量相似度计算。其中，向量相似度计算的方法包括余弦相似性、杰卡德系数相似性、欧几里得距离以及曼哈顿距离等，这里不做限定。

[0115] 步骤S103：在通过N个向量相似度从N个历史媒体对象中筛选出K个历史媒体对象时，建立K个历史媒体对象与待推荐媒体对象之间的相似关联关系。

[0116] 具体地，服务器可以基于N个向量相似度对N个历史媒体对象进行排序处理，得到与N个历史媒体对象相关联的历史媒体对象列表。从历史媒体对象列表中获取向量相似度满足相似检索条件的K个历史媒体对象，建立K个历史媒体对象与待推荐媒体对象之间的相似关联关系。其中，相似关联关系用于表征由K个历史媒体对象所构成的列表为相似媒体对象列表，且在相似媒体对象列表中的任意一个媒体对象均与待推荐媒体对象存在关联关系(即上述显示关联关系)。

[0117] 为便于理解，服务器(例如，上述图1所对应实施例中的服务器10d)可以根据上述计算出的N个向量相似度，由高到低对获取到的N个历史直播类媒体对象进行排序处理，得到与这N个历史直播类媒体对象相关联的历史媒体对象列表。进一步地，服务器可以基于相似检索条件(例如，筛选前K个的检索条件)，从上述历史媒体对象列表中筛选出前K个历史直播类媒体对象，也就是其中的TopK的历史直播类媒体对象，以将这K个历史直播类媒体对象与上述待推荐直播类媒体对象建立相似关联关系。

[0118] 应当理解，如图5所示，图5是本申请实施例提供的一种相似关联关系示意图。如图5所示的待处理媒体对象为需要加热推荐的待推荐媒体对象，为便于理解，这里以N＝3，K＝2为例，计算目标文本表征向量(即用于表征该待处理媒体对象的文本表征向量)与N个历史文本表征向量(此处N＝3，即3个历史文本表征向量5b，例如，图5所示的历史文本表征向量A、历史文本表征向量B以及历史文本表征向量C)的向量相似度，以得到3个向量相似度5c(例如，图5所示的向量相似度A、向量相似度B以及向量相似度C)。进而，将3个历史媒体对象5a(例如，图5所示的历史媒体对象A、历史媒体对象B以及历史媒体对象C)根据上述计算出的3个向量相似度5c由高到低对这3个历史媒体对象5a(具体可以是对获取到的这3个历史媒体对象5a的历史直播数据，即3个历史直播数据)进行排序处理，以得到与这3个历史直播数据相关联的历史媒体对象列表。进一步地，本申请实施例基于上述历史媒体对象列表筛选其前K个历史媒体对象的历史直播数据(此处K＝2，即2个历史直播数据)，得到与待处理媒体对象(即待推荐媒体对象)相关联的2个历史媒体对象。

[0119] 应当理解，如图5所述，本申请实施例在从历史媒体对象列表中筛选出前2个历史媒体对象(例如，图5所示的媒体对象A和媒体对象C，比如，具体还可以是这2个历史媒体对象的历史直播数据)之后，可以直接与待推荐媒体对象(例如，该待推荐媒体对象的待推荐直播数据)建立图5所示的相似关联关系。进而可以基于上述筛选出的这2个历史媒体对象(例如，图5所示的媒体对象A和媒体对象C)，构建得到相似媒体对象列表，其中，可以理解的是，图5所示的相似关联关系可以用于表征该相似媒体对象列表中的每个历史媒体对象(例如，每个历史直播类媒体对象)均与待推荐媒体对象(例如，待推荐直播类媒体对象)存在关联关系(即显示关联关系)。

[0120] 其中，本申请实施例可以根据上述计算出的N个向量相似度，为每个待推荐媒体对象检索出TopK个历史媒体对象(即K个历史高交互媒体对象)，以通过这K个历史高交互媒体对象构成上述相似媒体对象列表，这里的相似媒体对象列表可以为下述公式(8)中的语义近邻列表：

$r_i \leftarrow [J_1 \cdots J_K]$ (8)

[0121] 其中，$[J_1 \cdots J_K]$用于表示为当前直播间(例如，直播间i)所检索出的K个历史高交互媒体对象所构成的语义临近列表，其中，符号J的下标为检索出的这K个历史高交互媒体对象的身份识别(Identity document，ID)索引；K表示通过向量相似度检索方式所检索出的与当前直播间(例如，直播间i)最相似的K个历史媒体对象(例如，K个最相似的历史高交互直播间)。

[0122] 步骤S104：在获取到K个历史媒体对象的K个历史推荐特征时，基于相似关联关系和K个历史推荐特征，确定用于表征待推荐媒体对象的目标推荐特征，通过目标推荐特征对待推荐媒体对象进行媒体推荐。

[0123] 具体地，服务器可以在获取到K个历史媒体对象的K个历史推荐特征时，基于所述相似关联关系对这K个历史推荐特征进行特征池化处理，以得到与所述K个历史推荐特征相关联的池化处理特征；进一步的，服务器可以将池化处理特征，确定为用于表征所述待推荐媒体对象的目标推荐特征，以通过该目标推荐特征实现对该待推荐媒体对象的媒体推荐(例如，加热推荐)。

[0124] 比如，在K个历史媒体对象为K个历史直播类媒体对象时，可以从与这K个历史直播类媒体对象相关联的推荐特征数据库中获取到K个历史直播类媒体对象的K个历史推荐特征，进而可以对获取到这K个历史推荐特征进行池化处理，以得到用于表征待推荐直播类媒体对象的目标推荐特征。进一步地，本申请实施例还可以由服务器调用推荐系统中的目标推荐模型，基于该目标推荐特征确定出与上述待推荐直播类媒体对象相匹配的目标业务对象，以向该目标业务对象推荐该待推荐直播类媒体对象。

[0125] 本申请实施例可以识别待推荐媒体对象在目标采样窗口中的音频数据，得到与待推荐媒体对象相关联的目标音频数据。进一步地，本申请实施例通过语音转换理解组件中的语音识别器对上述目标音频数据进行语音识别处理，进而得到与目标音频数据相关联的目标语音识别文本。进一步地，本申请实施例可以通过语音转换理解组件中的语义理解模型，基于引导提示模板数据对上述目标语音识别文本进行信息提取，得到与上述目标语音识别文本相关联的目标语义理解信息。进而，对上述目标语义理解信息进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量。这样，本申请实施例可以获取用于表征N个历史媒体对象(例如，可以基于上述交互数据曲线分布获取位于头部区域中的N个历史媒体对象)的N个历史文本表征向量，进而将上述目标文本表征向量与N个历史文本表征向量中的每个历史文本表征向量分别进行向量相似计算，以得到N个向量相似度。进一步地，本申请实施例基于上述N个向量相似度，从上述N个历史媒体对象中筛选出K个与待推荐媒体对象相似度高的历史媒体对象，进而，建立待推荐媒体对象与筛选出的K个历史媒体对象的相似关联关系。本申请实施例可以进一步在获取到上述K个历史媒体对象的K个历史推荐特征时，基于上述相似关联关系对这K个历史推荐特征进行特征池化处理，以等效确定出用于表征该待推荐媒体对象的目标推荐特征，进而可以通过目标推荐特征向与待推荐媒体对象相匹配的目标业务对象进行媒体推荐。本申请实施例通过对目标音频数据依次进行语音识别处理、语义理解信息提取和文本向量表征处理，可以得到用于表征目标音频数据的目标文本表征向量，并通过将该目标文本表征向量分别与各个历史媒体对象的历史文本表征向量进行相似度计算，可以最终检索到与待推荐媒体对象高度相似的K个历史媒体对象，从而通过检索到的K个历史媒体对象的历史推荐特征，等效计算得到用于表征待推荐媒体对象的目标推荐特征，换言之，本申请实施例可以在待推荐媒体对象即使为新引入的缺乏交互数据量的冷启动直播间的情况下，也可以准确检索出的与该冷启动直播间的语义高度相似的其他历史直播间，进而可以通过检索出的这些的历史直播间中的历史直播数据所确定出的历史推荐特征，来等效计算得到该冷启动直播间的目标推荐特征(即可以得到模拟得到用于表征该冷启动直播间的高交互特征)，以便于后续可以通过推荐模型(即上述目标推荐模型)精确地检索到待推荐媒体对象所对应的目标业务对象，以向该目标业务对象定向推送该待推荐媒体数据，这样，不仅提高了目标推荐模型对待推荐媒体对象进行媒体推荐(例如，冷启动推荐)的精准度，还可以提升定向为该待推荐媒体对象匹配到合适的业务对象的匹配准确度，还达到向业务对象加热推荐该待推荐媒体对象的效果。

[0126] 在一个实施例中，当待推荐媒体对象的类型为直播类媒体对象时，将该待推荐直播类媒体对象作为在推荐业务(例如，直播推荐业务)中用于推广给与该直播类媒体对象相匹配的用户A(即当前确定出的与待

推荐媒体对象相匹配的目标业务对象)的目标推荐对象。这样，在与该直播类媒体对象相匹配的目标业务对象(例如，用户A)使用该业务处理平台(例如，社交软件平台)的其他业务(例如，XX业务)时，可以在目标业务对象(例如，用户A)正在使用的XX业务中插入该直播类媒体对象或者插入针对该直播类媒体对象的广告，进而实现向目标业务对象(例如，用户A)推广该直播类媒体对象。

[0127] 为便于理解，如图6所示，图6是本申请实施例提供的一种在业务处理平台上对直播类媒体对象进行媒体推荐的场景示意图。当目标业务对象6a(例如，图6中的用户A)在使用业务处理平台6c的其他业务(例如，图6中的XX业务)时，该目标业务对象6a所使用的本地设备(即第二本地设备，例如，图1所对应实施例中的用户终端20a)，可以在该XX业务所对应的XX页面上为该用户A显示目标推荐对象(例如，图6中的待推荐直播类媒体对象6b)，进而实现向用户A推荐该待推荐直播类媒体对象6b。

[0128] 进一步地，第二本地设备(例如，图1所对应实施例中的用户终端20a)响应于目标业务对象对图6所示的待推荐直播类媒体对象6b的触发操作，而跳转显示该待推荐直播类媒体对象的多媒体数据显示界面(即图7所示的在直播间进行直播的场景示意图)，这里的直播间即为待推荐直播类媒体对象。

[0129] 为便于理解，如图7所示，图7是本申请实施例提供的一种在直播间进行直播的场景示意图。其中，图7所示的待处理媒体对象可以为上述需要加热推荐的待推荐媒体对象(即待推荐直播类媒体对象)的多媒体数据显示界面具体包括待处理媒体对象头像显示区域7a和待处理媒体对象显示界面互动区域7b。应当理解，在待推荐直播类媒体对象为直播间时，该多媒体数据显示界面可以为目标业务对象点击上述图6中待推荐直播类媒体对象6b所在区域后，第二本地设备(例如，图1所对应实施例中的用户终端20a)响应于目标业务对象对上述图6待推荐直播类媒体对象6b的触发操作，而跳转显示的待推荐直播类媒体对象的多媒体数据显示界面。

[0130] 相应地，在另一个实施例中，当待推荐媒体对象的类型为点播类媒体对象时，将该点播类媒体对象作为在推荐业务(例如，直播推荐业务)中用于推广给与该点播类媒体对象相匹配的用户A(即当前确定出的与待推荐媒体对象相匹配的目标业务对象)的目标推荐对象。然后，本申请实施例可以通过将该点播类媒体对象所对应的点播链接封装为推荐通知消息，发送给该目标业务对象对应的终端设备，例如，第二本地设备，以使得第二本地设备可以将该推荐通知消息以弹窗、悬浮窗等方式显示在目标业务对象(例如，用户A)的使用页面上。应当理解，目标业务对象(例如，用户A)可通过对该推荐通知消息进行点击，进而使得该第二终端设备可以进入该点播类媒体对象的播放界面。

[0131] 为便于理解，如图8所示，图8是本申请实施例提供的一种在业务处理平台上对点播类媒体对象进行媒体推荐的场景示意图。第二本地设备(例如，图1所对应实施例中的用户终端20a)可以在目标业务对象8a(例如，图8中的用户A)的使用页面(即该用户A所访问的业务处理平台8c的显示页面)上，以图8所示的弹窗方式显示推荐通知消息8b。进而，目标业务对象8a(例如，图8所示的在业务处理平台8c中的用户A)可以通过点击该推荐通知消息，从而进入上述点播类媒体对象的多媒体数据显示界面(即上述播放界面)，从而向上述用户A推荐该点播类媒体对象。

[0132] 换言之，第二本地设备(例如，图1所对应实施例中的用户终端20a)可以响应于目标业务对象对该推荐通知消息的触发操作，以跳转显示该点播类媒体对象的多媒体数据显示界面。其中，点播类媒体对象的多媒体数据显示界面可以包括图9所示的待播放进度条9a。

[0133] 为便于理解，如图9所示，图9是本申请实施例提供的一种基于点播标识进行视频播放的场景示意图。应当理解，上述点播标识为用于标识待推荐点播类媒体对象的标识(例如，该点播标识具体可以为图9所示的待处理媒体对象名称)。该多媒体数据显示界面可以为目标业务对象(例如，图8所示的用户A)点击推荐通知消息8b后，由第二本地设备(例如，图1所对应实施例中的用户终端20a、20b、20c和图2所对应的用户终端20a)响应于目标业务对象的对该推荐通知消息的触发操作，而跳转显示的界面。

[0134] 在一个实施例中，在通过向量相似度检索得到TopK个历史直播类媒体对象(即历史媒体对象)之后，就可以基于这TopK个历史直播数据的历史推荐特征，等效计算得到目标推荐特征，应当理解，该目标推

荐特征可以用于查找与待推荐直播类媒体对象(即待推荐媒体对象)相匹配的在线匹配用户(即目标业务对象)，以使第二本地设备后续可以在社交软件平台(即业务处理平台)上为匹配到的在线匹配用户显示推荐显示列表，进而可以在该推荐显示列表上显示该待推荐直播类媒体对象(即待推荐媒体对象)。

[0135] 为便于理解，如图10所示，图10是本申请实施例提供的一种业务处理平台的推荐显示列表的示意图。其中，该推荐显示列表包含主播显示区域100a、推荐列表区域100b，进一步地，推荐列表区域100b中含有上述主播显示区域100a中每个主播所对应的媒体对象(例如，图10所示的主播A1对应的媒体对象A1、主播B1对应的媒体对象B1、主播C1对应的媒体对象C1以及主播D1对应的媒体对象D1)，其中,主播A1对应的媒体对象A1可以为由服务器所加热推荐的待处理直播类媒体对象100c(即上述待推荐直播类媒体对象)。

[0136] 这样，在上述目标推荐模型筛选出与待推荐媒体对象相匹配的目标业务对象后，第二本地设备(例如，图1所对应实施例中的用户终端20a)可以将上述接收到的待推荐媒体对象推送给目标匹配用户(即目标业务对象)，以使目标匹配用户可以在社交匹配平台上的推荐显示列表中查看到该待推荐媒体对象。需要注意的是，推荐列表区域100b中包括上述待推荐直播类媒体对象100c，这样，该目标匹配用户可以在图10所示的该待推荐直播类媒体对象100c的多媒体数据显示界面上，通过点击推荐列表区域100b中的媒体对象A1区域进行跳转显示。

[0137] 本申请实施例主要通过对待推荐媒体对象的音频数据(例如，待推荐媒体对象中的与说话人的声音相关的音频数据)进行语音识别处理，进而可以对识别到的语音文本进行处理，以便于后续可以查找到与待推荐媒体对象相匹配的目标业务对象。其中，若语音识别处理结果表示在上述待推荐直播数据中所提取到的目标音频数据中存在与说话人的声音相关音频数据，则输出与该待推荐媒体对象的目标音频数据相关联的目标语音识别文本。可选地，若语音识别处理结果表示在上述待推荐直播数据中所提取到的目标音频数据中未含有与说话人的声音相关的音频数据，则将上述目标音频数据标记为无效音频数据，后续将不再进行使用。应当理解，在将目标音频数据标记为无效音频数据时，后续可能分为两种情况，一种情况为直接结束，不再对该待推荐媒体对象进行处理。另一种情况为可以进一步获取待推荐媒体对象的视频数据，作为目标视频数据，并对上述目标视频数据中的文字进行识别，以将识别到的文字间接作为目标语音识别文本进行处理。

[0138] 其中，如图11所示，图11是本申请实施例提供的一种目标音频无效处理方法流程图。该方法可由服务器执行，这样，服务器(例如，上述图1所对应实施例中的服务器10d)在获取到待推荐媒体对象的目标音频数据后，通过语音转换理解组件中的语音识别器对该目标音频数据进行语音识别处理，以识别该目标音频数据中是否存在语音音频(即与说话人的声音相关的音频数据)。当识别到目标音频数据中不存在与说话人的声音相关的音频数据时，可以将该目标音频数据标记为无效音频数据，以直接结束该方法流程。比如，如图11所示，该方法至少可以包括步骤S2011-步骤S2016。

[0139] 步骤S2011：获取目标音频数据。

[0140] 其中，具体地，获取待推荐媒体对象的目标音频数据，并在对目标音频数据进行语音识别处理的过程中，检测或者识别该目标音频数据中是否存在语音音频，这里的语音音频可以为与说话人的声音相关的音频数据。

[0141] 步骤S2012：检测目标音频数据中是否存在音频。

[0142] 其中，具体地，使用语音转换理解组件中的语音识别器对上述目标音频数据进行语音识别处理，得到语音识别处理结果。其中，检测目标音频数据中是否存在音频(例如，语音音频)，如果存在语音音频，即在确定当前检测到目标音频数据中存在与说话人的声音相关的音频数据时，进一步执行步骤S2013。其中，服务器对目标音频数据进行语音识别处理的具体过程，可以参见上述图3所对应实施例中对语音识别处理的具体过程的描述，这里将不再继续进行赘述。

[0143] 步骤S2013：识别目标音频数据中的音频对应的文本。

[0144] 其中，具体地，当检测到目标音频中存在语音音频，即存在与说话人的声音相关的音频数据时，可以进一步识别得到该目标音频数据中的语音音频对应的文本(即识别得到该目标音频数据的目标语音识别文本)。

[0145] 步骤S2014：输出目标语音识别文本。

[0146] 其中，具体地，服务器在识别到上述目标音频数据中的音频对应的文本时，可以将识别到的文本的具体文本信息作为目标语音识别文本进行输出。这样，在完成该步骤后，可以进一步跳转执行步骤S2016。

[0147] 可选地，在执行完步骤S2012之后，还可以判断为否，即在确定当前检测到目标音频数据中不存在与说话人的声音相关的音频数据时，跳转执行步骤S2015。

[0148] 步骤S2015：将目标音频数据标记为无效音频数据。

[0149] 其中，具体地，服务器可以在检测到目标音频(即目标音频数据)中不存在音频(即语音音频，例如，与说话人的声音相关的音频数据)时，将该目标音频数据标记为无效音频数据。

[0150] 步骤S2016：结束。

[0151] 其中，服务器还可以进一步通过语音转换理解组件中的语义理解模型对目标语音识别文本进行文本语义理解处理，得到目标语音识别文本的目标语义理解信息。其中，对目标语音识别文本进行文本语义理解处理的具体过程，可以参见上述图3所对应实施例中对文本语义理解处理的具体过程的描述，这里将不再继续进行赘述。

[0152] 进一步地，如图12所示，图12是本申请实施例提供的另一种目标音频无效处理方法流程图。如图12所示，该方法至少可以包括步骤S2021-步骤S2028。

[0153] 步骤S2021：获取目标音频数据。

[0154] 其中，服务器获取目标音频数据的具体过程，可以参见上述图3所对应实施例中对获取目标音频数据的具体过程的描述，这里将不再继续进行赘述。

[0155] 步骤S2022：检测目标音频数据中是否存在音频(例如，语音音频)。

[0156] 步骤S2023：识别目标音频数据中的音频对应的文本。

[0157] 其中，具体地，当服务器检测到目标音频数据中存在音频(例如，语音音频，比如，与说话人的声音相关的音频数据)时，可以识别得到该目标音频数据中的音频对应的文本(即识别得到针对该目标音频数据的目标语音识别文本)。这样，服务器在完成该步骤后，可以进一步跳转执行步骤S2027。

[0158] 可选地，如图12所示，服务器在执行完步骤S2022之后，还可以在判断为否时，跳转执行步骤S2024。

[0159] 步骤S2024：将目标音频数据标记为无效音频数据。

[0160] 其中，具体地，服务器在检测到目标音频在不存在音频(例如，不存在语音音频，比如，不存在与说话人的声音相关的音频数据)时，将该目标音频数据标记为无效音频数据，以便于后续可以进一步执行步骤S2025。

[0161] 步骤S2025：获取目标音频数据对应的视频数据。

[0162] 其中，具体地，服务器可以获取上述待推荐媒体对象的目标音频数据所对应的视频数据，并将该视频数据作为目标视频数据。其中，服务器获取目标视频数据的具体过程，可以参见上述图3所对应实施例中对获取目标视频数据的具体过程的描述，这里将不再继续进行赘述。

[0163] 步骤S2026：识别目标视频数据中的文本。

[0164] 其中，具体地，服务器可以对上述待推荐媒体对象的目标视频数据进行文本识别处理。该服务器对目标视频数据进行文本识别处理的具体过程，可以参见上述图3所对应实施例中对识别得到视频文本内容的具体过程的描述，这里将不再继续进行赘述。

[0165] 步骤S2027：输出目标语音识别文本。

[0166] 步骤S2028：结束。

[0167] 本申请实施例主要为针对待处理媒体对象(即待推荐媒体对象)在冷启动或者长尾情况下对其进行加热的方法。该方法指出在获取到待推荐媒体对象的目标音频数据时，可以对其进行语音识别处理，以得到目标语音识别文本。进而，本申请实施例可以进一步对该目标语音识别文本进行语义理解处理后，再将其进行文本向量表征处理，得到目标文本表征向量。进一步地，本申请实施例在获取到用于表征N个历史媒体对象的N个历史文本表征向量时，可以将目标文本表征向量分别与上述N个历史文本表征向量进行向量相似度计算。进而可以基于上述计算出的向量相似度筛选出相似度最高的K个历史媒体对象，以与筛选出的K个历史媒体对象建立相似关联关系。进一步地，本申请实施例可以基于相似关联关系获取这K个历史媒体对象的K个历史推荐特征，进而可以通过对上述K个历史推荐特征进行池化处理，得到用于表征待推荐媒体对象的目标推荐特征，这样，服务器在基于上述目标推荐特征进行媒体推荐时，可以增强对该待推荐媒体对象的加热推荐效果。

[0168] 为便于理解，如图13所示，图13是本申请实施例提供的一种针对媒体对象的目标业务对象预估流程示意图。本申请实施例在获取到的待加热媒体对象(即上述待推荐媒体对象)为长尾/冷启媒体对象13a时，可以在目标采样窗口中采集到采集时长为窗口时长的目标音频数据，进而通过音频数据语音识别模块13b对目标音频数据进行语音识别处理，从而得到目标语音识别文本。需要注意的是，在使用上述目标采样窗口对待推荐媒体对象的多媒体数据中进行采样时本质上为实时动态采样，该实时动态采样的具体过程，可以参见上述图3所对应实施例中对通过不同动态采样时间点所对应的采样窗口，采样得到不同音频数据的具体过程的描述，这里将不再继续进行赘述。进一步地，如图13所示，本申请实施例可以基于音频数据语音识别模块13b的引导提示模板数据，使得语义理解模型可以对目标语音识别文本进行语义信息提取，得到目标语义理解信息。进而，将上述目标语义理解信息进行文本向量表征处理，得到目标文本表征向量。需要注意的是，这里的音频数据语音识别模块13b所起到的作用等同于上述语音转换理解组件。

[0169] 进一步地，本申请实施例可以通过向量相似度计算模块13c获取用于表征N个历史媒体数据的N个历史文本表征向量，并将上述目标文本表征向量分别与这N个历史文本表征向量进行向量相似度计算。需要注意的是，这里的N个历史文本表征向量的获取方式与上述目标文本表征向量的获取方式相同，但历史音频数据为通过历史采样窗口(比如，在历史动态采样时间点为历史媒体对象的多媒体数据的结束时间点所对应的采样窗口)所采集到的采样时长为窗口时长的音频数据，同理，这里的向量相似度计算模块13c所起到的作用等同于上述相似检索组件。

[0170] 进而，本申请实施例可以通过上述计算出的N个向量相似度筛选出相似度最高的K个历史媒体对象，并建立待推荐媒体对象与这K个历史媒体对象的相似关联关系，以得到高交互语义TopK列表(即相似媒体对象列表，该相似媒体对象列表用于存储具备较高交互数据量、且与该待推荐媒体对象存在较高语义相似度的K个历史媒体对象)。进一步地，本申请实施例可以基于目标推荐特征确定模块13d获取上述相似媒体对象列表中的K个历史媒体对象的K个历史推荐特征，并对上述K个历史推荐特征进行池化处理，从而得到用于加热推荐该待推荐媒体对象的目标推荐特征。然后，本申请实施例可以使用推荐模型13e基于上述目标推荐特征，从当前访问该业务处理平台(例如，社交软件平台)的所有在线用户中查找出与该待推荐媒体对象相匹配的目标业务对象，并将该待推荐媒体对象推送给当前查找出的目标业务对象。需要注意的是，推荐模型13e是通过上述N个历史媒体对象的N个历史交互行为信息13f对初始推荐模型进行训练得到的。

[0171] 进一步的，请参见图14，图14是本申请实施例提供的一种针对直播间的目标业务对象预估流程示意图。如图14所示，以待推荐媒体对象为待推荐直播类媒体对象(例如，待加热直播间)为例，可以将该待加热直播间统称为当前直播间，此时，历史媒体对象可以为历史直播类媒体对象(例如，图14所示的历史直播间)。该针对媒体对象的推荐方法至少可以包括以下步骤S301至步骤S309：

[0172] 步骤S301：时间窗采样。

[0173] 其中，具体地，针对待推荐直播类媒体对象的多媒体数据为待推荐直播数据而言，服务器可以首先使用目标采样窗口对上述待推荐直播数据进行采样(即进行时间窗采样)，以得到目标音频数据。其中，服务器采样得到目标音频数据的具体过程，可以参见上述图3所对应实施例中对采样得到目标音频数据的具体过程的描述，这里将不再继续进行赘述。在执行完步骤S301后，跳转执行步骤S303。

[0174] 步骤S302：高交互过滤。

[0175] 同理，针对历史媒体对象的多媒体数据为历史直播数据而言，服务器可以基于历史媒体对象数据库所对应的交互数据曲线分布，筛选出交互数据量处于前10%的历史媒体对象(即筛选出经高交互过滤而得到的高交互直播间)。换言之，本申请实施例可以在对历史媒体对象数据库中的所有历史媒体对象的交互数据量进行数据量统计的情况下，生成该交互数据曲线分布，进而可以从该交互数据曲线分布中过滤出位于头部的历史直播间统称为具有较高交互数据量的高交互直播间。进而，服务器可以获取筛选出的N个历史直播类媒体对象的N个历史音频数据。应当理解，历史音频数据为该服务器通过历史采样窗口(例如，记录到的在历史动态采样时间点为历史直播类媒体对象的历史直播数据的结束时间点时的采样窗口)所采集到的采样时长为窗口时长的音频数据。需要注意的是，上述交互数据曲线分布为以历史媒体对象数据库中的各个历史媒体对象和各个历史媒体对象的交互数据量为两坐标轴所确定的分布曲线图。具体可以参见上述图3所对应实施例中对筛选历史媒体对象的具体过程的描述，这里将不再继续进行赘述。

[0176] 步骤S303：语音识别处理。

[0177] 其中，具体地，本申请实施例可以针对上述待推荐直播类媒体对象的目标音频数据和上述N个历史直播类媒体对象的N个历史音频数据分别进行ASR处理(即进行语音识别处理)，以得到目标语音识别文本和N个历史语音识别文本。其中，服务器对每个历史媒体对象的历史音频数据进行语音识别处理的具体过程，可以参见上述图3所对应实施例中对目标音频数据进行语音识别处理的具体过程的描述，这里将不再继续进行赘述。

[0178] 步骤S304：语义理解信息提取。

[0179] 其中，具体地，服务器可以分别基于引导提示模板数据，使得语义理解模型(例如大语言模型)能够对目标语音识别文本和N个历史语音识别文本分别进行语义理解信息提取(即进行语义信息提取)，以提取得到目标语义理解信息和N个历史语义理解信息。应当理解，该目标语义理解信息是由对目标语音识别文本进行语义信息提取所得到的目标语义提取信息所确定的，同理，历史语义理解信息是由对历史语音识别文本进行语义信息提取所得到的历史语义提取信息所确定的。其中，服务器对每个历史语音识别文本进行语义信息提取的具体过程，可以参见上述图3所对应实施例中对目标语音识别文本进行语义信息提取的具体过程的描述，这里将不再继续进行赘述。

[0180] 步骤S305：文本向量化表征。

[0181] 其中，具体地，本申请实施例可以分别对上述目标语义理解信息和N个历史语义理解信息进行文本向量表征处理，以得到该目标语义理解信息对应的目标文本表征向量和每个历史语义理解信息对应的历史文本表征向量。其中，服务器对每个历史语义理解信息进行文本向量表征处理的具体过程，可以参见上述图3所对应实施例中对目标语义理解信息进行文本向量表征处理的具体过程的描述，这里将不再继续进行赘述。

[0182] 步骤S306：当前直播间表征。

[0183] 其中，具体地，本申请实施例可以在对目标语义理解信息进行文本向量表征处理之后，将处理得到的文本表征向量统称为当前直播间表征，即当前直播间表征可以用于表征当前待推荐直播类媒体对象(即当前直播间)的目标文本表征向量。

[0184] 步骤S307：高交互直播间表征。

[0185] 同理，具体地，本申请实施例也可以在经步骤S305对N个历史语义理解信息分别进行文本向量表征处理之后，将处理得到的每个文本表征向量统称为高交互直播间表征，即该高交互直播间表征用于表征经高交互过滤后所得到的具有高交互数据量的每个历史直播间的历史文本表征向量。进一步地，服务器可以将上述目标文本表征向量与上述N个历史文本表征向量分别进行向量相似度计算，并通过计算得到的向量相似度从这N个历史文本表征向量中筛选得到相似度最高的K个历史文本表征向量。

[0186] 步骤S308：得到直播间序列。

[0187] 其中，具体地，服务器将上述筛选出的K个历史文本表征向量所对应的K个历史直播类媒体对象分别与上述待推荐直播类媒体对象建立相似关联关系，以得到由这K个历史直播类媒体对象所构成的相似媒体对象列表(这里的相似媒体对象列表可以为图14所示的直播间序列)。

[0188] 步骤S309：推荐模型。

[0189] 其中，具体地，本申请实施例可以获取该相似媒体对象列表中K个历史媒体对象的K个历史推荐特征，进而对其进行池化处理，以得到用于表征待推荐直播类媒体对象的目标推荐特征。进而，服务器可以调用图14所示的推荐模型，基于上述目标推荐特征，从访问该业务处理平台(例如，社交软件平台)的所有在线用户(即所有业务对象)中，查找出与上述待推荐直播类媒体对象相匹配的目标业务对象，以向目标业务对象加热推荐该待推荐直播类媒体对象(即图14所示的当前直播间)。

[0190] 本申请实施例在通过目标采样窗口采集得到目标音频数据时，可以通过与该目标音频数据相关联的目标语音识别文本，计算得到用于表征该待推荐媒体对象的目标文本表征向量，这样，可以通过向量相似计算与检索的方式将目标文本表征向量与历史媒体对象的历史文本表征特征进行相似关联，从而可以通过筛选出的与上述待推荐媒体对象相关联的K个历史媒体对象的历史推荐特征，来等效计算得到待推荐媒体对象的目标推荐特征，这样，在待推荐媒体对象的交互数据量较小的情况下，仍可以使用语义高度相近的这K个历史媒体对象的历史推荐特征，来等效描述该待推荐媒体对象的目标推荐特征，这样，服务器所涉及的推荐系统在使用目标推荐模型进行媒体推荐时，可以直接通过当前等效计算出的该目标推荐特征向与待推荐媒体对象相匹配的目标业务对象推荐该待推荐媒体对象，从而可以提升媒体推荐的精准度。

[0191] 为便于理解，如图15所示，图15是本申请实施例提供的一种针对媒体对象的推荐方法的流程示意图。在一个实施例中，该针对媒体对象的推荐方法至少可以包括以下步骤S401至步骤S407：

[0192] 步骤S401：获取待推荐媒体对象的目标动态采样时间点和窗口时长，基于目标动态采样时间点和窗口时长，确定用于对待推荐媒体对象的多媒体数据进行采样的目标采样窗口，进而通过目标采样窗口从待推荐媒体对象的多媒体数据中，获取数据采样结束时间点为目标动态采样时间点，且采样时长为窗口时长的音频数据，将获取到的音频数据作为待推荐媒体对象在目标采样窗口中的目标音频数据。

[0193] 其中，目标动态采样时间点是指随着待处理媒体对象(即需要加热推荐的待推荐媒体对象)的媒体数据关联时长的变化而变化的时间点；目标动态采样时间点包括第一动态采样时间点，目标采样窗口包括第一动态采样时间点所对应的第一采样窗口，目标音频数据包括基于第一动态采样时间点和窗口时长所采集得到的第一音频数据。

[0194] 为便于理解，当待推荐媒体对象为直播类媒体对象，待推荐媒体对象的多媒体数据为待推荐直播数据时，历史媒体对象可以为历史直播类媒体对象，且历史媒体对象的多媒体数据可以为历史直播数据。此时，服务器可以在接收到第一本地设备发送的待推荐媒体对象的多媒体数据时，首先获取目标动态采样时

间点、窗口时长。可以理解的是，直接获取到的目标动态采样时间点和窗口时长可以为服务器所统一设定的默认值。

[0195] 进一步地，本申请实施例可以将用于对上述待推荐直播数据进行采样的采样窗口，作为目标采样窗口。进而，将上述目标动态采样时间点作为数据采样的结束时间点，获取通过目标采样窗口所采集到的采样时长为窗口时长的音频数据，并将在上述目标采样窗口中获取到的音频数据作为目标音频数据。

[0196] 应当注意的是，第一个作为目标动态采样时间点的时间点为从待推荐直播数据的开始时长起、至采样时长为上述窗口时长时所确定的第一动态采样时间点，且采样时长为窗口时长时的第一采样窗口为目标采样窗口，该目标采样窗口用于对上述待推荐直播数据进行截取和采集。

[0197] 步骤S402：获取用于进行动态采样的采样间隔时长，基于第一动态采样时间点和采样间隔时长，确定用于对待推荐媒体对象的多媒体数据进行采样的第二动态采样时间点，进而基于第二动态采样时间点和窗口时长，确定用于对待推荐媒体对象的多媒体数据进行采样的第二采样窗口，再通过第二采样窗口从待推荐媒体对象的多媒体数据中，获取数据采样结束时间点为第二动态采样时间点，且采样时长为窗口时长的音频数据，将获取到的音频数据作为待推荐媒体对象在目标采样窗口中的第二音频数据，在通过第二动态采样时间点更新目标动态采样时间点时，通过第二音频数据更新目标音频数据。

[0198] 其中，第二动态采样时间点为第一动态采样时间点的下一采样时间点；第二采样窗口为第一采样窗口的下一采样窗口，且第二采样窗口与第一采样窗口存在窗口重叠。

[0199] 其中，具体地，服务器所获取到的用于进行动态采样的采样间隔时长是指由两个相邻的分别作为目标动态采样时间点的时间点所确定的间隔时长。换言之，本申请实施例也可以基于上述第一动态采样时间点和采样间隔时长确定第二动态采样时间点，也就是说，第二动态采样时间点与第一动态采样时间点之间的间隔时长为该采样间隔时长。进一步地，本申请实施例可以将用于对上述待推荐直播数据进行采样的另一采样窗口(即另一目标采样窗口)统称为第二采样窗口，进而，可以通过第二采样窗口(即以第二动态采样时间点为数据采样结束时间点，采样时长为上述窗口时长的第二采样窗口)更新目标采样窗口，以实现对上述待推荐直播数据的数据采集，进而可以通过第二采样窗口采集到的直播数据中的音频数据，更新上述目标音频数据，以达到动态采样的效果。

[0200] 需要注意的是，上述采样间隔时长用于指示任意两个相邻的目标采样窗口之间存在窗口重叠，为便于理解，也就是说在上述采样间隔时长应小于或者等于目标采样窗口的窗口时长时，可以使上述待推荐直播数据能够被完整采集。

[0201] 为便于理解，如图16所示，图16是本申请实施例提供的一种基于待处理直播数据播放时间轴的动态采样示意图。应当理解，这里的待处理直播数据播放时间轴可以用于反映当前直播间的直播进度。如图16所示，目标采样窗口具体是指在第一采样窗口16b1和第二采样窗口16b2中所最新确定得到的采样窗口，目标动态采样时间点具体可以包括第一动态采样时间点16a1和第二动态采样时间点16a2。比如，如图16所示，当上述待推荐直播数据的直播进度进行到T2时刻时，可以将T2时刻作为第一动态采样时间点16a1，随即可以将T2时刻所对应的采样窗口作为第一采样窗口16b1，即可以通过该第一采样窗口16b1采样得到采样结束的时间点为T2时刻时的时间点、且采样时长为窗口时长16d的待推荐直播数据，也就是可以通过目标采样窗口16B截取以目标动态采样时间点16A为止(即采样结束的时间点)、且采样时长为窗口时长16d的多媒体数据。进而，提取上述目标采样窗口16B中采集到的多媒体数据的音频数据，并将其作为目标音频数据。进一步地，随着待推荐直播数据的继续进行，本申请实施例可以动态判断该待推荐直播数据的当前直播进度与上述T2时刻所表征的直播进度之间的间隔时长是否达到采样间隔时长16c，如果达到该采样间隔时长16c，则可以通过该T2时刻和该采样间隔时长16c，确定该待推荐直播数据的直播进度当前已进行到T3时刻，从而可以将T3时刻作为第二动态采样时间点16a2，并将T3时刻所对应的采样窗口作为第二采样窗口16b2，然后，可以通过该第二采样窗口16b2采样得到采样结束的时间点为T3时刻时的时间点、且采样时长为窗口时长16d的另一待推荐直播数据，进而可以将第二动态采样时间点16a2作为新的目标动态采样时间

点16A，以替换掉以第一动态采样时间点16a1为目标动态采样时间点16A的采样结束时间点，从而可以以第二采样窗口16b2替换掉第一采样窗口16b1作为新的目标采样窗口的方式，达到动态更新目标采样窗口16B的目的，进而将从第二采样窗口16b2中待推荐直播数据提取到的音频数据作为另一目标音频数据，以执行下述步骤S403−步骤407。可选地，在一种可实现方式中，为避免存储资源的浪费，本申请实施例还在通过第二采样窗口16b2采集到新的音频数据(即另一目标音频数据)时，通过该新的音频数据替换原来通过第一采样窗口16b1所采集到的旧的音频数据，即可以将该新的音频数据直接作为当前最新的目标音频数据。

[0202] 步骤S403：对目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量。

[0203] 其中，获取用于对目标音频数据进行语音转换处理的语音转换理解组件和相似检索组件。其中，语音转换理解组件包括语音识别器和用于进行语义提取的语义理解模型且相似检索组件包括文本表征器。

[0204] 进一步地，通过上述语音转换理解组件的语音识别器对上述目标音频数据进行语音识别处理，得到与上述目标音频数据相关联的目标语音识别文本。进而，通过上述语音转换理解组件中的语义理解模型对目标语音识别文本进行信息提取，从而得到目标语义理解信息。进一步地，通过上述相似检索组件的文本表征其对上述目标语义理解信息进行文本向量表征处理，得到用于表征待推荐直播类媒体对象的目标文本表征向量。

[0205] 步骤S404：获取用于表征N个历史媒体对象的N个历史文本表征向量，对目标文本表征向量和每个历史文本表征向量分别进行向量相似计算，得到N个向量相似度。

[0206] 其中，N为大于1的正整数；一个历史文本表征向量是通过对一个历史音频数据对应的历史语音识别文本进行文本向量表征处理后所记录得到的，一个历史音频数据是指一个历史媒体对象在记录到的历史采样窗口中的音频数据；历史采样窗口的窗口时长与目标采样窗口的窗口时长保持一致。

[0207] 其中，具体地，上述目标文本表征向量是通过语音转换处理的语音转换理解组件和相似检索组件对上述目标音频数据进行语音转换处理后得到的。进而，获取用于表征N个历史媒体对象的N个历史文本表征向量，且在获取到上述N个历史文本表征向量后，将上述用于表征待推荐直播类媒体对象的目标文本表征向量分别与用于表征N个历史直播类媒体对象的N个历史文本表征向量进行向量相似度计算。其中，向量相似度的计算方法包括余弦相似性、杰卡德系数相似性、欧几里得距离以及曼哈顿距离等，这里不做限定。

[0208] 步骤S405：在通过N个向量相似度从N个历史媒体对象中筛选出K个历史媒体对象时，建立K个历史媒体对象与待推荐媒体对象之间的相似关联关系。

[0209] 其中，具体地，基于上述计算出的N个向量相似度对N个历史直播类媒体对象进行排序处理，进而从中筛选出与上述待推荐直播类媒体对象相关联的K个历史直播类媒体对象，进而将上述K个历史直播类媒体对象与上述待推荐直播类媒体对象建立相似关联关系。具体实现方式可以参见上述图3所对应实施例中对筛选K个历史媒体对象的具体过程的描述，这里将不再继续进行赘述。

[0210] 步骤S406：从与K个历史媒体对象相关联的推荐特征数据库中获取K个历史媒体对象的K个历史推荐特征。

[0211] 其中，具体地，上述K个历史直播类媒体对象的K个历史推荐特征为服务器从与K个历史直播类媒体对象相关联的推荐特征数据库中获取到的。

[0212] 其中，上述推荐特征数据库是通过以下方法得到：

[0213] 获取N个历史媒体对象对应的N个历史交互行为信息。其中，一个历史媒体对象对应一个历史交互行为信息，一个历史交互行为信息用于表征一个历史媒体对象中的各个交互物品的物品交互详情。将N个历史交互行为信息所表征的物品交互详情作为用于训练初始推荐模型的训练样本，获取用于训练初始推荐

模型的推荐指标。通过训练样本和推荐指标，对初始推荐模型进行模型训练，得到目标推荐模型；目标推荐模型用于提取得到N个历史媒体对象的N个历史推荐特征。其中，一个历史推荐特征为目标推荐模型对一个历史交互行为信息所表征的物品交互详情进行交互向量提取所得到的。将目标推荐模型所提取到的N个历史推荐特征，添加至推荐特征数据库。

[0214] 其中，具体地，服务器获取上述N个历史直播类媒体对象中每个历史直播类媒体对象所对应的历史交互行为信息。进而，将上述获取到的N个历史直播类媒体对象的历史交互行为信息表征的用户物品交互的详情作为初始推荐模型的训练样本。上述初始推荐模型是还未训练的用于给长尾或冷启动媒体对象进行加热推荐的模型。且在对上述初始推荐模型进行训练时，获取用于训练的与上述初始推荐模型相关联的推荐指标。

[0215] 进一步地，通过上述初始推荐模型的推荐指标，对上述初始推荐模型进行训练，得到用于提取待推荐媒体对象推荐特征与历史媒体对象推荐特征的目标推荐模型。通过上述目标推荐模型对上述N个历史直播类媒体对象进行处理，得到上述N个历史直播类媒体对象的N个历史推荐特征，将上述N个历史推荐特征添加至推荐特征数据库。

[0216] 需要注意的是，待推荐媒体对象(例如，上述待推荐直播类媒体对象)为业务处理平台中的交互数据量小于交互数据量阈值的多媒体对象。也就是说，待推荐媒体对象(例如，上述待推荐直播类媒体对象)为冷启动/长尾媒体对象。

[0217] 步骤S407：在获取到所述K个历史媒体对象的K个历史推荐特征时，基于所述相似关联关系对所述K个历史推荐特征进行特征池化处理，得到与所述K个历史推荐特征相关联的池化处理特征，将所述池化处理特征，确定为用于表征所述待推荐媒体对象的目标推荐特征，通过目标推荐特征对待推荐媒体对象进行媒体推荐。

[0218] 应当理解，在上述将K个历史直播类媒体对象的K个历史推荐特征进行池化处理之前，可以先根据相似关联关系获取用于进行池化处理的池化处理策略。该池化处理策略包括平均池化处理策略、最大池化处理策略以及随机池化处理策略等，此处不做限定。

[0219] 其中，若上述池化处理时所采用的池化处理策略为平均池化处理策略时，可以由推荐模型按照该平均池化处理策略对K个历史直播数据的K个历史推荐特征进行平均池化处理，得到平均池化特征。进而，基于上述平均池化特征得到与上述待推荐直播数据的目标推荐特征。

[0220] 应当理解，推荐模型对K个历史推荐特征进行池化处理，得到待推荐媒体对象的目标推荐特征的具体实现方式，参见如下公式(9)：

$$\hat{r}_i = \frac{1}{K}\sum_{k=1}^{K} M(J_k) \qquad (9)$$

[0221] 其中，符号J的下标k表示在为当前直播间(例如，直播间i)所确定的语义临近列表(即相似媒体对象列表)中的第k个历史高交互媒体对象的ID索引，k为小于或者等于K的正整数，即k的取值为1～K；K表示通过向量相似度检索方式所检索出的K个最相似的直播间；M表示在待推荐媒体对象推荐场景下部署的当前已训练好的推荐模型(即上述目标推荐模型)；$M(J_k)$表示通过该目标推荐模型所提取到的第k个历史媒体对象的推荐表征。

[0222] 进一步地，在将目标推荐模型部署应用在服务器所提供的业务处理平台时，可以由服务器获取访问业务处理平台的业务对象的对象关联数据信息(例如，由服务器获取各个业务对象通过第二本地设备所授权发送的历史行为交互信息，在推荐业务为直播推荐业务时，这里的历史行为交互信息是基于用户在一个或者多个直播间中针对某些物品执行的交互操作所确定的，例如，这里的交互操作包括点赞操作、收藏操作以及物品交易操作等任一项操作)，将对象关联数据信息和目标推荐特征输入目标推荐模型，通过目标推荐模型在从对象关联数据信息中提取到用于表征业务对象的对象向量表征特征(这里的对象向量表征特征可以用于反映业务对象对某些直播间中的物品的偏好情况)时，在对象向量表征特征中筛选与目标推荐

特征相匹配的目标对象向量表征特征。这样，服务器可以将目标对象向量表征特征所对应的业务对象，确定为在业务对象中所筛选出的与待推荐媒体对象相匹配的目标业务对象，并向目标业务对象推荐待推荐媒体对象。

[0223] 其中，具体地，上述业务处理平台包括社交软件平台，在上述业务处理平台为社交软件平台的情况时，将上述训练好的目标推荐模型部署在上述社交软件平台中。且在部署好上述目标推荐模型的同时，筛选出的目标业务对象可以通过第二本地设备(例如，图1所对应实施例中的用户终端20a)接收由服务器(例如，上述图1所对应实施例中的服务器10d)所加热推荐的待推荐媒体对象，并可以在第二本地设备的社交软件平台上查看到该待推荐媒体对象的媒体数据(例如，上述待推荐直播数据)。比如，服务器可以获取当前访问该社交软件平台中的所有用户(即上述业务对象)的历史行为交互信息，这里的历史行为交互信息可以用于记录这些用户的物品交易信息及其用户偏好(即上述对象关联数据信息)。进一步地，服务器将目标推荐特征与获取到的这些用户的物品交易信息及其用户偏好输入目标推荐模型中，以使目标推荐模型在上述物品交易信息及其用户偏好中提取用于表征该用户的对象向量表征特征。进而，在社交软件平台的所有用户的对象向量表征特征中筛选出与上述目标推荐特征相匹配的对象向量表征特征，且将其作为目标对象向量表征特征。进一步地，将上述目标对象向量表征特征所对应的用户作为与上述待推荐直播数据相匹配的目标业务对象，进而向上述目标业务对象推荐上述待推荐媒体对象。

[0224] 其中，上述目标推荐模型查找与待推荐媒体对象相匹配的目标业务对象的具体实现方式如下：

$$y_{u,i} = similarity\left(\hat{r}_i, M(u)\right) \qquad (10)$$

[0225] 其中，M(u)表示某个用户(例如，在线用户u)在推荐模型M中的表征(即对象向量表征特征)。最后，通过公式(10)所指示的特征匹配的方式，计算目标推荐特征$\hat{r}_i$与该对象向量表征特征之间的特征匹配度，来评估该在线用户u对冷启动/长尾直播间的偏好分数$y_{u,i}$，以圈定满足需求的用户量级，进而实现个性化加热推荐的推荐准确度。

[0226] 应当理解，将对象关联数据信息和目标推荐特征输入目标推荐模型，并通过目标推荐模型对对象关联数据信息进行信息向量表征提取，得到用于表征业务对象的对象向量表征特征。将对象向量表征特征和目标推荐特征进行特征匹配，得到特征匹配结果。若特征匹配结果指示对象向量表征特征中存在与目标推荐特征相匹配的对象向量表征特征，则将与目标推荐特征相匹配的对象向量表征特征，确定为目标对象向量表征特征。

[0227] 其中，具体地，在将上述社交软件平台(即上述业务处理平台)用户(即上述业务对象)的物品交易信息及其用户偏好(即上述对象关联数据信息)和目标推荐特征输入部署在社交软件平台的目标推荐模型时，该目标推荐模型对上述用户的物品交易信息及其用户偏好进行信息向量表征提取，从而提取得到用于表征用户信息的对象向量表征特征。进一步地，从上述所有用户中筛选出与待推荐媒体对象相关联的用户，进而目标推荐模型将上述对象向量表征特征与待推荐媒体对象的目标推荐特征进行匹配，进而得到特征匹配结果。相应地，若特征匹配结果指示对象向量表征特征中存在与目标推荐特征相匹配的向量表征特征，则将该对象向量表征特征作为目标对象向量表征特征。

[0228] 进一步地，若待推荐媒体对象为直播类媒体对象，则将待推荐直播类媒体对象作为在直播推荐业务中的用于推荐给目标业务对象的目标推荐对象，将目标推荐对象发送给目标业务对象所对应的终端设备，以使终端设备在直播推荐业务所对应的直播推荐列表中显示目标推荐对象。可选地，若待推荐媒体对象为点播类媒体对象，则将点播类媒体对象所对应的点播链接封装为推荐通知消息，将推荐通知消息发送给目标业务对象所对应的终端设备，以使终端设备在直播推荐业务所对应的直播推荐列表上显示推荐通知消息中的点播链接；点播链接用于指示终端设备响应于针对推荐通知消息的触发操作，显示点播类媒体对象。

[0229] 需要注意的是，为了综合评估本申请实施例提出方法的有效性，当待推荐媒体对象为直播类媒体对象时，在业务处理平台的离线和在线两个场景进行充分的方案评估。本次评估使用的ASR算法能够用于进行语音识别，且使用的大语言模型可以用于进行语义理解，且使用向量相似度算法(例如，余弦相似度)所

筛选出的高度语义相似的历史直播类媒体对象的数量为K，这样当K＝20时，也就是筛选出20个与待推荐直播类媒体对象相关联的历史直播类媒体对象。

[0230] 首先，离线评估时主要关注目标推荐模型的受试者工作特征曲线下方的面积大小(Area Under Curve，AUC)、加权平均后各组受试者工作特征曲线下方与坐标轴围成的面积(Group AUC，GAUC)和精度(Precision at Cut-off，PCOC)三个指标。

[0231] 其中，受试者工作特征曲线(Receiver Operating Characteristic curve，ROC)曲线下的面积中，ROC曲线是以假阳性率(false positive rate，FPR)为横轴，真正率(true positive rate，TPR)为纵轴绘制的曲线。AUC值的范围在0到1之间，AUC值越高，表示模型在区分正负样本方面的能力越强，比如，在本申请实施例中，能够通过ROC曲线智能确定出用于进行二分类(例如，具有较高交互数据量的头部直播间或者具有较低交互数据量的尾部直播间)的最佳划分阈值，进而可以将确定出的最佳划分阈值作为推荐系统所配置的上述交互数据量阈值，换言之，本申请实施例可以通过ROC曲线下的面积大小来计算得到AUC值，进而可以通过AUC值反映通过不同业务对象的训练样本(例如，用户的历史行为交互信息)离线训练得到目标推荐模型时，由该目标推荐模型对正样本(例如，冷启动和/或长尾直播间等低交互直播间或者低热度直播间)和负样本(例如，高交互直播间或者高热度直播间)进行点击预测的能力，具体可以参加下述表1。这里的正负样本可以根据实际训练需求而确定，例如，正样本可以是高热度直播间，负样本可以是冷启动直播间等，这里将不对参与训练得到目标推荐模型的正负样本的具体类型进行限定。

[0232] 其中，GAUC为对AUC的扩展，通常用于评估推荐系统在不同业务对象组或不同类别上的表现。GAUC是对每个组的AUC进行加权平均，考虑了不同业务对象组的样本量。GAUC越高越好，表示模型在各个组别上的表现都较为均衡，比如，如下表1所示，点击PV(即点击浏览量或者交互数据量)为0-100的直播间可以划分为一个组，点击PV为100-500的直播间可以划分为另一个组。

[0233] 其中，PCOC指在推荐列表中，前X个已推荐待推荐媒体对象的精确率。它衡量的是在推荐的前X个媒体对象中，有多少比例是业务对象实际感兴趣的媒体对象。PCOC越高越好，表示推荐系统在前X个已推荐待推荐媒体对象中的准确性越高。

[0234] 下述表格将基于点击PV中的页面浏览量(page view，PV)对本申请实施例提出的方法进行评估。在本申请实施例中，显示不同热度直播间在使用本申请实施例提出的方法后，相对基线取得的相对增益可以参见下述表1。应当理解，上述直播间为待推荐直播类媒体对象。

[0235] 表1

点击 PV	0-100	100-500	500-2000	2000-5000	5000+
AUC	+0.049%	+0.047%	+0.023%	+0.015%	+0.014%
GAUC	+0.149%	+0.079%	+0.034%	+0.019%	+0.016%
PCOC	+0.189%	+0.112%	+0.064%	+0.043%	+0.042%

[0236] 如上述表1所示，本申请实施例在不同热度直播间实施时，相对基线(即相对于采用现有技术的推荐模型来评估这些指标时所预测的点击浏览量)都取得了不同的增益。

[0237] 进一步地，对本申请实施例在进行在线评估时，主要关注冷启渠道、长尾渠道直播间(即上述待推荐直播类媒体对象)的加热效果。在本申请实施例中，显示冷启动渠道分发相对基线取得的相对收益可以参见下述表2。

[0238] 表2

冷启动分发	观看 PV	1min 观看 PV	1min 观看 PV 占比
近邻检索=10	+0.866%	+0.629%	+0.825%
近邻检索=20	+1.524%	+1.173%	+1.248%

[0239] 在本申请实施例中，显示长尾内容分发相对基线取得的相对收益可以参见下述表3。

[0240] 表3

长尾内容分发	观看 PV	1min 观看 PV	1min 观看 PV 占比
PV<500 观看	+2.698%	+ 6.153%	+3.287%
PV<100 观看	+4.034%	+6.721%	+2.658%

[0241] 如上述表2、表3所示，本申请实施例在冷启动渠道、长尾内容分发相对基线都取得了不同的增益。

[0242] 本申请实施例通过目标采样窗口从待推荐媒体对象的多媒体数据中采集的以数据采样结束时间点为所述目标动态采样时间点，且采样时长为所述窗口时长的音频数据，将其作为目标音频数据。且上述目标音频数据基于采样间隔时长更新目标动态采样时间点，进而更新与待推荐媒体对象相关联的目标音频数据。进一步地，可以识别在目标采样窗口中获取到的目标音频数据，得到与上述目标音频数据相关联的目标语音识别文本。进一步地，本申请实施例对与音频数据相关联的目标语音识别文本进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量。获取用于表征N个历史媒体对象的N个历史文本表征向量，进而对上述目标文本表征向量与N个历史文本表征向量中的每个历史文本表征向量分别进行向量相似计算，得到N个向量相似度。进一步地，本申请实施例基于上述N个向量相似度，从上述N个历史媒体对象中筛选出K个与待推荐媒体对象相似度高的历史媒体对象，进而，建立待推荐媒体对象与上述筛选出的K个历史媒体对象的相似关联关系。本申请实施例从与K个历史媒体对象相关联的推荐特征数据库中获取到上述K个历史媒体对象的K个历史推荐特征，进而获取用于进行池化处理的池化处理策略将上述K个历史推荐特征进行池化处理，得到上述待推荐媒体对象的目标推荐特征。进一步地，获取上述N个历史媒体对象的历史交互行为信息，从而在将所述目标推荐模型部署应用在业务处理平台时，获取访问业务处理平台的业务对象的对象关联数据信息，进而通过上述对象关联数据信息和上述目标推荐特征从业务处理平台的业务对象中筛选出与上述待推荐媒体对象相匹配的目标业务对象，进而通过向上述目标业务对象进行媒体推荐。本申请实施例可以对待推荐媒体对象的多媒体数据进行动态采样，得到与待推荐媒体对象相关联的目标音频数据，且将上述目标音频数据中的内容向量化，并与历史音频数据中内容的向量进行相似度计算，从而检索到与待推荐媒体对象关联度高的K个历史媒体对象，从而从与K个历史媒体对象相关联的推荐特征数据库中获取K个历史媒体对象的历史推荐特征，进而对其进行池化处理，得到用于表征待推荐媒体对象的目标推荐特征，再通过N个历史媒体对象的N个历史交互行为信息和待推荐媒体对象的目标推荐特征筛选出与待推荐媒体对象相匹配的目标业务对象，对其推送上述待推荐媒体数据。进而，提高了目标推荐模型对冷启动待推荐媒体对象和业务对象匹配的准确度，还达到加热待推荐媒体对象的效果。

[0243] 进一步地，请参见图17，图17是本申请实施例提供的一种针对媒体对象的推荐装置的结构示意图。该针对媒体对象的推荐装置可以是运行于计算机设备中的一个计算机程序(包括程序代码)，例如该针对媒体对象的推荐装置1可以为一个应用软件；该针对媒体对象的推荐装置1可以用于执行本申请实施例提供的方法中的相应步骤。其中，该针对媒体对象的推荐装置1可以包括：音频数据语音识别模块11、向量相似度计算模块12、相似关联关系建立模块13、目标推荐特征确定模块14。

[0244] 音频数据语音识别模块11，用于获取待推荐媒体对象在目标采样窗口中的目标音频数据，对目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征待推荐媒体对象的目标文本表征向量；

[0245] 向量相似度计算模块12，用于获取用于表征N个历史媒体对象的N个历史文本表征向量，对目标文本表征向量和N个历史文本表征向量分别进行向量相似计算，得到N个向量相似度；每个历史媒体对象的交互数据量均高于待推荐媒体对象的交互数据量；N为大于1的正整数；N个历史文本表征向量中的每个历史文本表征向量是通过对每个历史音频数据对应的历史语音识别文本进行文本向量表征处理后所记录得到的，每个历史音频数据是指每个历史媒体对象在记录到的历史采样窗口中的音频数据；历史采样窗口的窗口时长与目标采样窗口的窗口时长保持一致；

[0246] 相似关联关系建立模块13，用于在通过N个向量相似度从N个历史媒体对象中筛选出K个历史媒体对象时，建立K个历史媒体对象与待推荐媒体对象之间的相似关联关系；K为小于N的正整数；

[0247] 目标推荐特征确定模块14，用于在获取到K个历史媒体对象的K个历史推荐特征时，基于相似关联关系和K个历史推荐特征，确定用于表征待推荐媒体对象的目标推荐特征，通过目标推荐特征对待推荐媒体对象进行媒体推荐。

[0248] 其中，音频数据语音识别模块11、向量相似度计算模块12、相似关联关系建立模块13、目标推荐特征确定模块14的具体实现方式，可以参见上述图3所对应实施例中对步骤S101-步骤S104与图15所对应实施例中对步骤S401-步骤S407的描述，这里将不再继续进行赘述。

[0249] 其中，音频数据语音识别模块11包括：目标音频数据确定单元110、目标音频数据更新单元111、组件获取单元112、语音识别单元113、语义理解单元114、文本向量表征单元115、无效音频数据标记单元116、目标语音识别文本更新单元117。

[0250] 目标音频数据确定单元110，用于获取待推荐媒体对象的目标动态采样时间点和窗口时长，基于目标动态采样时间点和窗口时长，确定用于对待推荐媒体对象的多媒体数据进行采样的目标采样窗口；目标动态采样时间点是指随着待推荐媒体对象的媒体数据关联时长的变化而变化的时间点；

[0251] 目标音频数据确定单元110，还用于通过目标采样窗口从待推荐媒体对象的多媒体数据中，获取数据采样结束时间点为目标动态采样时间点，且采样时长为窗口时长的音频数据；

[0252] 目标音频数据确定单元110，还用于将获取到的音频数据作为待推荐媒体对象在目标采样窗口中的目标音频数据。

[0253] 其中，可选地，所述目标动态采样时间点包括第一动态采样时间点，所述目标采样窗口包括所述第一动态采样时间点所对应的第一采样窗口，所述目标音频数据包括基于所述第一动态采样时间点和所述窗口时长所采集得到的第一音频数据；

[0254] 目标音频数据更新单元111，用于获取用于进行动态采样的采样间隔时长，基于第一动态采样时间点和采样间隔时长，确定用于对待推荐媒体对象的多媒体数据进行采样的第二动态采样时间点；第二动态采样时间点为第一动态采样时间点的下一采样时间点；

[0255] 目标音频数据更新单元111，还用于基于第二动态采样时间点和窗口时长，确定用于对待推荐媒体对象的多媒体数据进行采样的第二采样窗口；第二采样窗口为第一采样窗口的下一采样窗口，且第二采样窗口与第一采样窗口存在窗口重叠；

[0256] 目标音频数据更新单元111，还用于通过第二采样窗口从待推荐媒体对象的多媒体数据中，获取数据采样结束时间点为第二动态采样时间点，且采样时长为窗口时长的音频数据；

[0257] 目标音频数据更新单元111，还用于将获取到的音频数据作为待推荐媒体对象在目标采样窗口中的第二音频数据，在通过第二动态采样时间点更新目标动态采样时间点时，通过第二音频数据更新目标音频数据。

[0258] 其中，具体的，组件获取单元112，用于获取用于对目标音频数据进行语音转换处理的语音转换理解组件和相似检索组件；语音转换理解组件包括语音识别器和用于进行语义提取的语义理解模型，相似检索组件包括文本表征器。

[0259] 语音识别单元113，用于通过语音识别器对目标音频数据进行语音识别处理，得到语音识别处理结果；

[0260] 语义理解单元114，用于若语音识别处理结果指示识别到目标音频数据的目标语音识别文本，则通过语义理解模型对目标语音识别文本进行文本语义理解处理，得到目标语音识别文本的目标语义理解信息；

[0261] 文本向量表征单元115，用于通过文本表征器对目标语义理解信息进行向量表征，得到用于表征待推荐媒体对象的目标文本表征向量。

[0262] 可选地，无效音频数据标记单元116，用于若上述语音识别处理结果指示未识别到目标音频数据的目标语音识别文本，则将目标音频数据标记为无效音频数据。

[0263] 可选地，目标语音识别文本更新单元117，用于通过目标采样窗口从待推荐媒体对象的多媒体数据中，获取待推荐媒体对象的视频数据，将获取到的视频数据作为待推荐媒体对象在目标采样窗口中的目标视频数据；

[0264] 目标语音识别文本更新单元117，还用于对目标视频数据中的视频帧进行视频内容识别，得到用于表征目标视频数据的视频文本内容，通过视频文本内容更新目标语音识别文本。

[0265] 其中，具体的，语义理解单元114，具体用于获取用于引导语义理解模型进行模型输出的引导提示模板数据，并将引导提示模版数据和目标语音识别文本输入语义理解模型，通过语义理解模型按照引导提示模版数据中的数据内容，对目标语音识别文本进行语义信息提取，得到目标语音识别文本的目标语义提取信息，并通过目标语义提取信息确定目标语音识别文本的目标语义理解信息。

[0266] 进一步的，语义理解单元114，具体用于若待推荐媒体对象的对象类型为视频直播类型，则将与视频直播类型相关联的提示模版数据作为引导提示模板数据；

[0267] 语义理解单元114，还具体用于若待推荐媒体对象的对象类型为视频点播类型，则将与视频点播类型相关联的提示模版数据作为引导提示模板数据。

[0268] 其中，上述目标音频数据确定单元110、目标音频数据更新单元111、组件获取单元112、语音识别单元113、语义理解单元114、文本向量表征单元115、无效音频数据标记单元116、目标语音识别文本更新单元117的具体实现方式可以参见上述图3所对应实施例中对步骤S101-步骤S104与图15所对应实施例中对步骤S401-步骤S407的描述，这里将不再继续进行赘述。

[0269] 进一步地，向量相似度计算模块12包括：交互数据曲线分布获取单元121、历史媒体对象确定单元122和历史文本表征向量获取单元123。

[0270] 交互数据曲线分布获取单元121，用于获取历史媒体对象数据库所对应的交互数据曲线分布；其中，交互数据曲线分布是基于历史媒体对象数据库中的各个历史媒体对象和各个历史媒体对象的交互数据量所确定的；

[0271] 历史媒体对象确定单元122，用于通过交互数据曲线分布，从历史媒体对象数据库中获取交互数据量达到交互数据阈值的N个历史媒体对象；

[0272] 历史文本表征向量获取单元123，用于从与N个历史媒体对象相关联的历史文本向量表征数据库中，获取N个历史媒体对象的N个历史文本表征向量；每个历史文本表征向量是通过语音转换理解组件和相似检索组件，对每个历史媒体对象的历史音频数据进行语音转换处理后所得到的。

[0273] 其中，上述交互数据曲线分布获取单元121、历史媒体对象确定单元122和历史文本表征向量获取单元123的具体实现方式可以参见上述图3所对应实施例中对步骤S101-步骤S104与图15所对应实施例中对步骤S401-步骤S407的描述，这里将不再继续进行赘述。

[0274] 其中，相似关联关系建立模块13包括：历史媒体对象列表确定单元131和相似关联关系建立单元132；

[0275] 历史媒体对象列表确定单元131，用于基于N个向量相似度对N个历史媒体对象进行排序处理，得到与N个历史媒体对象相关联的历史媒体对象列表；

[0276] 相似关联关系建立单元132，用于从历史媒体对象列表中获取向量相似度满足相似检索条件的K个历史媒体对象，建立K个历史媒体对象与待推荐媒体对象之间的相似关联关系；相似关联关系用于表征由K个历史媒体对象所构成的列表为相似媒体对象列表，且在相似媒体对象列表中的任意一个媒体对象均与待推荐媒体对象存在关联关系。

[0277] 其中，上述历史媒体对象列表确定单元131和相似关联关系建立单元132的具体实现方式可以参见上述图3所对应实施例中对步骤S101-步骤S104与图15所对应实施例中对步骤S401-步骤S407的描述，这里将不再继续进行赘述。

[0278] 其中，目标推荐特征确定模块14包括：目标推荐特征确定单元141和池化处理单元142。

[0279] 目标推荐特征确定单元141，用于在获取到K个历史媒体对象的K个历史推荐特征时，基于相似关联关系对K个历史推荐特征进行特征池化处理，得到与K个历史推荐特征相关联的池化处理特征；

[0280] 池化处理单元142，用于将池化处理特征，确定为用于表征待推荐媒体对象的目标推荐特征；

[0281] 其中，池化处理单元142，具体用于基于相似关联关系获取用于进行池化处理的池化处理策略；

[0282] 池化处理单元142，还具体用于在池化处理策略为平均池化处理策略时，对K个历史推荐特征进行平均池化处理，得到平均池化特征；平均池化特征为与K个历史推荐特征相关联的池化处理特征。

[0283] 其中，上述目标推荐特征确定单元141和池化处理单元142的具体实现方式可以参见上述图3所对应实施例中对步骤S101-步骤S104与图15所对应实施例中对步骤S401-步骤S407的描述，这里将不再继续进行赘述。

[0284] 其中，装置1还包括：历史推荐特征获取模块15；

[0285] 历史推荐特征获取模块15，用于从与K个历史媒体对象相关联的推荐特征数据库中获取K个历史媒体对象的K个历史推荐特征。

[0286] 其中，历史推荐特征获取模块15包括：历史交互行为信息获取单元151、推荐指标获取单元152、目标推荐模型训练单元153和推荐特征数据库更新单元154。

[0287] 历史交互行为信息获取单元151，用于获取N个历史媒体对象对应的N个历史交互行为信息；一个历史媒体对象对应一个历史交互行为信息，一个历史交互行为信息用于表征一个历史媒体对象中的各个交互物品的物品交互详情；

[0288] 推荐指标获取单元152，用于将N个历史交互行为信息所表征的物品交互详情作为用于训练初始推荐模型的训练样本，获取用于训练初始推荐模型的推荐指标；

[0289] 目标推荐模型训练单元153，用于通过训练样本和推荐指标，对初始推荐模型进行模型训练，得到目标推荐模型；目标推荐模型用于提取得到N个历史媒体对象的N个历史推荐特征，一个历史推荐特征为目标推荐模型对一个历史交互行为信息所表征的物品交互详情进行交互向量提取所得到的；

[0290] 推荐特征数据库更新单元154，用于将目标推荐模型所提取到的N个历史推荐特征，添加至推荐特征数据库。

[0291] 其中，上述历史交互行为信息获取单元151、推荐指标获取单元152、目标推荐模型训练单元153和推荐特征数据库更新单元154的具体实现方式可以参见上述图3所对应实施例中对步骤S101-步骤S104与图15所对应实施例中对步骤S401-步骤S407的描述，这里将不再继续进行赘述。

[0292] 其中，进一步地，待推荐媒体对象为业务处理平台中的交互数据量小于交互数据量阈值的多媒体对象，业务处理平台所对应的推荐模型为目标推荐模型，目标推荐模型是通过N个历史媒体对象对应的N个历史交互行为信息，对业务处理平台上所部署的初始推荐模型进行模型训练后所得到的；

[0293] 目标推荐特征确定模块14包括：关联数据信息确定单元143、对象向量特征确定单元144和媒体对象推荐单元145。

[0294] 关联数据信息确定单元143，用于在将目标推荐模型部署应用在业务处理平台时，获取访问业务处理平台的业务对象的对象关联数据信息；

[0295] 对象向量特征确定单元144，用于将对象关联数据信息和目标推荐特征输入目标推荐模型，通过目标推荐模型在从对象关联数据信息中提取到用于表征业务对象的对象向量表征特征时，在对象向量表征特征中筛选与目标推荐特征相匹配的目标对象向量表征特征；

[0296] 媒体对象推荐单元145，用于将目标对象向量表征特征所对应的业务对象，确定为在业务对象中所筛选出的与待推荐媒体对象相匹配的目标业务对象，并向目标业务对象推荐待推荐媒体对象。

[0297] 其中，具体的，对象向量特征确定单元144，具体用于将对象关联数据信息和目标推荐特征输入目标推荐模型，通过目标推荐模型，并通过目标推荐模型对对象关联数据信息进行信息向量表征提取，得到用于表征业务对象的对象向量表征特征；

[0298] 对象向量特征确定单元144，还具体用于将对象向量表征特征和目标推荐特征进行特征匹配，得到特征匹配结果；

[0299] 对象向量特征确定单元144，还具体用于若特征匹配结果指示对象向量表征特征中存在与目标推荐特征相匹配的对象向量表征特征，则将与目标推荐特征相匹配的对象向量表征特征，确定为目标对象向量表征特征。

[0300] 其中，具体的，媒体对象推荐单元145，具体用于若待推荐媒体对象为直播类媒体对象，则将直播类媒体对象作为在直播推荐业务中的用于推荐给目标业务对象的目标推荐对象，将目标推荐对象发送给目标业务对象所对应的终端设备，以使终端设备在直播推荐业务所对应的直播推荐列表中显示目标推荐对象；

[0301] 媒体对象推荐单元145，还具体用于若待推荐媒体对象为点播类媒体对象，则将点播类媒体对象所对应的点播链接封装为推荐通知消息，将推荐通知消息发送给目标业务对象所对应的终端设备，以使终端设备在直播推荐业务所对应的直播推荐列表上显示推荐通知消息中的点播链接；点播链接用于指示终端设备响应于针对推荐通知消息的触发操作，显示点播类媒体对象。

[0302] 其中，上述关联数据信息确定单元143、对象向量特征确定单元144和媒体对象推荐单元145的具体实现方式可以参见上述图3所对应实施例中对步骤S101-步骤S104与图15所对应实施例中对步骤S401-步骤S407的描述，这里将不再继续进行赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。

[0303] 可以理解的是，在本申请的具体实施方式中，可能涉及到用户、企业、机构等对象的业务数据(例如，用户的历史行为交互信息、对象关联数据信息、用户的物品交易信息、用户偏好等信息)，当本申请以上实施例运用到具体产品或技术中时，需要获得用户、企业、机构等对象的许可或同意，且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。

[0304] 进一步地，请参见图18，图18是本申请实施例提供的一种计算机设备的结构示意图。如图18所示，该计算机设备1000可以为用户终端，例如，上述图1所对应实施例中的用户终端10a、用户终端20a，还可以为服务器，例如，上述图1所对应实施例中的服务器10d，这里将不对其进行限制。为便于理解，本申请以计算机设备为用户终端为例，该计算机设备1000可以包括：处理器1001，网络接口1004和存储器1005，此外，该计算机设备1000还可以包括：用户接口1003，和至少一个通信总线1002。其中，通信总线1002用于实现这些组件之间的连接通信。其中，用户接口1003还可以包括标准的有线接口、无线接口。网络接口1004可选的可以包括标准的有线接口、无线接口(如WI-FI接口)。存储器1005可以是高速RAM存储器，也可以是非不稳定的存储器(non-volatile memory)，例如至少一个磁盘存储器。存储器1005可选的还可以是

至少一个位于远离前述处理器1001的存储装置。如图18所示，作为一种计算机可读存储介质的存储器1005中可以包括操作系统、网络通信模块、用户接口模块以及设备控制应用程序。

[0305] 其中，该计算机设备1000中的网络接口1004还可以提供网络通讯功能，且可选用户接口1003还可以包括显示屏(Display)、键盘(Keyboard)。在图18所示的计算机设备1000中，网络接口1004可提供网络通讯功能；而用户接口1003主要用于为用户提供输入的接口；而处理器1001可以用于调用存储器1005中存储的设备控制应用程序，以实现针对媒体对象在冷启动或者长尾情况下对其进行加热。

[0306] 应当理解，本申请实施例中所描述的计算机设备1000可执行前文图3或者图15所对应实施例中对针对媒体对象的推荐方法的描述，也可执行前文图17所对应实施例中对推荐装置1的描述，在此不再赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。

[0307] 此外，这里需要指出的是：本申请实施例还提供了一种计算机可读存储介质，且计算机可读存储介质中存储有前文提及的针对媒体对象的推荐装置1所执行的计算机程序，且计算机程序包括计算机指令，当处理器执行计算机指令时，能够执行前文图3或者图15所对应实施例中对针对媒体对象的推荐方法的描述，因此，这里将不再进行赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。对于本申请所涉及的计算机可读存储介质实施例中未披露的技术细节，请参照本申请方法实施例的描述。作为示例，计算机指令可被部署在一个计算机设备上执行，或者在位于一个地点的多个计算机设备上执行，又或者，在分布在多个地点且通过通信网络互连的多个计算机设备上执行，分布在多个地点且通过通信网络互连的多个计算机设备可以组成区块链系统。

[0308] 此外，需要说明的是：本申请实施例还提供了一种计算机程序产品或计算机程序，该计算机程序产品或者计算机程序可以包括计算机指令，该计算机指令可以存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器可以执行该计算机指令，使得该计算机设备执行前文图3或者图15所对应实施例中对针对媒体对象的推荐方法的描述，因此，这里将不再进行赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。对于本申请所涉及的计算机程序产品或者计算机程序实施例中未披露的技术细节，请参照本申请方法实施例的描述。

[0309] 本申请实施例中，术语“模块”或“单元”是指有预定功能的计算机程序或计算机程序的一部分，并与其他相关部分一起工作以实现预定目标，并且可以通过使用软件、硬件(如处理电路或存储器)或其组合来全部或部分实现。同样的，一个处理器(或多个处理器或存储器)可以用来实现一个或多个模块或单元。此外，每个模块或单元都可以是包含该模块或单元功能的整体模块或单元的一部分。

[0310] 以上所揭露的仅为本申请较佳实施例而已，当然不能以此来限定本申请之权利范围，因此依本申请权利要求所作的等同变化，仍属本申请所涵盖的范围。

权利要求书

[1] 一种针对媒体对象的推荐方法，其特征在于，包括： 获取待推荐媒体对象在目标采样窗口中的目标音频数据，对所述目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征所述待推荐媒体对象的目标文本表征向量； 获取用于表征N个历史媒体对象的N个历史文本表征向量，对所述目标文本表征向量和所述N个历史文本表征向量分别进行向量相似计算，得到N个向量相似度；所述每个历史媒体对象的交互数据量均高于所述待推荐媒体对象的交互数据量；N为大于1的正整数；所述N个历史文本表征向量中的每个历史文本表征向量是通过对每个历史音频数据对应的历史语音识别文本进行文本向量表征处理后所记录得到的，所述每个历史音频数据是指所述每个历史媒体对象在记录到的历史采样窗口中的音频数据；所述历史采样窗口的窗口时长与所述目标采样窗口的窗口时长保持一致； 在通过所述N个向量相似度从所述N个历史媒体对象中筛选出K个历史媒体对象时，建立所述K个历史媒体对象与所述待推荐媒体对象之间的相似关联关系；K为小于N的正整数； 在获取到所述K个历史媒体对象的K个历史推荐特征时，基于所述相似关联关系和所述K个历史推荐特征，确定用于表征所述待推荐媒体对象的目标推荐特征，通过所述目标推荐特征对所述待推荐媒体对象进行媒体推荐。

[2] 根据权利要求1所述的方法，其特征在于，所述获取待推荐媒体对象在目标采样窗口中的目标音频数据，包括： 获取所述待推荐媒体对象的目标动态采样时间点和窗口时长，基于所述目标动态采样时间点和所述窗口时长，确定用于对所述待推荐媒体对象的多媒体数据进行采样的目标采样窗口；所述目标动态采样时间点是指随着所述待推荐媒体对象的媒体数据关联时长的变化而变化的时间点； 通过所述目标采样窗口从所述待推荐媒体对象的多媒体数据中，获取数据采样结束时间点为所述目标动态采样时间点，且采样时长为所述窗口时长的音频数据； 将获取到的音频数据作为所述待推荐媒体对象在所述目标采样窗口中的目标音频数据。

[3] 根据权利要求1-2任一项所述的方法，其特征在于，所述目标动态采样时间点包括第一动态采样时间点，所述目标采样窗口包括所述第一动态采样时间点所对应的第一采样窗口，所述目标音频数据包括基于所述第一动态采样时间点和所述窗口时长所采集得到的第一音频数据； 所述方法还包括： 获取用于进行动态采样的采样间隔时长，基于所述第一动态采样时间点和所述采样间隔时长，确定用于对所述待推荐媒体对象的多媒体数据进行采样的第二动态采样时间点；所述第二动态采样时间点为所述第一动态采样时间点的下一采样时间点； 基于所述第二动态采样时间点和所述窗口时长，确定用于对所述待推荐媒体对象的多媒体数据进行采样的第二采样窗口；所述第二采样窗口为所述第一采样窗口的下一采样窗口，且所述第二采样窗口与所述第一采样窗口存在窗口重叠； 通过所述第二采样窗口从所述待推荐媒体对象的多媒体数据中，获取数据采样结束时间点为所述第二动态采样时间点，且采样时长为所述窗口时长的音频数据； 将获取到的音频数据作为所述待推荐媒体对象在所述目标采样窗口中的第二音频数据，在通过所述第二动态采样时间点更新所述目标动态采样时间点时，通过所述第二音频数据更新所述目标音频数据。

[4] 根据权利要求1-3任一项所述的方法，其特征在于，所述对所述目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征所述待推荐媒体对象的目标文本表征向量，包括： 获取用于对所述目标音频数据进行语音转换处理的语音转换理解组件和相似检索组件；所述语音转换理解组件包括语音识别器和用于进行语义提取的语义理解模型；所述相似检索组件包括文本表征器； 通过所述语音识别器对所述目标音频数据进行语音识别处理，得到语音识别处理结果； 若所述语音识别处理结果指示识别到所述目标音频数据的目标语音识别文本，则通过所述语义理解模型对所述目标语音识别文本进行文本语义理解处理，得到所述目标语音识别文本的目标语义理解信息； 通过所述文本表征器对所述目标语义理解信息进行向量表征，得到用于表征所述待推荐媒体对象的目标文本表征向量。

[5] 根据权利要求1-4任一项所述的方法，其特征在于，所述方法还包括： 若所述语音识别处理结果指示未识别到所述目标音频数据的目标语音识别文本，则将所述目标音频数据标记为无效音频数据。

[6] 根据权利要求1-5任一项所述的方法，其特征在于，所述方法还包括： 若所述语音识别处理结果指示未识别到所述目标音频数据的目标语音识别文本，则通过所述目标采样窗口从所述待推荐媒体对象的多媒体数据中，获取所述待推荐媒体对象的视频数据，将获取到的视频数据作为所述待推荐媒体对象在所述目标采样窗口中的目标视频数据； 对所述目标视频数据中的视频帧进行视频内容识别，得到用于表征所述目标视频数据的视频文本内容，通过视频文本内容更新所述目标语音识别文本。

[7] 根据权利要求1-6任一项所述的方法，其特征在于，所述通过所述语义理解模型对所述目标语音识别文本进行文本语义理解处理，得到所述目标语音识别文本的目标语义理解信息，包括： 获取用于引导语义理解模型进行模型输出的引导提示模板数据； 将所述引导提示模版数据和所述目标语音识别文本输入所述语义理解模型，通过所述语义理解模型按照所述引导提示模版数据中的数据内容，对所述目标语音识别文本进行语义信息提取，得到所述目标语音识别文本的目标语义提取信息； 通过所述目标语义提取信息确定所述目标语音识别文本的目标语义理解信息。

[8] 根据权利要求1-7任一项所述的方法，其特征在于，所述获取用于引导语义理解模型进行模型输出的引导提示模板数据，包括： 若所述待推荐媒体对象的对象类型为视频直播类型，则将与所述视频直播类型相关联的提示模版数据作为所述引导提示模板数据； 若所述待推荐媒体对象的对象类型为视频点播类型，则将与所述视频点播类型相关联的提示模版数据作为所述引导提示模板数据。

[9] 根据权利要求1-8任一项所述的方法，其特征在于，所述目标文本表征向量是通过语音转换理解组件和相似检索组件，对所述目标音频数据进行语音转换处理后所得到的； 所述获取用于表征N个历史媒体对象的N个历史文本表征向量，包括： 获取历史媒体对象数据库所对应的交互数据曲线分布；所述交互数据曲线分布是基于所述历史媒体对象数据库中的各个历史媒体对象和所述各个历史媒体对象的交互数据量所确定的； 通过所述交互数据曲线分布，从所述历史媒体对象数据库中获取交互数据量达到交互数据阈值的N个历史媒体对象； 从与所述N个历史媒体对象相关联的历史文本向量表征数据库中，获取所述N个历史媒体对象的N个历史文本表征向量；所述每个历史文本表征向量是通过所述语音转换理解组件和所述相似检索组件，对所述每个历史媒体对象的历史音频数据进行语音转换处理后所得到的。

[10] 根据权利要求1-9任一项所述的方法，其特征在于，所述在通过所述N个向量相似度从所述N个历史媒体对象中筛选出K个历史媒体对象时，建立所述K个历史媒体对象与所述待推荐媒体对象之间的相似关联关系，包括： 基于所述N个向量相似度对所述N个历史媒体对象进行排序处理，得到与所述N个历史媒体对象相关联的历史媒体对象列表； 从所述历史媒体对象列表中获取向量相似度满足相似检索条件的K个历史媒体对象，建立所述K个历史媒体对象与所述待推荐媒体对象之间的相似关联关系；所述相似关联关系用于表征由所述K个历史媒体对象所构成的列表为相似媒体对象列表，且在所述相似媒体对象列表中的任意一个媒体对象均与所述待推荐媒体对象存在关联关系。

[11] 根据权利要求1-10任一项所述的方法，其特征在于，所述在获取到所述K个历史媒体对象的K个历史推荐特征时，基于所述相似关联关系和所述K个历史推荐特征，确定用于表征所述待推荐媒体对象的目标推荐特征，包括： 在获取到所述K个历史媒体对象的K个历史推荐特征时，基于所述相似关联关系对所述K个历史推荐特征进行特征池化处理，得到与所述K个历史推荐特征相关联的池化处理特征； 将所述池化处理特征，确定为用于表征所述待推荐媒体对象的目标推荐特征。

[12] 根据权利要求1-11任一项所述的方法，其特征在于，所述的方法还包括： 从与所述K个历史媒体对象相关联的推荐特征数据库中获取所述K个历史媒体对象的K个历史推荐特征。

[13] 根据权利要求1-12任一项所述的方法，其特征在于，所述方法还包括： 获取N个历史媒体对象对应的N个历史交互行为信息；一个历史媒体对象对应一个历史交互行为信息，一个历史交互行为信息用于表征一个历史媒体对象中的各个交互物品的物品交互详情； 将所述N个历史交互行为信息所表征的物品交互详情作为用于训练初始推荐模型的训练样本，获取用于训练所述初始推荐模型的推荐指标； 通过所述训练样本和所述推荐指标，对所述初始推荐模型进行模型训练，得到目标推荐模型；所述目标推荐模型用

于提取得到所述N个历史媒体对象的N个历史推荐特征；一个历史推荐特征为所述目标推荐模型对一个历史交互行为信息所表征的物品交互详情进行交互向量提取所得到的； 将所述目标推荐模型所提取到的N个历史推荐特征，添加至所述推荐特征数据库。

[14] 根据权利要求1-13任一项所述的方法，其特征在于，所述待推荐媒体对象为业务处理平台中的交互数据量小于交互数据量阈值的多媒体对象；所述业务处理平台所对应的推荐模型为目标推荐模型；所述目标推荐模型是通过所述N个历史媒体对象对应的N个历史交互行为信息，对所述业务处理平台上所部署的初始推荐模型进行模型训练后所得到的； 所述通过所述目标推荐特征对所述待推荐媒体对象进行媒体推荐，包括： 在将所述目标推荐模型部署应用在所述业务处理平台时，获取访问所述业务处理平台的业务对象的对象关联数据信息； 将所述对象关联数据信息和所述目标推荐特征输入所述目标推荐模型，通过所述目标推荐模型在从所述对象关联数据信息中提取到用于表征所述业务对象的对象向量表征特征时，在所述对象向量表征特征中筛选与所述目标推荐特征相匹配的目标对象向量表征特征； 将所述目标对象向量表征特征所对应的业务对象，确定为在所述业务对象中所筛选出的与所述待推荐媒体对象相匹配的目标业务对象，并向所述目标业务对象推荐所述待推荐媒体对象。

[15] 根据权利要求1-14任一项所述的方法，其特征在于，所述将所述对象关联数据信息和所述目标推荐特征输入所述目标推荐模型，通过所述目标推荐模型在从所述对象关联数据信息中提取到用于表征所述业务对象的对象向量表征特征时，在所述对象向量表征特征中筛选与所述目标推荐特征相匹配的目标对象向量表征特征，包括： 将所述对象关联数据信息和所述目标推荐特征输入所述目标推荐模型，通过所述目标推荐模型，并通过所述目标推荐模型对所述对象关联数据信息进行信息向量表征提取，得到用于表征所述业务对象的对象向量表征特征； 将所述对象向量表征特征和所述目标推荐特征进行特征匹配，得到特征匹配结果； 若所述特征匹配结果指示所述对象向量表征特征中存在与所述目标推荐特征相匹配的对象向量表征特征，则将与所述目标推荐特征相匹配的对象向量表征特征，确定为所述目标对象向量表征特征。

[16] 根据权利要求1-15任一项所述的方法，其特征在于，与所述业务处理平台相关联的目标推荐业务包括直播推荐业务； 所述向所述目标业务对象推荐所述待推荐媒体对象，包括： 若所述待推荐媒体对象为直播类媒体对象，则将所述直播类媒体对象作为在所述直播推荐业务中的用于推荐给所述目标业务对象的目标推荐对象，将所述目标推荐对象发送给所述目标业务对象所对应的终端设备，以使所述终端设备在所述直播推荐业务所对应的直播推荐列表中显示所述目标推荐对象； 若所述待推荐媒体对象为点播类媒体对象，则将所述点播类媒体对象所对应的点播链接封装为推荐通知消息，将所述推荐通知消息发送给所述目标业务对象所对应的终端设备，以使所述终端设备在所述直播推荐业务所对应的直播推荐列表上显示所述推荐通知消息中的所述点播链接；所述点播链接用于指示所述终端设备响应于针对所述推荐通知消息的触发操作，显示所述点播类媒体对象。

[17] 一种针对媒体对象的推荐装置，其特征在于，包括： 音频数据语音识别模块，用于获取待推荐媒体对象在目标采样窗口中的目标音频数据，对所述目标音频数据所对应的目标语音识别文本进行文本向量表征处理，得到用于表征所述待推荐媒体对象的目标文本表征向量； 向量相似度计算模块，用于获取用于表征N个历史媒体对象的N个历史文本表征向量，对所述目标文本表征向量和所述N个历史文本表征向量分别进行向量相似计算，得到N个向量相似度；所述每个历史媒体对象的交互数据量均高于所述待推荐媒体对象的交互数据量；N为大于1的正整数；所述N个历史文本表征向量中的每个历史文本表征向量是通过对每个历史音频数据对应的历史语音识别文本进行文本向量表征处理后所记录得到的，所述每个历史音频数据是指所述每个历史媒体对象在记录到的历史采样窗口中的音频数据；所述历史采样窗口的窗口时长与所述目标采样窗口的窗口时长保持一致； 相似关联关系建立模块，用于在通过所述N个向量相似度从所述N个历史媒体对象中筛选出K个历史媒体对象时，建立所述K个历史媒体对象与所述待推荐媒体对象之间的相似关联关系；K为小于N的正整数； 目标推荐特征确定模块，用于在获取到所述K个历史媒体对象的K个历史推荐特征时，基于所述相似关联关系和所述K个历史推荐特征，确定用于表征所述待推荐媒体对象的目标推荐特征，通过所述目标推荐特征对所述待推荐媒体对象进行媒体推荐。

[18] 一种计算机设备，其特征在于，包括存储器和处理器； 所述存储器与所述处理器相连，所述存储器用于存储计算机程序，所述处理器用于调用所述计算机程序，以使得所述计算机设备执行权利要求1-16任一项所述的方法。

[19] 一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有计算机程序，所述计算机程序适于由处理器加载并执行，以使得具有所述处理器的计算机设备执行权利要求1-16任一项所述的方法。

[20] 一种计算机程序产品，其特征在于，包括计算机程序/指令，所述计算机程序/指令被处理器执行时实现权利要求1-16任一项所述的方法。