



(12)发明专利申请

(10)申请公布号 CN 110309427 A

(43)申请公布日 2019. 10. 08

(21)申请号 201810553549.1

(22)申请日 2018.05.31

(71)申请人 腾讯科技(深圳)有限公司

地址 518000 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 丘志杰 饶君 张博 林乐宇

冯喆 陈磊 胡澜涛 刘书凯

刘毅 孙振龙 王良栋

(74)专利代理机构 深圳市深佳知识产权代理事

务所(普通合伙) 44285

代理人 王仲凯

(51)Int. Cl.

G06F 16/9536(2019.01)

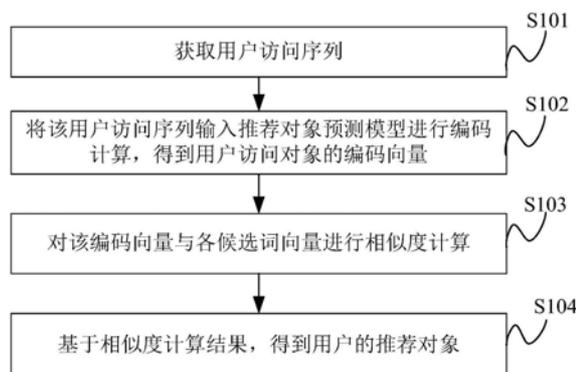
权利要求书3页 说明书21页 附图10页

(54)发明名称

一种对象推荐方法、装置及存储介质

(57)摘要

本发明实施例提供一种对象推荐方法、装置、存储介质及计算机设备,本实施例基于循环神经网络,对多个样本用户对应的用户访问序列进行训练,得到推荐对象预测模型,对用户访问序列进行编码计算,再对得到的编码向量与各候选词向量进行相似度计算,得到该用户的推荐对象。可见,基于循环神经网络的在编码计算的特点,使得本实施例这种模型编码方式,兼顾了用户的长期历史兴趣和短期历史兴趣,且考虑到了用户在应用平台上访问对象的访问顺序,能够更准确定位用户的兴趣变迁和兴趣积累,解决现有ItemCF对象推荐方法导致所得推荐对象多样性和个性化损失的问题。



1. 一种对象推荐方法,其特征在于,所述方法包括:
 - 获取用户访问序列,所述用户访问序列是基于用户访问应用平台输出的对象生成的;
 - 将所述用户访问序列输入推荐对象预测模型进行编码计算,得到所述用户访问对象的编码向量,所述推荐对象预测模型是基于循环神经网络,对多个样本用户对应的用户访问序列训练得到;
 - 对所述编码向量与各候选词向量进行相似度计算,所述候选词向量对应候选集合中的候选对象或候选关键词;
 - 基于相似度计算结果,得到所述用户的推荐对象。
2. 根据权利要求1所述的方法,其特征在于,所述循环神经网络包括有多个门控循环单元层,或多个长短期记忆网络层。
3. 根据权利要求1或2所述的方法,其特征在于,所述用户访问序列为访问对象序列,所述获取用户访问序列包括:
 - 获取用户的多条历史访问数据,所述历史访问数据是基于用户对应用平台输出对象的访问操作产生的;
 - 由所述多条历史访问数据分别包含的对象标识,构成访问对象序列。
4. 根据权利要求3所述的方法,其特征在于,所述方法还包括:
 - 获取多个候选访问对象;
 - 将各候选访问对象输入语言模型,得到相应的候选词向量。
5. 根据权利要求1或2所述的方法,其特征在于,所述用户访问序列为关键词序列,所述获取用户访问序列,包括:
 - 获取用户的多条历史访问数据;
 - 利用各历史访问数据,得到相应访问对象对应的关键词簇;
 - 由各访问对象对应的关键词簇,构成关键词序列。
6. 根据权利要求5所述的方法,其特征在于,所述方法还包括:
 - 获取多个候选关键词;
 - 将所述多个候选关键词分别输入语言模型,得到相应的候选词向量。
7. 根据权利要求5所述的方法,其特征在于,所述基于相似度计算结果,得到所述用户的推荐对象,包括:
 - 获取相似度最高的第一数量个候选词向量对应的候选关键词;
 - 由获取的第一数量个候选关键词,构成候选关键词簇;
 - 利用所述候选关键词簇包含的一候选关键词,与所述推荐对象预测模型中上一时刻隐藏层输出作为所述推荐对象预测模型中当前时刻隐藏层输入,继续进行编码计算;
 - 将得到的新的编码向量作为用户访问对象的编码向量,执行所述对所述编码向量与各候选词向量进行相似性计算步骤,直至相似度计算次数达到第二数量,或构成的候选关键词簇满足预设条件;
 - 利用构成的第二数量个候选关键词簇中同一维度候选关键词,生成第一数量个候选关键词序列;
 - 基于所述第一数量个候选关键词序列包含的候选关键词,得到所述用户的推荐对象。
8. 根据权利要求7所述的方法,其特征在于,所述基于所述第一数量个候选关键词序列

包含的候选关键词,得到所述用户的推荐对象,包括:

获取已构建的关键词映射到对象的倒排索引;

查询所述倒排索引,得到各候选关键词序列中的每一个候选关键词的倒排列表,所述倒排列表用于表征该候选关键词映射到的至少一个候选推荐对象;

基于得到的各候选关键词的倒排列表,统计各候选关键词序列中的各候选推荐对象被映射的次数;

基于统计结果,从获取的多个候选推荐对象中,筛选所述用户的推荐对象,其中,被映射次数越多的候选推荐对象被筛选为推荐对象的概率越大。

9. 根据权利要求7所述的方法,其特征在于,所述基于所述第一数量个候选关键词序列包含的候选关键词,得到所述用户的推荐对象,包括:

按照各候选关键词序列中各候选关键词的词序,对所述第一数量个候选关键词序列包含的候选关键词进行拼接;

获取多个候选访问对象分别对应的关键词序列,所述候选访问对象至少包含任一候选关键词序列中的一候选关键词;

对拼接得到的候选关键词序列与各候选访问对象对应的关键词序列进行文本相似度计算;

基于文本相似度计算结果,从所述多个候选访问对象中,筛选所述用户的推荐对象,其中,所述相似度越高的候选访问对象被筛选为推荐对象的概率越大。

10. 根据权利要求1或2所述的方法,其特征在于,所述方法还包括:

按照特定规则,对所述用户的推荐对象进行初步筛选,得到初选推荐对象;

获取各初选推荐对象对应的编码向量,及当前访问对象的编码向量;

对各初选推荐对象对应的编码向量与所述当前访问对象的编码向量进行相似度计算;

选择相似度最高的预设数量个初选推荐对象作为目标推荐对象;

将所述目标推荐对象发送至所述用户的客户端进行展示。

11. 一种对象推荐装置,其特征在于,所述装置包括:

序列获取模块,用于获取用户访问序列,所述用户访问序列是基于用户访问应用平台输出的对象生成的;

编码计算模型,用于将所述用户访问序列输入推荐对象预测模型进行编码计算,得到所述用户访问对象的编码向量,所述推荐对象预测模型是基于循环神经网络,对多个样本用户对应的用户访问序列训练得到;

第一相似度计算模块,用于对所述编码向量与各候选词向量进行相似度计算;

推荐对象选择模型,用于基于相似度计算结果,得到所述用户的推荐对象。

12. 根据权利要求11所述的装置,其特征在于,所述序列获取模块包括:

第一数据获取单元,用于获取用户的多条历史访问数据,所述历史访问数据是基于用户对应用平台输出对象的访问操作产生的;

第一序列构成单元,用于由所述多条历史访问数据分别包含的对象标识,构成访问对象序列。

13. 根据权利要求11所述的装置,其特征在于,所述序列获取模块包括:

第二数据获取单元,用于获取用户的多条历史访问数据;

关键词簇获取单元,用于利用各历史访问数据,得到相应访问对象对应的关键词簇;
第二序列构成单元,用于由各访问对象对应的关键词簇,构成关键词序列。

14.一种存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行,实现如权利要求1-10任意一项所述的对象方法的各步骤。

15.一种计算机设备,其特征在于,所述计算机设备包括:

通信接口;

存储器,用于存储实现如权利要求1-10任意一项所述的对象方法的计算机程序;

处理器,用于记载并执行所述存储器存储的计算机程序,所述计算机程序用于实现以下步骤:

获取用户访问序列,所述用户访问序列是基于用户访问应用平台输出的对象生成的;

将所述用户访问序列输入推荐对象预测模型进行编码计算,得到所述用户访问对象的编码向量,所述推荐对象预测模型是基于循环神经网络,对多个样本用户对应的用户访问序列训练得到;

对所述编码向量与各候选词向量进行相似度计算;

基于相似度计算结果,得到所述用户的推荐对象。

一种对象推荐方法、装置及存储介质

技术领域

[0001] 本发明涉及数据处理技术领域,具体涉及一种对象推荐方法、装置及存储介质。

背景技术

[0002] 如今,互联网的普及给用户带来了大量的信息,满足了用户在信息时代对信息的需求,但随着网络的迅速发展,网络信息量的大幅增长,用户面对大量信息时,很难从中获得对自己真正有用的那部分信息,对信息的使用效率反而降低。对此,技术人员提出了推荐系统,即根据用户的信息需求、兴趣等,将用户感兴趣的信息、产品等推荐给用户,实现个性化信息推荐,目前已被广泛应用到很多领域,如新闻推荐、商务推荐、娱乐推荐、学习推荐、生活推荐等。

[0003] 目前,常用的推荐方法主要有基于协同过滤(Collaborative Filtering,简称CF)的推荐方法,即通过计算Item-Item(Item可以是一篇文章或一个视频等对象)相似度的方式,得到与用户访问过的对象,最相似的K个Item作为推荐对象。

[0004] 其中,Item-Item相似度计算通常是采用Item CF的方法实现,即利用Item在用户的访问对象序列中的共现关系,计算两个Item的相似度。当不同用户在一段时间内访问对象Item相同,但各访问对象的访问顺序不同,基于Item CF的推荐方法得到的各用户的推荐对象将会相同,无法实现个性化推荐,还会影响推荐准确率,且得到的推荐对象无法兼顾用户的长期兴趣和短期兴趣。

发明内容

[0005] 本发明实施例提供一种对象推荐方法、装置、存储介质及计算机设备,可以实现对用户兴趣变迁和兴趣积累的准确定位,满足不同用户个性化推荐需求,且所得推荐对象兼顾了用户的长期兴趣和短期兴趣,提高了向用户推荐对象的准确性。

[0006] 为实现上述目的,本发明实施例提供如下技术方案:

[0007] 一种对象推荐方法,所述方法包括:

[0008] 获取用户访问序列,所述用户访问序列是基于用户访问应用平台输出的对象生成的;

[0009] 将所述用户访问序列输入推荐对象预测模型进行编码计算,得到所述用户访问对象的编码向量,所述推荐对象预测模型是基于循环神经网络,对多个样本用户对应的用户访问序列训练得到;

[0010] 对所述编码向量与各候选词向量进行相似度计算;

[0011] 基于相似度计算结果,得到所述用户的推荐对象。

[0012] 一种对象推荐装置,所述装置包括:

[0013] 序列获取模块,用于获取用户访问序列,所述用户访问序列是基于用户访问应用平台输出的对象生成的;

[0014] 编码计算模型,用于将所述用户访问序列输入推荐对象预测模型进行编码计算,

得到所述用户访问对象的编码向量,所述推荐对象预测模型是基于循环神经网络,对多个样本用户对应的用户访问序列训练得到;

[0015] 第一相似度计算模块,用于对所述编码向量与各候选词向量进行相似度计算;

[0016] 推荐对象选择模型,用于基于相似度计算结果,得到所述用户的推荐对象。

[0017] 一种存储介质,其上存储有计算机程序,所述计算机程序被处理器执行,实现如上所述的对象方法的各步骤。

[0018] 一种计算机设备,所述计算机设备包括:

[0019] 通信接口;

[0020] 存储器,用于存储实现上所述的对象方法的计算机程序;

[0021] 处理器,用于记载并执行所述存储器存储的计算机程序,所述计算机程序用于实现以下步骤:

[0022] 获取用户访问序列,所述用户访问序列是基于用户访问应用平台输出的对象生成的;

[0023] 将所述用户访问序列输入推荐对象预测模型进行编码计算,得到所述用户访问对象的编码向量,所述推荐对象预测模型是基于循环神经网络,对多个样本用户对应的用户访问序列训练得到;

[0024] 对所述编码向量与各候选词向量进行相似度计算;

[0025] 基于相似度计算结果,得到所述用户的推荐对象。

[0026] 基于上述技术方案,本发明实施例提供一种对象推荐方法、装置、存储介质及计算机设备,本实施例基于循环神经网络,对多个样本用户对应的用户访问序列进行训练,得到推荐对象预测模型,对用户访问序列进行编码计算,再对得到的编码向量与各候选词向量进行相似度计算,得到该用户的推荐对象。可见,基于循环神经网络的在编码计算的特点,使得本实施例这种模型编码方式,兼顾了用户的长期历史兴趣和短期历史兴趣,且考虑到了用户在应用平台上访问对象的访问顺序,能够更准确定位用户的兴趣变迁和兴趣积累,解决现有Item CF推荐方法导致所得推荐对象多样性和个性化损失的问题。

附图说明

[0027] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据提供的附图获得其他的附图。

[0028] 图1为一种循环神经网络中每个GRU层的网络结构示意图;

[0029] 图2为本发明实施例提供一种对象推荐方法的流程示意图;

[0030] 图3为本发明实施例提供一种推荐对象预测模型的网络架构示意图;

[0031] 图4为本发明实施例提供的另一种对象推荐方法的流程示意图;

[0032] 图5为本发明实施例提供的另一种对象推荐方法的流程示意图;

[0033] 图6为本发明实施例提供的另一种对象推荐方法的流程示意图;

[0034] 图7为本发明实施例提供一种关键词序列生成方法示意图;

[0035] 图8为本发明实施例提供的另一种对象推荐方法的流程示意图;

- [0036] 图9为本发明实施例提供的另一种对象推荐方法的流程示意图；
- [0037] 图10为本发明实施例提供的又一种对象推荐方法的流程示意图；
- [0038] 图11为本发明实施例提供的一种对象推荐方法的应用流程示意图；
- [0039] 图12为本发明实施例提供的一种对象推荐装置的结构示意图；
- [0040] 图13为本发明实施例提供的另一种对象推荐装置的结构示意图；
- [0041] 图14为本发明实施例提供的另一种对象推荐装置的结构示意图；
- [0042] 图15为本发明实施例提供的又一种对象推荐装置的结构示意图；
- [0043] 图16为本发明实施例提供的一种计算机设备的硬件结构示意图。

具体实施方式

[0044] 本发明的发明人研究发现：当用户A在一定时间内阅读文章有[X、Y、Z]，用户B阅读的文章有[Y、Z、X]，X、Y和Z可以是用户阅读文章的文章ID，但并不局限于此，在获取这两个用户各自的推荐文章时，主要是由于基于Item CF的对象推荐方法仅考虑历史阅读文章内容，并未考虑用户阅读文章的阅读顺序，导致基于相似度计算结果得到的推荐对象相同，无法准确体现这两个用户各自的兴趣变迁和兴趣积累。

[0045] 而且，发明人还发现在进行相似度计算时，很难控制使用多少个候选对象进行相似性计算，若选择最近一段时间的候选对象进行相似度计算，所得推荐对象只能表征用户短期兴趣；若选择历史较长的候选对象进行相似度计算，不仅计算量大，且会导致得到的推荐对象非常多，需要进一步利用排序算法对得到的大量推荐对象进行排序筛选，过程比较繁琐，影响了获取用户的推荐对象的效率。

[0046] 基于上述分析，发明人提出一种新的对象推荐方法，不仅会考虑用户在应用平台访问对象（即阅读对象、观看视频等）的访问顺序，更加精确地定位用户的兴趣变迁和兴趣积累，而且兼顾了用户的长期历史兴趣和短期历史兴趣，提高了用户的推荐结果的准确性和效率，满足不同用户的个性化推荐需求。

[0047] 下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0048] 为了方便理解本实施例提供的推荐方法，在此对本实施例涉及到的循环神经网络(Recurrent neural Network, RNN)的原理进行简单说明。RNN是一种节点定向连接成环的人工神经网络，其内部状态可以展示动态时序行为，且不同于前馈神经网络的是，多层反馈RNN可以利用内部的记忆来处理任意时序的输入序列，使其能够更容易处理如不分段的手写识别、语音识别等。

[0049] 其中，RNN作为一种可处理变长数据的神经网络，可以将任意长度的历史信息编码到一个隐藏层(hidden layer)，即神经网络的中间输出，表征输入的某种隐含表现形式，通常为一个向量或矩阵。如下公式(1)所示，RNN可以对高纬度数据的压缩降维。尤其是随着近几年RNN中使用LSTM(Long Short-Term Memory, 长短期记忆网络)和GRU(Gated Recurrent Unit, 门控循环单元)的广泛应用，RNN被成功用来解决自然语言处理NLP(Natural Language Process)的技术问题，如机器翻译，序列预测，语音信号处理等。

$$[0050] \quad h_t = g(Wx_t + Uh_{t-1}) \quad (1)$$

[0051] 公式(1)中, x_t 可以表示当前输入的代表向量(在本实施例中可以是访问对象或访问对象中的关键词的embedding向量,本实施例也可以将该embedding向量记为词向量), h_{t-1} 可以表示上一时刻隐藏层输出,初始时刻的 h_{t-1} 为一个零向量, W 和 U 分别表示映射矩阵。在实际应用中,这个计算公式可以用来将当前输入与历史信息进行计算融合,得到一个新的编码向量(本实施例中,该编码向量可以是encoding向量)。

[0052] 可见,在RNN的应用中,在每一个时刻,会将历史信息和当前输入进行有效的编码计算,得到一个新的数据表达向量,即编码向量。在本实施例中,该历史信息可以是用户访问应用平台输出的对象产生的历史访问数据(如阅读文章、观看视频等产生的数据等),当前输入可以是当前时刻的访问对象的词向量或该访问对象的关键词的词向量等。

[0053] 其中,在循环神经网络隐藏层的编码计算过程中,如上述公式表示,其是由上一时刻的输出加上本时刻的输入,经过一个激活函数如tanh函数,得到下一时刻的输出,以此类推,最后输出的即为本发明需要的编码向量。

[0054] 由于传统RNN并未考虑到随着展开层数的增加,网络深度会变得很深,导致反向梯度传播出现异常,如梯度弥散和梯度爆炸。为了解决该问题,本实施例使用的循环神经网络可以是在传统RNN中增加了GRU或LSTM模型,得到的变型循环神经网络,在本实施例的循环神经网络中通常包括多个GRU层或多个LSTM层,具体网络结构本实施例不作详述。本实施例在此仅以包括多个GRU层的循环神经网络为例,来说明获取编码向量的计算过程,关于GRU的原理本实施例在此不做详述。

[0055] 在利用GRU计算隐藏层时,如图1所示,可以认为上一时刻隐藏层输出是当前隐藏层输出的线性加权组合的一部分,具体参照公式(2):

$$[0056] \quad h_t = (1 - z_t)h_{t-1} + z_t \hat{h}_t \quad (2)$$

[0057] 公式(2)中, \hat{h}_t 可以表示当前(即 t 时刻)的候选隐藏输出,即当前隐藏层的中间输出,需要与上一时刻的隐藏层输出进行加权融合,得到最终的隐藏层输出;加权因子 z_t 可以由上一时刻隐藏层输出 h_{t-1} 和当前的输入 x_t 进行自动适配计算得到,计算公式如下公式(3):

$$[0058] \quad z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (3)$$

[0059] 可见,随着上一时刻隐藏层输出 h_{t-1} 和当前的候选隐藏层输入的变化,加权因子 z_t (即GRU的更新门)也可以所有改变,但该加权因子 z_t 最终被映射到(0,1]区间内, z_t 越大表示越看重当前的输入信息,给予更高的权重,本实施例对该 z_t 的具体数值不作限定。其中,在GRU中,公式(3)中的 U 和 W 通常都比较小, σ 可以是系数因子,如图1中的左侧第二个 σ ,本实施例对 U 、 W 、 σ 具体数值不作限定。

[0060] 另外,对于上述公式(2)中的当前的候选隐藏输出 \hat{h}_t ,可以利用激活函数,如双曲正切函数tanh计算得到,其可以保留原始输入的符号信息和大小信息,具体计算公式(4)如下:

$$[0061] \quad h = \tanh(Wx_t + U(r_t \otimes h_{t-1})) \quad (4)$$

[0062] 在公式(4)中 r_t 可以重置门,且上述公式中的 z_t 可以是更新门,即GRU模型中的两个

门,该重置门可以采用公式(5)所示的方式计算得到,但并不局限于此。

$$[0063] \quad r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (5)$$

[0064] 公式(5)中 σ 可以是系数因子,如图1中的左侧第一个 σ 本实施例对其具体数值不作限定。并且,需要说明,关于上述各实施例中相同字母表示的含义相同,可以参照上述对公式(1)对应部分的解释。

[0065] 结合上述公式(2)~(5)的描述,以及图1所示的GRU的结构,其包含两个门,即重置门 r_t 和更新门 z_t ,由重置门 r_t 乘以上一时刻隐藏层的输出 h_{t-1} ,看是否要重置或重置多大程度,之后,和当前时刻的输入 x_t 拼接,经过激活函数 \tanh 运算,得到隐含变量 h_t ,然后将上一时刻的输出 h_{t-1} 与该隐含变量 h_t 进行线性组合,得到当前时刻的输出,依次类推,得到预测所需的编码向量,其中,进行线性组合的两者权重和为1,且该隐含变量 h_t 的权重即为更新门的输出,表征更新强度多大。

[0066] 在本实施例实际应用中,在获取当前用户的推荐对象之前,需要先训练得到对象推荐预测模型,结合上述循环神经网络原理描述,可知该循环神经网络在对输入进行编码计算时,会考虑输入的时间顺序,满足了对用户的预测结果的个性化要求,且模型输入无论是较长或较短数据,其在编码期间就实现了输出结果的排序,使得由此得到的推荐结果,能够兼顾用户的长期历史兴趣和短期历史兴趣。

[0067] 基于此,在本实施例中,可以获取多个样本用户对应的用户访问序列,基于上述循环神经网络,对这些用户访问序列进行模型训练,得到推荐对象预测模型,本实施例对该模型训练的具体实现过程不作限定。

[0068] 其中,样本用户的用户访问序列可以由该样本用户在应用平台上的访问对象的对象标识构成,如按照访问时间顺序,由从历史访问数据中提取的访问对象ID生成用户访问序列,此时该用户访问序列可以称为访问对象序列,本实施例可以将该访问对象序列作为训练数据,实现模型训练。

[0069] 在对模型不断训练优化过程中,可以在本次得到的模型预测结果(如推荐对象)与实际访问对象的差值满足预设条件,即本次训练模型满足约束条件,可以将本次训练得到的模型作为推荐对象预测模型。但对该约束条件的内容不作限定,且本实施例也可以限定优化次数或训练数据条数,以控制模型的训练优化处理,本实施例对如何得到推荐对象预测模型的实现方法不作限定。

[0070] 作为一可选实施例,为了提高推荐结果的实时性,本实施例在获取训练数据过程中,可以获取每个用户的每个session内的序列作为训练数据。该session信息可以是用户当前界面刷新到下一次刷新之前的时间窗口,其内部的信息在时间差异上通常都较小,使用这种训练数据进行模型训练,再利用得到的推荐对象预测模型,得到预测的推荐对象能够在更加短的时间间隔内得到用户的反馈,以保证推荐结果实时性,更加容易抓住用户的短期兴趣。按照这种构思,本实施例可以扩大训练数据的历史长度,获取用户的长期兴趣。

[0071] 进一步,为了使训练数据更加丰富,本实施例可以针对一个样本用户的历史访问数据,构成多条训练数据。假设由用户在应用平台上访问对象的对象标识,得到的访问对象序列为 $[x_1, x_2, x_3, x_4, x_5, x_6]$,本实施例可以由此构成出如下多条训练数据如: $([x_1], x_2)$, $([x_1, x_2], x_3)$, $([x_1, x_2, x_3], x_4)$, $([x_1, x_2, x_3, x_4], x_5)$ 以及 $([x_1, x_2, x_3, x_4, x_5], x_6)$,其中,圆括号左边的为当前已知访问对象序列,圆括号右边的是需要预测的目标对象。在进行

模型训练过程中,可以利用这多条训练数据进行模型训练,以提高所得推荐对象预测模型的预测准确性,本实施例在此对如何基于循环神经网络,对上述多条训练数据进行模型训练的过程不作详述。

[0072] 需要说明,对于本实施例得到的推荐对象预测模型,可以随着时间的推移,利用样本用户更新的用户访问序列,对该推荐对象预测模型进行优化,以提高了预测准确性,本实施例在此对其优化过程不做详述。

[0073] 参照图2,为本发明实施例提供了一种对象推荐方法的流程示意图,该方法可以应用于服务侧,即该方法可以由服务器执行实现,具体可以包括但并不局限于以下步骤:

[0074] 步骤S101,获取用户访问序列;

[0075] 其中,该用户访问序列可以基于用户访问应用平台输出的对象生成,本实施例对该用户访问序列的具体生成方式不作限定,且根据需要,该用户访问序列包含的序列元素可以是访问对象的对象标识,或者是访问对象的关键词,对于不同内容的用户访问序列,其生成方式往往是不同的,具体可以参照下文相应实施例的描述。

[0076] 在本发明中,可以将序列元素是对象标识的用户访问序列称为访问对象序列,将序列元素为关键词的用户访问序列称为关键词序列,实施例的用户访问序列的类型并不局限于本实施例列举的这两种序列。

[0077] 步骤S102,将该用户访问序列输入推荐对象预测模型进行编码计算,得到用户访问对象的编码向量;

[0078] 如上文描述,该推荐对象预测模型可以基于循环神经网络,对多个样本用户对应的用户访问序列训练得到,对于不同的训练数据,得到的推荐对象预测模型的表示形式可以不同,模型输出数据表示的含义可以不同,但该模型的处理逻辑可以相同,本实施例对推荐对象预测模型对输入序列的编码计算过程不作详述。

[0079] 其中,在向推荐对象预测模型输入用户访问序列时,该用户访问序列中的各序列元素是依次输入该推荐对象预测模型的,此时,推荐对象预测模型的一个输入可以是相应时刻的序列元素。

[0080] 以图3所示的推荐对象预测模型架构示意图为例,对用户访问序列输入模型的处理流程进行减压说明,本实施例的推荐对象预测模型可以包括多个GRU层,即图3中的多个GRU Cell,结合上述GRU原理的描述,每一GRU层的输入为前一时刻隐藏层输出和当前输入,输出为下一时刻隐藏层状态信息。

[0081] 其中,在每一GRU层的计算过程中,通常是利用其包含的重置门和更新门,实现候选隐藏层的计算,且其控制的是保留多少前一时刻隐藏层的信息,以及控制加入多少候选隐藏层的信息,以得到输出。因此,采用本实施例提供的推荐对象预测模型实现对用户访问序列的编码计算,能够利用多个GRU层灵活控制长短距离的依赖信息,适合刻画序列数据,即在保留了用户在应用平台上很久之前的阅读文章同时,又能够突出近期阅读文章,使得本实施例得到的用来预测用户推荐对象的编码向量,不仅兼顾用户的长期兴趣和短期兴趣,能够考虑用户访问对象的访问顺序,提高了预测准确性,其满足了不同用户各个个性化推荐要求。

[0082] 需要说明,步骤S102得到的编码向量可以是推荐对象预测模型中最后一层隐藏层输出的编码向量。且上述推荐对象预测模型并不局限于图3所示的架构示意图,其中间层也

可以包括多个LSTM层,可以基于LSTM的原理,实现隐藏层计算,得到用户访问对象的编码向量,本实施例对这种推荐对象预测模型的编码计算过程不作详述。

[0083] 由于LSTM是一种时间递归神经网络,通常适合处理和预测时间序列中间隔和延迟相对较长的重要事件,其通常是引入三个门控装置,来处理记忆单元的记忆/遗忘、输入程度、输出程度的问题,结构比较复杂;而GRU可以引入Reset Gate(重置门)和Update Gate(更新门),需要的参数少,训练速度更快,结构也相对简单,本实施例可以根据实际需要,选择基于哪种循环神经网络实现推荐对象模型的训练,及对当前用户的用户访问序列的编码计算,本文仅以图3所示的架构示意图为例进行说明。

[0084] 步骤S103,对该编码向量与各候选词向量进行相似度计算;

[0085] 需要说明,该编码向量与候选词向量维度保持一致,当然,上述编码计算过程中,所产生的词向量与最终得到的编码向量维度也是相同的,从而保证相似性计算的正常进行,本实施例对向量的维度的具体内容及维数不作限定。

[0086] 在不同场景实施例中,该候选词向量获取过程可以不同,比如由用户历史访问对象直接得到的词向量,或者是由用户历史访问的关键词得到的词向量等等,本实施例对词向量的具体生成方式不作限定,比如采用Word2Vec获取访问对象或关键词的词向量等,但并不局限于此。

[0087] 其中,Word2Vec可以通过给每一个词语分配一个稠密的向量,与离散型特征处理-独热码one-hot标识方法形成对比,能够保持词语与词语之间的语义维度信息,本实施例对推荐对象预测模型中的Embedding Layer对每一个访问对象或关键计算,生成对应的embedding的具体实现方法不作详述。

[0088] 可选的,本实施例可以采用如余弦相似度(即Cosine相似度)这种相似度算法,实现对两个向量之间的相似度计算,但并不局限于这一种相似度计算方式,本实施例仅以此为例进行相似度计算的说明。

[0089] 其中,余弦相似度计算采用的公式如下公式(6):

$$[0090] \quad \text{Cos}(u,v) = \frac{\sum_{i=1}^n u_i \times v_i}{\sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2}} \quad (6)$$

[0091] 公式(6)中,u,v分别表示用户访问对象的编码向量和候选词向量,这两个向量的维度相同,而 u_i 表示该编码向量中第i维度的特征值, v_i 表示该候选词向量中第i维度的特征值。

[0092] 步骤S104,基于相似度计算结果,得到用户的推荐对象;

[0093] 可选的,在本实施例中,对于存储的多个候选对象来说,其对应的词向量与编码向量的相似度越高,表明相应候选对象成为该对象的推荐对象的概率越大,即相应候选对象成为预测推荐对象的可能性越大。可见,相似度越大的候选对象,用户对其感兴趣的可能性越大。

[0094] 基于此,本实施例可以选择达到预设阈值的相似度对应的候选对象,作为该用户的推荐对象;或者选择若干个最高相似度对应的候选对象,作为用户的推荐对象等等,本实施例对如何利用相似性计算结果,得到该用户的推荐对象的实现方式不作限定。

[0095] 作为另一实施例,若上述候选词向量是关键词的词向量,同理,相似度越高,说明该用户对相应关键词越感兴趣,将相应关键词选为候选关键词,以确定推荐对象的概率越大,所以说,在这种情况下,相似度大小可以表示用户对相应关键词感兴趣的概率大小,进而可以表示包含相应关键词的候选对象成为推荐对象的概率大小,本实施例可以选择相似度较高的若干个关键词,来确定出用户最可能感兴趣的推荐对象,具体实现过程可以参照下文相应实施例的描述。

[0096] 综上所述,本实施例基于循环神经网络,对多个样本用户对应的用户访问序列进行训练,得到推荐对象预测模型,实现对用户访问序列进行编码计算,不仅兼顾到用户的长期历史兴趣和短期历史兴趣,且考虑到了用户在应用平台上访问对象的访问顺序,这样,基于得到的编码向量与各候选词向量的相似度计算结果,得到该用户的推荐对象能够准确定位用户的兴趣变迁和兴趣积累,解决了现有Item CF推荐方法导致所得推荐对象多样性和个性化损失的问题。

[0097] 作为本发明一可选实施例,在进行推荐对象预测模型的训练过程中,可以将样本用户在应用平台上上述的历史访问对象Item,直接成为访问对象序列,由该访问对象序列得到进行模型训练的训练数据,基于循环神经网络进行模型训练,得到推荐对象预测模型。这种情况下,本发明可以采用图4所示的对象推荐方法,来获取当前用户的推荐对象,以便将其作为候选项实现后续初选逻辑,进而经过排序逻辑得到向用户客户端推送的目标推荐对象,方便用户快速且准确访问所需的对象。

[0098] 如图4所示的另一种对象推荐方法的流程示意图,该方法也可以应用在服务器侧,具体可以包括但并不局限于以下步骤:

[0099] 步骤S201,获取用户的多条历史访问数据;

[0100] 在实际应用中,用户登录应用平台,应用平台通常会输出多个访问对象(如文章、视频、图片等),由于访问对象的内容过多,往往是输出各访问对象的概要或标签,需要用户进入访问对象的显示界面,才能展示出该访问对象的具体内容。

[0101] 比如,新闻阅读应用平台上,通常会输出很多条新闻的标题,用户可以根据各条新闻的标题,选择感兴趣的新闻,进入该新闻内容显示界面,阅读这条新闻的详细内容。在视频播放应用平台上,也会显示很多个视频,为了方便用户选择,每一个视频也会有相应的标题或简要说明,用户选择感兴趣的视频才会进入该视频播放界面,才会播放视频内容。

[0102] 可见,本实施例上述应用平台可以是用户常用的各种APP的应用平台,如音视频应用平台、浏览器应用平台、即时通信应用平台等社交应用平台等等,相应地,该应用平台输出的供用户访问的访问对象可以是一个视频、一篇文章等,可以记为Item。

[0103] 而且,用户在应用平台上进行对象访问操作过程中,将会产生相应的访问数据,该访问数据可以作为历史访问数据存储到应用平台的数据库中,用来表明用户在应用平台上的历史访问行为。需要说明,用户在应用平台上进行操作期间,除了会产生历史访问数据外,还可以产生其他历史行为数据,本实施例主要对历史访问数据进行分析,对其他行为数据不作详述。

[0104] 其中,上述历史访问数据包括访问对象的对象标识、访问对象的内容、标题等信息,该对象标识表明用户访问哪个对象,如阅读了哪篇文章,观看了哪个视频等,即用来区分各访问对象。因此,本实施例中的对象标识可以是对象ID等,本实施例对历史访问数据包

含的内容,及对象标识指代的内容不作限定。

[0105] 可选的,应用平台存储历史行为数据(其包含历史访问数据)时,可以根据用户标识进行分类存储,即不同用户在应用平台上进行操作,所产生的历史行为数据,可以与该用户的用户标识关联后存储,以便后续快速查找到某用户的历史行为数据,但本实施例对历史行为数据的具体存储方式不作限定,用户标识可以包括但并不局限于用户账号。基于此,本实施例可以直接从应用平台的数据库中,查询与目标用户的用户标识关联的历史访问数据,但并不局限于这一种获取方式。

[0106] 步骤S202,由多个历史访问数据分别包含的对象标识,构成访问对象序列;

[0107] 如上所述,本实施例获取的每一条历史访问数据可以包括表明当前访问对象的对象标识,如对象ID,在获取有效时长内的多条历史访问数据后,可以提取各历史访问数据包含的对象标识,并按照各条历史访问数据的生成时间(即相应访问对象的访问时间),生成相应的访问对象序列,也就是说,该对象序列中的元素可以是用户在应用平台上的各访问对象的对象标识构成,且这些元素之间按照生成时间分布,如按照生成时间的先后顺序,对各访问对象的对象标识进行排序,得到对象序列,本实施例对对象序列的具体生成方式不作限定。

[0108] 举例说明:如用x1、x2、x3、x4、x5、x6等表示各访问对象的对象ID,本实施例生成的访问对象序列可以是[x1, x2, x3, x4, x5, x6],可见,由该访问对象序列的序列元素内容,可以确定用户在应用平台上有效时间内访问对象有哪些。若访问对象为文章,可以得知用户在有效时间内阅读了哪些文章,且阅读的这些文章将按照阅读顺序进行排序,以便精确定位用户的兴趣变化和兴趣积累。步骤S203,将该访问对象序列输入推荐对象预测模型进行编码计算,得到用户访问对象的编码向量

[0109] 其中,该推荐对象预测模型可以基于循环神经网络,对多个样本用户对应的访问对象序列训练得到,具体训练过程可以参照上文实施例相应部分的描述。

[0110] 在得到推荐对象预测模型之后,可以将该访问对象序列包含的各序列元素依次输入该推荐对象预测模型,该推荐对象预测模型对输入序列的编码计算过程可以参照上述对RNN原理的描述过程,本实施例在此不作详述。

[0111] 参照图3所示的推荐对象预测模型架构示意图,在本实施例中,输入层输入的是用户的访问对象序列,其序列元素为该用户曾在应用平台上访问的访问对象Item,经Embedding Layer的降维计算后,将生成各访问对象Item对应的embedding向量(本实施例可以将其记为词向量)。其中,循环神经网络的Embedding Layer的实际上就是数据降维处理层,本实施例具体可以利用Word2Vec实现,具体实现方法不作限定。

[0112] 在得到的各Item对应的embedding向量后,即各时刻对应的embedding向量,可以输入多个GRU Cell构成的中间层,在每层计算过程中,利用如上述公式(2)、(3)、(4)和(5)进行编码计算,即利用上一时刻隐藏层输出(其通常是一个编码向量)及当前时刻的输入信息(即当前时刻对应的词向量)进行不断迭代更新,最后输出的encoding向量即为用户访问对象的编码向量。

[0113] 需要说明的,关于对访问对象序列进行编码计算,以得到编码向量的方式,并不局限于图3所示的模型架构示意图,也可以利用多个LSTM层构成得到推荐对象预测模型实现,可以参照LSTM原理实现,本实施例在此不再详述。

[0114] 另外,在本实施例编码计算过程中,需要每个访问对象Item生成的embedding向量的维度与encoding向量的维度保持一致,以保证后续能够进行相似性计算,且本实施例对向量的维度的具体内容及维数不作限定。

[0115] 步骤S204,获取多个候选访问对象;

[0116] 可选的,本实施例可以选择应用平台输出的优质访问对象作为候选对象,构成候选集合,其中,优质对象可以是应用平台上访问率较大的对象、最近社交网络中的热门话题相关的对象、最近一段时间访问量较多的对象、或者随机在应用平台输出对象采集到的对象等等,本实施例对选择优质访问对象(即候选集合的组成元素)的方式不作限定,并不局限于本实施例列举的这几种选择方式。

[0117] 在实际应用中,候选集合中的候选对象可以随着时间的变化,按照选择候选对象的方式,对候选集合的组成元素不断更新,以提高后续据此得到用户的推荐对象的准确性,本实施例对候选集合的更新方式不作限定。

[0118] 步骤S205,将各候选访问对象输入语言模型,得到相应的候选词向量;

[0119] 可选的,该语言模型可以是Word2Vec,但并不局限于此。若将候选集合中的每一个元素(即候选对象)记为 X_n , n 为正整数,具体数值大小不作限定,由这些候选对象可以得到的候选集合可以为 $[X_1, X_2, X_3 \cdots X_n]$,之后,可以计算每一个候选对象的embedding向量(此时可以记为候选词向量),该实现方法与上述计算各访问对象Item对应的embedding向量的计算方法可以相同,也可以不同,本实施例对获取候选词向量的实现方法不作限定。

[0120] 需要说明,上述步骤S204和步骤S205可以在相似度计算之前的任意步骤执行,并不局限于本实施例描述的该位置。本实施例可以对获取的候选集合进行存储,当需要进行相似度计算时,可以计算该候选集合中各候选访问对象对应的候选词向量,即本实施例描述的方式。

[0121] 当然,本实施例也可以预先计算出该候选集合中各候选访问对象对应的候选词向量,直接对各候选词向量进行存储,这样,当需要与编码向量进行相似度计算时,直接获取预先存储的各候选访问对象对应的候选词向量,不需要在线计算,提高了工作效率。这种情况下,步骤S204和步骤S205可以合并为:获取多个候选访问对象分别对应的候选词向量,其他步骤相同,本实施例不再单独举例说明。

[0122] 步骤S206,对该编码向量与各候选对象的候选词向量进行相似度计算;

[0123] 需要说明,本实施例对如何计算两个向量之间的相似度的方法不作限定,可以采用如上文描述的Cosine相似度计算方法,也可以采用距离的倒数计算方法,或者皮尔森相关系数计算方法等等,本实施例在此不做一一详述。

[0124] 步骤S207,基于相似度计算结果,从多个候选对象中,筛选该用户的推荐对象。

[0125] 本实施例得到各候选访问对象与预测结果的相似度后,即各候选词向量与编码向量之间的相似度,由于该相似度大小可以表征用户对相应候选访问对象感兴趣的概率大小,相似度越大,用户越可能对相应候选访问对象感兴趣,选择访问相应的访问对象。因此,本实施例可以直接选择向相似度最高的若干个候选访问对象作为推荐对象,或者相似度达到预设阈值的候选访问对象作为推荐对象等等,本实施例对基于相似度计算结果,得到用户的推荐对象的实现方式不作限定。

[0126] 举例说明,用户U1在某应用平台上阅读文章的文章序列为 $[X_{u1}, X_{u2}, X_{u3}, X_{u4}, X_{u5},$

x_{u6}],该应用平台的候选集合包含的候选文章有10篇,如 $[X1, X2, X3, X4, X5, X6, X7, X8, X9, X10]$,按照上述方式对用户阅读文章的文章序列进行编码计算,得到编码向量 E_{u1} ,计算 E_{u1} 与每一个候选文章对应的候选词向量 E_j 之间的Cosine相似度, j 表示第几个候选文章,确定相似度最高的三个词向量(或相似度达到预设阈值的词向量)对应的候选文章为 $X3, X5$ 和 $X7$,本实施例可以将这三个候选文章作为推荐文章,由此构成本次召回逻辑得到的召回结果,以供推荐系统中后续筛选逻辑从中最终向用户 $U1$ 客户端推送的目标推荐文章。

[0127] 综上所述,本实施例利用用户的访问对象构成模型的训练数据,基于循环神经网络训练得到的推荐对象预测模型,实现用户访问对象的访问对象序列的编码计算,得到的编码向量兼顾了用户的长期历史兴趣和短期历史兴趣,进而提高了最终向用户推送的目标推荐对象的准确性,且由于考虑到了用户访问对象的访问顺序,能够更准确定位用户的兴趣变迁和兴趣积累,解决了现有Item CF的对象推荐方法导致召回结果多样性和个性化损失的问题。

[0128] 本发明的发明人提出上述实施例描述的对象推荐方法后,发明人发现这种方法虽然解决了现有技术基于Item CF的对象推荐方法的问题,但是在进行相似性计算过程中,候选集合包含的候选访问对象数量级往往都很大,特别是像针对目前各种即时通信应用程序,其应用平台输出的访问对象数量是千万级别,且随着应用平台用户规模的增加和访问对象来源的扩展,候选集合只会越来越大,将来会到亿级别甚至十亿级别,更或甚至到百亿级别,如果要针对每一个候选访问对象的候选词向量都进行相似度计算,线上系统根本无法做到实时计算。可见,上述实施例提出的对象推荐方法中,候选集合数量巨大成为其实施的重要难题。

[0129] 针对这个问题,发明人提出对应用平台能够输出的访问对象,按照一定的采样策略筛选出若干候选访问对象,构成候选集合,本实施例对该采样策略不作限定,但是,无论按照什么内容的采样策略,上述可选实施例描述的对象推荐方法都会带来预测精度的损失。

[0130] 因此,为了进一步提高预测精度,本发明人提出对上述方案的改进,以访问对象为用户阅读的文章为例进行分析,考虑到每篇文章都是由关键词构成,虽然应用平台输出的文章数量会越来越多,但是构成文章的关键词的总数量是一个相对稳定的集合,其通常不会因为用户规模的增大和文章来源的扩展而发生大的变化。所以,本实施例可以将上述实施例对用户的访问对象序列(此处可以是文章序列)进行编码计算,更改为对关键词序列进行编码计算,具体实现过程可以参照下文实施例描述的方法,但并不局限于下文实施例的实现方式。

[0131] 参照图5和图6,为本发明实施例提供的另一种对象推荐方法的流程示意图,该方法仍可以应用于服务器,具体可以包括以下步骤:

[0132] 步骤S301,获取用户的多条历史访问数据;

[0133] 步骤S302,基于各历史访问数据,得到相应访问对象对应的关键词簇;

[0134] 结合上述对历史访问数据的分析,本实施例可以基于获取的历史访问数据,得知用户在过去的各时刻对应用平台输出对象的访问情况,即曾经访问了哪些对象,即各访问对象的访问顺序。得到用户在应用平台上的访问对象的标题或内容后,可以从中提取相应的关键词,生成相应的关键词簇,本实施例对如何从各访问对象提取相应关键词的方式不

作限定,但在本实施例中,提取各访问对象包含的关键词时采用的提取规则一致。

[0135] 举例说明:假设用户U2的对象序列为 $[x_1, x_2, x_3]$,即在有效时长内,获取的用户U2在应用平台上的访问对象有三个,可以由相应的对象ID构成该对象序列,从每一个对象ID对应的访问对象中提取的关键词可以记为 $Tag_{n,n}$, n 为整数,在实际应用中,若对象为文章,可以从文章的标题中提取包含的至少一个关键词,当然也可以从文章内容中提取至少一个关键词等等,且对于关键词的提取可以采用自然语言处理技术实现,具体实现过程本实施例不作详述。

[0136] 本实施例在得到各访问对象对应的关键词簇后,可以生成如下表1所示的访问对象-关键词映射关系,以便后续直接查询。可见,上述步骤S302可以利用各历史访问数据,得到各访问对象的对象标识,之后,直接查询访问对象-关键词映射关系,获取相应访问对象包含的关键词,得到各访问对象对应的关键词簇。需要说明,访问对象-关键词映射关系的表示方式并不局限于下表1所示的形式。

[0137] 表1访问对象-关键词映射关系

[0138]

| 对象标识 | 关键词 |
|-------|---------------------|
| x_1 | Tag11, Tag12 |
| x_2 | Tag21, Tag22, Tag23 |
| x_3 | Tag31, Tag32 |

[0139] 步骤S303,由各访问对象对应的关键词簇,构成关键词序列;

[0140] 可选的,为了更好的表征访问对象的边界信息,本实施例可以在每个访问对象对应的关键词簇中,添加虚拟的关键词头部和尾部,分别记为head-tag和tail-tag,其中,在各访问对象对应的关键词簇中加入head-tag和tail-tag的内容可以相同,且其内容确定后可以固定不变,本实施例对该head-tag和tail-tag分别表示的内容不作限定。

[0141] 基于此,参照图7,仍以上述用户U2的访问对象序列为 $[x_1, x_2, x_3]$,及表1所示的访问对象-关键词映射关系为例进行说明,对应各访问对象的关键词簇分别为 $[Tag_{11}, Tag_{12}]$; $[Tag_{21}, Tag_{22}, Tag_{23}]$; $[Tag_{31}, Tag_{32}]$,每一个关键词簇可以理解为一篇文章的Tag序列,即每个关键词簇可以映射到至少一个访问对象,如图7所示的访问对象1、访问对象2等。本实施例可以按照各关键词簇对应访问对象的访问时间,将各访问对象的关键词簇作为序列元素,生成关键词序列,即 $[head-tag, Tag_{11}, Tag_{12}, tail-tag, head-tag, Tag_{21}, Tag_{22}, Tag_{23}, tail-tag, head-tag, Tag_{31}, Tag_{32}, tail-tag]$ 。

[0142] 由此可见,针对用户U2在应用平台的历史访问数据,得到的用来输入模型的序列由 $[x_1, x_2, x_3]$,变成了 $[head-tag, Tag_{11}, Tag_{12}, tail-tag, head-tag, Tag_{21}, Tag_{22}, Tag_{23}, tail-tag, head-tag, Tag_{31}, Tag_{32}, tail-tag]$ 。

[0143] 步骤S304,将关键词序列输入推荐对象预测模型进行编码计算,得到该用户访问对象的编码向量;

[0144] 本实施例可以获取多个样本用户对应的关键词序列,得到模型训练所用的训练数据,从而基于循环神经网络对这些训练数据进行模型训练,得到步骤S304所使用的推荐对象预测模型,模型训练过程可以参照上述实施例相应部分的描述。

[0145] 其中,将关键词序列输入推荐对象预测模型的编码处理过程,与上述对对象序列

的编码处理过程类似,区别在于输入由上述实施例的访问对象,变成了访问对象的关键词,本实施例对编码过程不做详述。因此,本实施例得到编码向量可以用来预测用户感兴趣的访问对象的关键词,并不能直接预测出用户感兴趣的访问对象。

[0146] 步骤S305,对该编码向量与各候选关键词对应的候选词向量进行相似度计算;

[0147] 与上述可选实施例的步骤S206不同的是,本实施例与编码向量进行相似度计算的是候选关键词对应的候选词向量,而不是候选访问对象对应的候选词向量,与此同时,本实施例中的候选集合中的元素不再是候选访问对象,而变成了候选关键词。

[0148] 其中,本实施例生成候选集合,选择的候选关键词可以是应用平台上访问次数较高的关键词、与社交网络热点主题相关的关键词等,本实施例对如何选择候选关键词的方式不做限定。而关于向量之间的相似度计算方法,与上述实施例描述的相似度计算方法类似,如余弦相似度计算方法,但并不局限于此,本实施例对步骤S305的具体实现过程不再赘述。

[0149] 步骤S306,获取相似度最高的第一数量个候选词向量对应的候选关键词;

[0150] 本实施例对第一数量的具体数值不作限定,其可以根据经验或试验设定。而且,在筛选候选关键词,也可以预先设置筛选的相似度标准,即预设的相似度阈值,那么,步骤S306可以变为获取相似度达到预设相似度阈值的候选词向量对应的候选关键词,当然还可以采用其他方式实现候选关键词的筛选,并不局限于本文给出的方式。

[0151] 步骤S307,由获取的第一数量个候选关键词,构成候选关键词簇;

[0152] 需要说明,该步骤S307是为了方便描述本实施例提供的对象推荐方法,所以,在本实施例实际应用中,可以直接利用筛选出的候选关键词进行后续步骤,并不一定要执行步骤S307,或者说在执行步骤S306过程中,隐含执行了步骤S307。

[0153] 步骤S308,利用候选关键词簇包含的一候选关键词,与推荐对象预测模型中上一时刻隐藏层输出作为当前时刻隐藏层输入,继续进行编码计算;

[0154] 本实施例可以将上述相似度计算过程认为解码过程,本实施例可以采用多次解码的方式,来获得多个候选关键词序列,用来得到用户的推荐对象。

[0155] 可选的,本实施例可以利用seq2seq技术实现多次解码。该seq2seq实际上是一个Encoder-Decoder(即编码器-解码器)结构的循环神经网络,其输入是一个序列,输出也是一个序列。其中,Encoder层的处理可以将一个可变长度的信号序列变为固定长度的向量表达,即将输入的输入编码成一个向量,Decoder层的处理可以将这个固定长度的向量变成可变长度的目标的信号序列,即结合编码生成的向量预测可能的输出对象,在本实施例中为通过用户的编码向量预测写一个序列,具体实现过程本实施例不做详述。

[0156] 基于此,本实施例得到候选关键词簇后,可以任意选择其中的一个候选关键词,或按照顺序选择一个候选关键词,或者选择相似度最高的候选关键词等,作为模型输入继续进行编码计算,即将选择的一候选关键词作为当前输入,利用推荐对象预测模型中上一时刻隐藏层输出,继续进行编码计算,得到新的编码向量,与上述实施例描述的将各序列元素依次输入推荐对象预测模型后,其隐藏层的计算方式类似,本实施例在此不再详述。

[0157] 步骤S309,将得到的新的编码向量作为用户访问对象的编码向量,返回步骤S305;

[0158] 步骤S310,检测相似度计算次数达到第二数量,利用构成的第二数量个候选关键词簇中同一维度候选关键词,生成第一数量个候选关键词序列;

[0159] 本实施例对多次编码次数不作限定,即对第二数量的数值不作限定,可以根据经验或试验设定第二数量的数值,也可以根据本次编码得到的候选关键词的相似度大小确定等等;且需要进行编码还可以通过判断构成的候选关键词簇是否满足预设条件,如候选关键词簇中的候选关键词的相似度是否达到一定阈值,或每次得到的候选关键词的相似度趋于稳定等条件实现,本实施例对如何确定是否继续进行编码的条件不做限定,并不局限于本实施例描述的方式。

[0160] 如上述分析,在本实施例实际应用中,假设每次解码计算,即每次通过相似度计算,筛选出相似度最高的K个候选关键词,即第一数量为K,按照上文描述方式实现T次解码后,即第二数量为T,将得到T个长度为K的候选关键词簇,之后,可以从这些关键词簇中,依次提取同一维度的候选关键词,生成相应的候选关键词序列,如将每次得到的K个候选关键词序列作为矩阵的一列,如此得到T列数据,取同一行的候选关键词构成一个预测得到的候选关键词序列,记为 $[Tagm1, Tagm2, Tagm3, \dots, TagmT]$, $m=1, 2, 3, \dots, K$,可见,本实施例可以得到K个长度为T的候选关键词序列。

[0161] 需要说明,本实施例的候选关键词序列生成方式并不局限于上文描述的同纬度候选关键词提取方式,也可以从每一个候选关键词簇中任意提取一个候选关键词生成,这样将得到大量候选关键词序列,后续其他步骤的处理方式相同,本实施例不再单独描述。

[0162] 步骤S311,基于第一数量个候选关键词序列包含的候选关键词,得到用户的推荐对象。

[0163] 可选的,本实施例可以预先建立应用平台能够输出的各对象的关键词到访问对象的倒排索引,具体构建方法本实施例不作详述。其中,倒排索引可以称为反向索引、置入档案或反向档案,可以被用来存储在全文搜索下,某个单词在一个文档或一组文档中的存储位置的映射,本实施例存储的则是关键词到访问对象的映射。

[0164] 基于此,在得到多个后续关键词序列,即 $[Tagm1, Tagm2, Tagm3, \dots, TagmT]$, $m=1, 2, 3, \dots, K$,可以在应用平台能够输出的访问对象中,依次拉取包含每一个候选关键词的候选访问对象,即获取各候选关键词映射到的至少一个候选访问对象,从而生成各候选关键词对应的倒排列表,如拉取包含 $Tagm1$ 的候选访问对象 $x2$,此时,该候选访问对象 $x2$ 可能还会包含其他候选关键词,也就是说,候选关键词可能会映射到同一个候选访问对象,之后,通过计算各候选访问对象被映射的次数,根据该次数大小,从映射到的候选访问对象中,筛选出用户的推荐对象,具体实现过程不做限定。

[0165] 当然,若候选关键词序列保留了词序,本实施例也可以将这些候选关键词序列按照词序进行拼接,之后,可以利用文本相似性计算方式,计算拼接后的候选关键词序列与各候选访问对象对应的关键词序列之间的相似度,进而选择相似度最高的预设数量个候选访问对象作为用户的推荐对象。

[0166] 可见,步骤S311的实现方式并不局限于某一种方式,来得到用户的推荐对象,可以采用上文给出的任一种方式实现,且并不局限于上文描述的两种实现方式。

[0167] 综上所述,本实施例利用用户访问对象中的关键词,构成模型的训练数据,基于循环神经网络训练得到的推荐对象预测模型,实现对用户访问对象的关键词序列的编码计算,得到预测用户感兴趣的关键词的编码向量,基于循环神经网络的特点,这种编码方式兼顾了用户的长期历史兴趣和短期历史兴趣,进而提高了最终向用户推送的目标推荐对象的

准确性,且由于考虑到了用户访问对象的访问顺序,能够更准确定位用户的兴趣变迁和兴趣积累,解决了现有Item CF的对象推荐方法导致召回结果多样性和个性化损失的问题。

[0168] 而且,由于本实施例在相似度计算的候选词向量是关键词的词向量,相对于访问对象的词向量,在数量上要少很多,且随着时间的变化,需要计算的候选词向量也不会有很大变化,减少了计算量,其提高了对象推荐的准确性及稳定性。

[0169] 可选的,对于上述可选实施例的步骤S311的实现方法提供了以下两种方式,但并不局限于下文给出的两种实现方式:

[0170] 方式一:

[0171] 参照图6及图8所示的流程示意图,该方法可以包括以下步骤:

[0172] 步骤A1,获取已构建的关键词映射到对象的倒排索引;

[0173] 本实施例对当前应用平台中的关键词映射到对象的倒排索引的构建方式不做限定。

[0174] 步骤A2,查询该倒排索引,得到各候选关键词序列中的每一个候选关键词的倒排列表;

[0175] 其中,该倒排列表用于表征该候选关键词映射到的至少一个候选推荐对象。如上文描述的方式,可以针对每一个候选关键词,从应用平台能够输出的访问对象中,拉回至少包含给候选关键词的候选访问对象,或者说是获取各候选关键词分别映射到的候选访问对象,从而生成各候选关键词的倒排列表,通过该倒排列表能够快速得到包含相应候选关键词的候选访问对象有哪些。

[0176] 步骤A3,基于得到的各候选关键词的倒排列表,统计各候选关键词序列中的各候选推荐对象被映射的次数;

[0177] 结合上述分析可知,对于本实施例拉回的候选访问对象,其包含的候选关键词(即上述得到的各候选关键词序列中的候选关键词)数量越多,即该候选访问对象被拉回的次数越多,说明该候选访问对象越可能被推荐给用户,即成为推荐对象的概率越大。所以,本实施例可以基于各候选访问对象被拉回的次数,按照一定规则,确定该候选访问对象对应的分值,之后,根据各候选访问对象的分值大小,来实现后续步骤。在本实施例中,候选对象的分值越大,说明用户对该候选访问对象感兴趣的概率越大,将其作为推荐对象的概率越大。

[0178] 可选的,在实际应用中,本实施例还可以统计各候选访问对象对各候选关键词序列的关键词覆盖率,该关键词覆盖率越大,说明相应候选访问对象成为推荐对象的概率越大,本实施例可以在此基础上直接进行后续处理,也可以据悉得到各后续访问对象的分值,由分值大小来表示相应候选访问对象成为推荐对象的概率等等,本实施例对该关键词覆盖率的具体计算方法不作限定。

[0179] 步骤A4,基于统计结果,从获取的多个候选推荐对象中,筛选所述用户的推荐对象。

[0180] 在本实施例中,由于被映射次数越多的候选推荐对象被筛选为推荐对象的概率越大,也就是说,候选访问对象对应的候选关键词数量越多,说明其被推荐给用户的概率越大,因此,本实施例可以筛选对应候选关键词数量最多的p个候选对象作为用户的推荐对象。

[0181] 具体的,本实施例可以对各候选访问对象被映射的次数进行排序,按照被映射次数从大到小的顺序,选择p个候选访问对象作为推荐对象,但并不局限于这种实现方式。

[0182] 方式二:

[0183] 参照图6及图9所示的流程示意图,该方法可以包括以下步骤:

[0184] 步骤B1,按照各候选关键词序列中各候选关键词的词序,对第一数量个候选关键词序列包含的候选关键词进行拼接;

[0185] 步骤B2,获取多个候选访问对象分别对应的关键词序列,所述候选访问对象至少包含任一候选关键词序列中的一候选关键词;

[0186] 需要说明,在本实施例实际应用中,步骤B2可以在步骤B1之前执行,并不局限于本实施例这种步骤顺序。

[0187] 可选的,当候选访问对象的关键词序列过多时,为了减少相似度计算工作量,本实施例可以采用倒列表方式对候选访问对象进行过滤,如筛选出至少包含一个候选关键词序列中的候选关键词的候选访问对象,用来完成后续的相似度计算。

[0188] 步骤B3,对拼接得到的候选关键词序列与各候选访问对象对应的关键词序列进行文本相似度计算;

[0189] 在本实施例实际应用中,对于拼接后得到的关键词序列可以认为是一个文本,因此,本实施例可以采用文本相似度计算方式,计算候选关键词序列与各候选访问对象对应的关键词序列之间的相似度,以筛选出推荐对象。本实施例对文本相似度计算方法的具体实现过程不作详述。

[0190] 步骤B4,基于文本相似度计算结果,从多个候选访问对象中,筛选用户的推荐对象。

[0191] 其中,所述相似度越高的候选对象被筛选为推荐对象的概率越大,因此,本实施例可以从候选访问对象中,筛选相似度最高的预设数量个候选访问对象为推荐对象,也可以筛选相似度达到预设阈值的候选访问对象为推荐对象等等,本实施例对步骤B4的具体实现过程不作限定。

[0192] 基于上述各实施例得到的用户的推荐对象,在实际应用中,可以将其作为目标推荐对象的候选项,也就是说,在得到用户的推荐对象后,还可以利用一些逻辑对其做进一步筛选,而上述各实施例描述的对象推荐方法的步骤在实际应用中的作用可以是:依据特定用户的画像信息,按照各种精准个性化、泛个性化、热度等维度进行数据拉取,即拉取用户可能感兴趣的推荐对象,通常此时得到的推荐对象的数量比较多,可以进一步筛选,具体实现方法可以参照图10所示的流程示意图,但并不局限于图10所示的方法。

[0193] 如图10所示,为本发明实施例提供的又一对象推荐方法的流程示意图,该方法中获取用户的推荐对象实现过程,可以参照上述各实施例的描述,本实施例在此不再赘述,在此仅对得到用户的推荐对象之后的处理过程进行描述,因此,该方法还可以包括以下步骤:

[0194] 步骤S401,按照特定规则,对用户的推荐对象进行初步筛选,得到初选推荐对象;

[0195] 其中,特定规则可以是用户访问对象相关性、时效性、地域、多样性等方面的因素确定,本实施例对该特定规则包含的具体内容不做限定。

[0196] 如图11所示的应用场景的推荐系统的示意图,该应用场景具体可以是某即时通信客户端中信息展示平台对其输出信息的推荐场景,如图11所示,该推荐系统可以包括召回

逻辑、初选逻辑及排序逻辑(即图11中的Rank)等多个功能模型,其中,召回逻辑实现的是上述各实施例描述的用户推荐对象的获取过程,初选逻辑用来实现步骤S401的实现过程,而排序逻辑用来实现对初选推荐对象的后续排序处理过程。

[0197] 步骤S402,获取各初选推荐对象对应的编码向量,及当前访问对象的编码向量;

[0198] 在本实施例实际应用中,除了直接计算初选推荐对象与当前浏览对象之间的相似度,按照相似度大小对各初选推荐对象进行排序的方式外,本实施例还可以采用上文描述的相似度计算方式,即得到访问对象对应的编码向量,从而通过向量之间的相似度,实现对各初选推荐对象的排序。

[0199] 基于此,本实施例可以按照上文描述的方式,得到各初选推荐对象对应的编码向量,如将各初选推荐对象作为序列元素,生成访问对象序列,将其依次输入推荐对象预测模型进行编码计算,得到相应的编码向量,但并不局限于这一种编码方式,还可以采用上述关键词序列的编码方式,具体实现过程可以参照上文各实施例描述的编码向量的生成步骤。

[0200] 步骤S403,对各初选推荐对象对应的编码向量与当前访问对象的编码向量进行相似度计算;

[0201] 可选的,本实施例可以采用余弦相似度计算方法,得到向量之间的相似度,具体实现过程可以参照上文实施例相应部分的描述,但向量之间的相似度计算方法并不局限于这种实现方法。

[0202] 步骤S404,选择相似度最高的预设数量个初选推荐对象作为目标推荐对象;

[0203] 需要说明,本实施例从多个初选推荐对象中,筛选目标推荐对象的实现方法并不局限于本实施例描述的方式,与上述实施例描述的从多个候选访问对象,筛选用户的推荐对象的方法类型,本实施例在此不作一一列举。

[0204] 步骤S405,将该目标推荐对象发送至用户的客户端进行展示。

[0205] 综上所述,本实施例按照上述实施例描述的方式,得到的用户的推荐对象不仅能够准确定位出用户的兴趣变迁和兴趣积累,还能够兼顾用户的长期兴趣和短期兴趣,从这样的推荐对象中筛选出的用户的目标推荐对象能够更加符合用户的当前需求,使得采用本实施例这种对象推荐方法的应用服务器,能够更好地为用户提供推荐服务。

[0206] 以上述图11所示的应用推荐系统为例,对上述实施例提供的对象推荐方法进行场景下的具体说明,本实施例仅以访问对象为应用平台输出的文章,且基于对象序列作为训练数据为例进行说明,对于该应用的服务器来说,其可以获取该应用平台上多个用户的文章序列,并由此得到训练模型,基于循环神经网络进行模型训练,得到推荐对象预测模型,之后,服务器可以针对使用该应用的任一用户,可以记为目标用户,可以按照上述方式获取目标用户的文章序列,将其依次输入该推荐对象预测模型,得到用来预测用户可能感兴趣的文章的编码向量,并将其与预设的多个候选文章的词向量进行相似度计算,从而选择相似度最高的若干个候选文章作为目标用户的推荐文章,即图11执行召回逻辑的召回结果。当然,在模型训练过程中,也可以使用多个用户对应的关键词序列,构成训练数据,这样,在预测目标用户的推荐文章时,获取的则是目标用户的关键词序列,将其输入推荐对象预测模型后,输出的编码向量可以预测用户可能感兴趣的关键词,之后,分别计算该编码向量与各候选关键词的词向量之间的相似度(可以认为是一次解码过程),选择K个相似度最高的候选关键词,经多次解码后,生成多个长度为T的关键词序列,之后,按照上文给出两种文字

召回逻辑的处理方式,如倒排索引或文本相似度计算方式,得到目标用户的推荐文章,即该目标用户的文章召回结果。

[0207] 之后,可以利用初选逻辑对召回的多个推荐文章进行筛选,得到多个初选文章,再利用Rank逻辑对多个初选文章进行排序,此时也可以利用上文描述的获取推荐文章的方式,对初选文章进行排序,得到目标推荐文章,并将其展示在目标用户使用客户端显示界面上。

[0208] 在实际应用中,上述处理过程可以由服务器在线下实现,当某用户使用客户端连通服务器时,服务器可以根据该用户的用户标识,直接将该用户标识关联的目标推荐文章反馈至客户端展示,但并不局限于这种实现方式。

[0209] 参照图12,为本实施例提供一种对象推荐装置的结构示意图,该装置可以应用于服务器中,该装置可以包括但并不局限于以下组成结构:

[0210] 序列获取模块11,用于获取用户访问序列;

[0211] 其中,该用户访问序列是基于用户访问应用平台输出的对象生成的;

[0212] 编码计算模型12,用于将所述用户访问序列输入推荐对象预测模型进行编码计算,得到所述用户访问对象的编码向量;

[0213] 其中,该推荐对象预测模型是基于循环神经网络,对多个样本用户对应的用户访问序列训练得到,且该循环神经网络包括有多个门控循环单元层,或多个长短期记忆网络层,具体网络结构和原理可以参照上述方法实施例相应部分的描述。

[0214] 第一相似度计算模块13,用于对所述编码向量与各候选词向量进行相似度计算;

[0215] 推荐对象选择模块14,用于基于相似度计算结果,得到所述用户的推荐对象。

[0216] 可选的,参照图13所示,训练推荐对象预测模型所需的训练数据内容不同,获取的用户访问序列的序列元素内容将不同,进而得到用户推荐对象的方式也会相应改变。

[0217] 基于此,上述序列获取模块11可以包括:

[0218] 第一数据获取单元1110,用于获取用户的多条历史访问数据,所述历史访问数据是基于用户对应用平台输出对象的访问操作产生的;

[0219] 第一序列构成单元1111,用于由所述多条历史访问数据分别包含的对象标识,构成访问对象序列。

[0220] 在这种情况下,如图13所示,本实施例提供的装置还可以包括:

[0221] 候选访问对象获取模块15,用于获取多个候选访问对象;

[0222] 第一词向量获取模块16,用于将各候选访问对象输入语言模型,得到相应的候选词向量

[0223] 相应地,上述第一相似度计算模块13具体可以用于对所述编码向量与候选访问对象对应的候选词向量进行相似度计算。

[0224] 其中,关于向量之间的相似度计算方法可以参照上文方法实施例相应部分的描述。

[0225] 推荐对象选择模块14具体可以用于选择相似度最高的预设数量个候选访问对象,作为用户的推荐对象;或者选择相似度达到预设阈值的候选访问对象作为用户的推荐对象等等,本实施例对如何利用相似度,从多个候选访问对象中,筛选用户的推荐对象的方式不做限定。

- [0226] 作为本申请另一实施例,如图14所示,上述序列获取模块11可以包括:
- [0227] 第二数据获取单元1120,用于获取用户的多条历史访问数据;
- [0228] 关键词簇获取单元1121,用于利用各历史访问数据,得到相应访问对象对应的关键词簇;
- [0229] 第二序列构成单元1122,用于由各访问对象对应的关键词簇,构成关键词序列。
- [0230] 本实施例中,获取关键词序列的实现过程可以参照上文方法实施例相应部分的描述。
- [0231] 此时,上述装置还可以包括:
- [0232] 候选关键词获取模块,用于获取多个候选关键词;
- [0233] 第二词向量获取模块,用于将所述多个候选关键词分别输入语言模型,得到相应的候选词向量。
- [0234] 相应地,上述第一相似度计算模块13具体可以用于对所述编码向量与各候选访问关键词对应的候选词向量进行相似度计算。
- [0235] 在实际应用中,由于应用平台输出的访问对象的关键词数量是相对稳定的,并不会随着应用平台的用户及输出对象来源的增加,而急速增加,这使得本实施例相似度计算的计算量比较小,且由此得到的用户的推荐对象也相对稳定。
- [0236] 可选的,在候选集合为候选关键词的情况下,如图14所述,上述推荐对象选择模块14可以包括:
- [0237] 候选关键词获取单元141,用于获取相似度最高的第一数量个候选词向量对应的候选关键词;
- [0238] 关键词簇生成单元142,用于由获取的第一数量个候选关键词,构成候选关键词簇;
- [0239] 编码计算单元143,用于利用所述候选关键词簇包含的一候选关键词,及所述推荐对象预测模型中上一时刻隐藏层输出继续进行编码计算;
- [0240] 相似度计算终止单元144,用于将得到的新的编码向量作为用户访问对象的编码向量,执行所述对所述编码向量与各候选词向量进行相似性计算步骤,直至相似度计算次数达到第二数量,或构成的候选关键词簇满足预设条件;
- [0241] 关键词序列生成单元145,用于利用构成的第二数量个候选关键词簇中同一维度候选关键词,生成第一数量个候选关键词序列;
- [0242] 推荐对象获取单元146,用于基于所述第一数量个候选关键词序列包含的候选关键词,得到所述用户的推荐对象。
- [0243] 可选的,上述推荐对象获取单元146具体可以包括:
- [0244] 倒排索引获取子单元,用于获取已构建的关键词映射到对象的倒排索引;
- [0245] 查询子单元,用于查询所述倒排索引,得到各候选关键词序列中的每一个候选关键词的倒排列表;
- [0246] 其中,该倒排列表用于表征该候选关键词映射到的至少一个候选推荐对象。
- [0247] 统计子单元,用于基于得到的各候选关键词的倒排列表,统计各候选关键词序列中的各候选推荐对象被映射的次数;
- [0248] 第一筛选子单元,用于基于统计结果,从获取的多个候选推荐对象中,筛选所述用

户的推荐对象。

[0249] 其中,被映射次数越多的候选推荐对象被筛选为推荐对象的概率越大。

[0250] 作为另一可选实施例,上述推荐对象获取单元146也可以包括:

[0251] 拼接子单元,用于按照各候选关键词序列中各候选关键词的词序,对所述第一数量个候选关键词序列包含的候选关键词进行拼接;

[0252] 关键词序列获取子单元,用于获取多个候选访问对象分别对应的关键词序列,所述候选访问对象至少包含任一候选关键词序列中的一候选关键词;

[0253] 相似度计算子单元,用于对拼接得到的候选关键词序列与各候选访问对象对应的关键词序列进行文本相似度计算;

[0254] 第二筛选子单元,用于基于文本相似度计算结果,从所述多个候选访问对象中,筛选所述用户的推荐对象。

[0255] 其中,所述相似度越高的候选访问对象被筛选为推荐对象的概率越大。

[0256] 需要说明,关于上述装置实施例中各功能模块或单元的功能实现过程,可以参照上述方法实施例相应部分的描述。

[0257] 可选的,在上述各实施例的基础上,如图15所示,上述装置还可以包括:

[0258] 初步筛选模块17,用于按照特定规则,对所述用户的推荐对象进行初步筛选,得到初选推荐对象;

[0259] 编码向量获取模块18,用于获取各初选推荐对象对应的编码向量,及当前访问对象的编码向量;

[0260] 第二相似度计算模块19,用于对各初选推荐对象对应的编码向量与所述当前访问对象的编码向量进行相似度计算;

[0261] 目标推荐对象选择模块120,用于选择相似度最高的预设数量个初选推荐对象作为目标推荐对象;

[0262] 目标推荐对象发送模块121,用于将所述目标推荐对象发送至所述用户的客户端进行展示。

[0263] 综合上述,本实施例基于循环神经网络,对多个样本用户对应的用户访问序列进行训练,得到推荐对象预测模型,实现对用户访问序列进行编码计算,不仅兼顾到用户的长期历史兴趣和短期历史兴趣,且考虑到了用户的应用平台上访问对象的访问顺序,这样,基于得到的编码向量与各候选词向量的相似度计算结果,得到该用户的推荐对象能够准确定位用户的兴趣变迁和兴趣积累,解决了现有Item CF推荐方法导致所得推荐对象多样性和个性化损失的问题。

[0264] 其中,根据不同需要,可以灵活选择进行模型训练所使用的训练数据的内容,即用户访问序列的内容,提高了对象推荐方法的灵活性。

[0265] 本发明实施例还提供一种计算机设备,该计算机设备的硬件结构可以如图16,该计算机设备的硬件结构可以包括:通信接口1、存储器2和处理器3;

[0266] 在本发明实施例中,通信接口1、存储器2、处理器3可以通过通信总线实现相互间的通信,且该通信接口1、存储器2、处理器3及通信总线的数量可以为至少一个。

[0267] 可选的,通信接口1可以为通信模块的接口,如GSM模块的接口;

[0268] 处理器3可能是一个中央处理器CPU,或者是特定集成电路ASIC(Application

Specific Integrated Circuit), 或者是被配置成实施本发明实施例的一个或多个集成电路。

[0269] 存储器2可能包含高速RAM存储器, 也可能还包括非易失性存储器(non-volatile memory), 例如至少一个磁盘存储器。

[0270] 其中, 存储器2存储有计算机程序, 处理器3调用存储器2所存储的计算机程序, 以实现上述应用于计算机设备的对象推荐方法的各步骤;

[0271] 可选的, 该计算机程序主要可用于:

[0272] 获取用户访问序列, 所述用户访问序列是基于用户访问应用平台输出的对象生成的;

[0273] 将所述用户访问序列输入推荐对象预测模型进行编码计算, 得到所述用户访问对象的编码向量, 所述推荐对象预测模型是基于循环神经网络, 对多个样本用户对应的用户访问序列训练得到;

[0274] 对所述编码向量与各候选词向量进行相似度计算;

[0275] 基于相似度计算结果, 得到所述用户的推荐对象。

[0276] 在本实施例实际应用中, 上述计算机设备可以是应用服务器, 如各种即时通信客户端对应的应用服务器等。

[0277] 本发明实施例还提供一种存储介质, 该存储介质记录有适于计算机设备的处理器执行的计算机程序, 以实现上述应用于计算机设备的对象推荐方法的各步骤, 该对象推荐方法的实现过程可以参照上述方法实施例相应部分的描述。

[0278] 本说明书中各个实施例采用递进的方式描述, 每个实施例重点说明的都是与其他实施例的不同之处, 各个实施例之间相同相似部分互相参见即可。对于实施例公开的装置、计算机设备而言, 由于其与实施例公开的方法相对应, 所以描述的比较简单, 相关之处参见方法部分说明即可。

[0279] 专业人员还可以进一步意识到, 结合本文中所公开的实施例描述的各示例的单元及算法步骤, 能够以电子硬件、计算机软件或者二者的结合来实现, 为了清楚地说明硬件和软件的可互换性, 在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行, 取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能, 但是这种实现不应认为超出本发明的范围。

[0280] 结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块, 或者二者的结合来实施。软件模块可以置于随机存储器(RAM)、内存、只读存储器(ROM)、电可编程ROM、电可擦除可编程ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的任意其它形式的存储介质中。

[0281] 对所公开的实施例的上述说明, 使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的, 本文中所定义的一般原理可以在不脱离本发明的核心思想或范围的情况下, 在其它实施例中实现。因此, 本发明将不会被限制于本文所示的这些实施例, 而是要符合与本文所公开的原理和和特点相一致的最宽的范围。

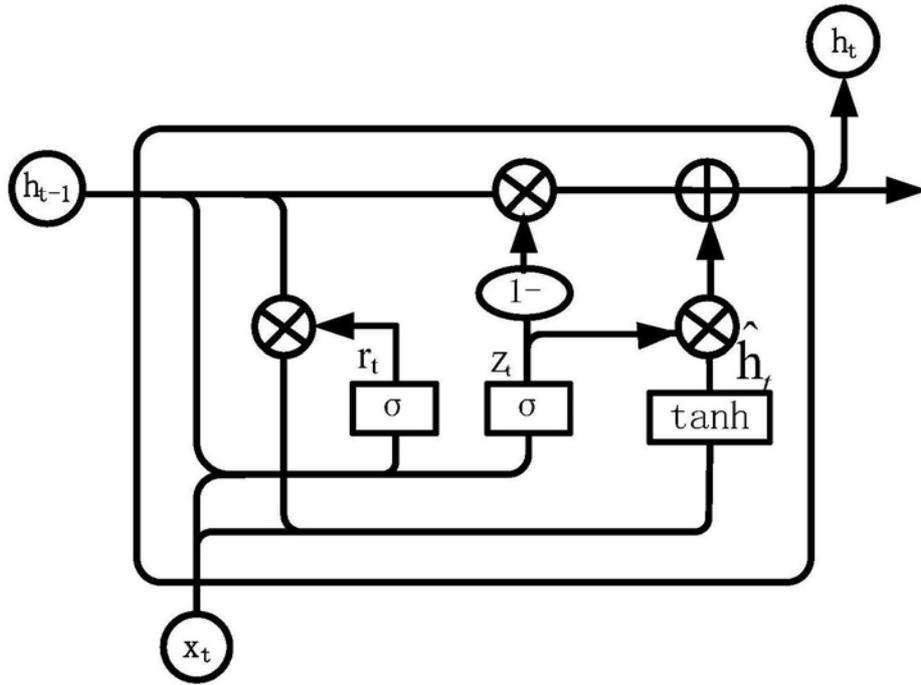


图1

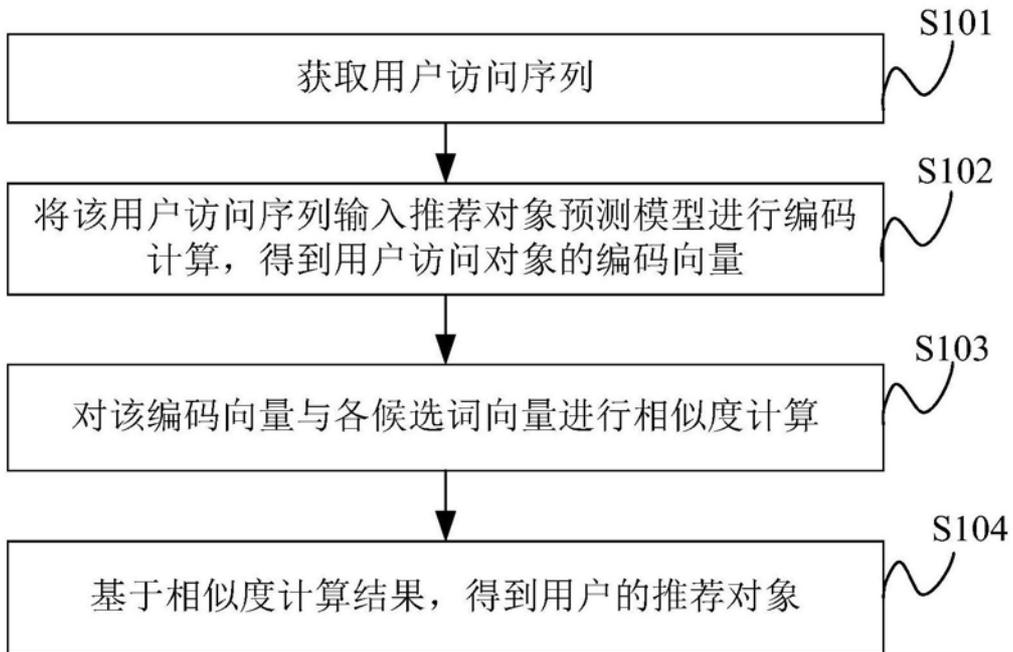


图2

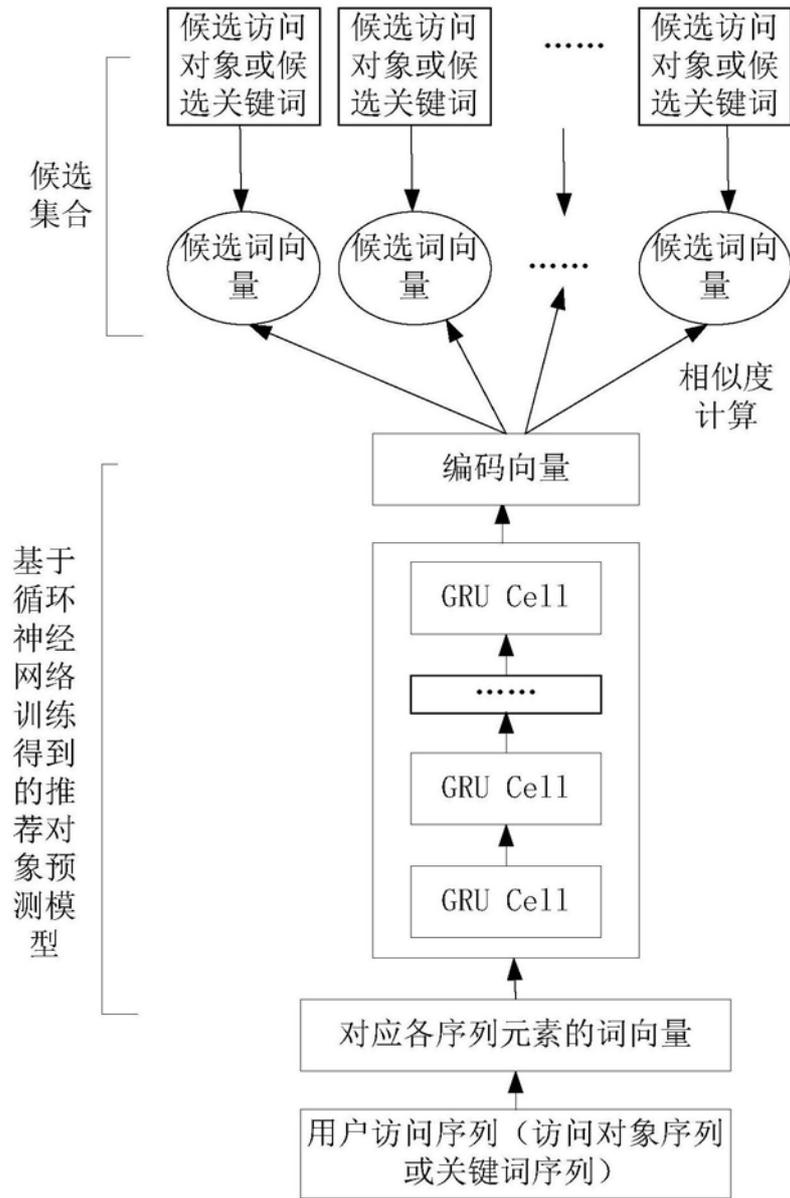


图3

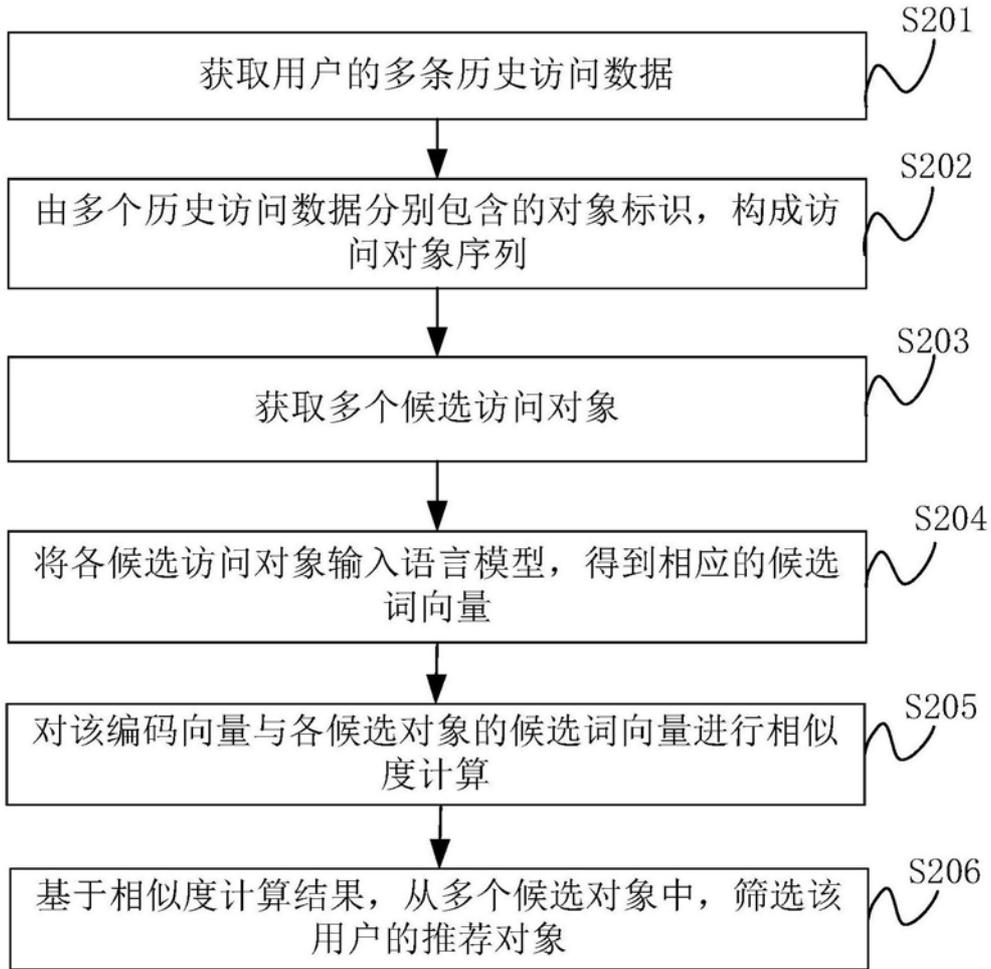


图4

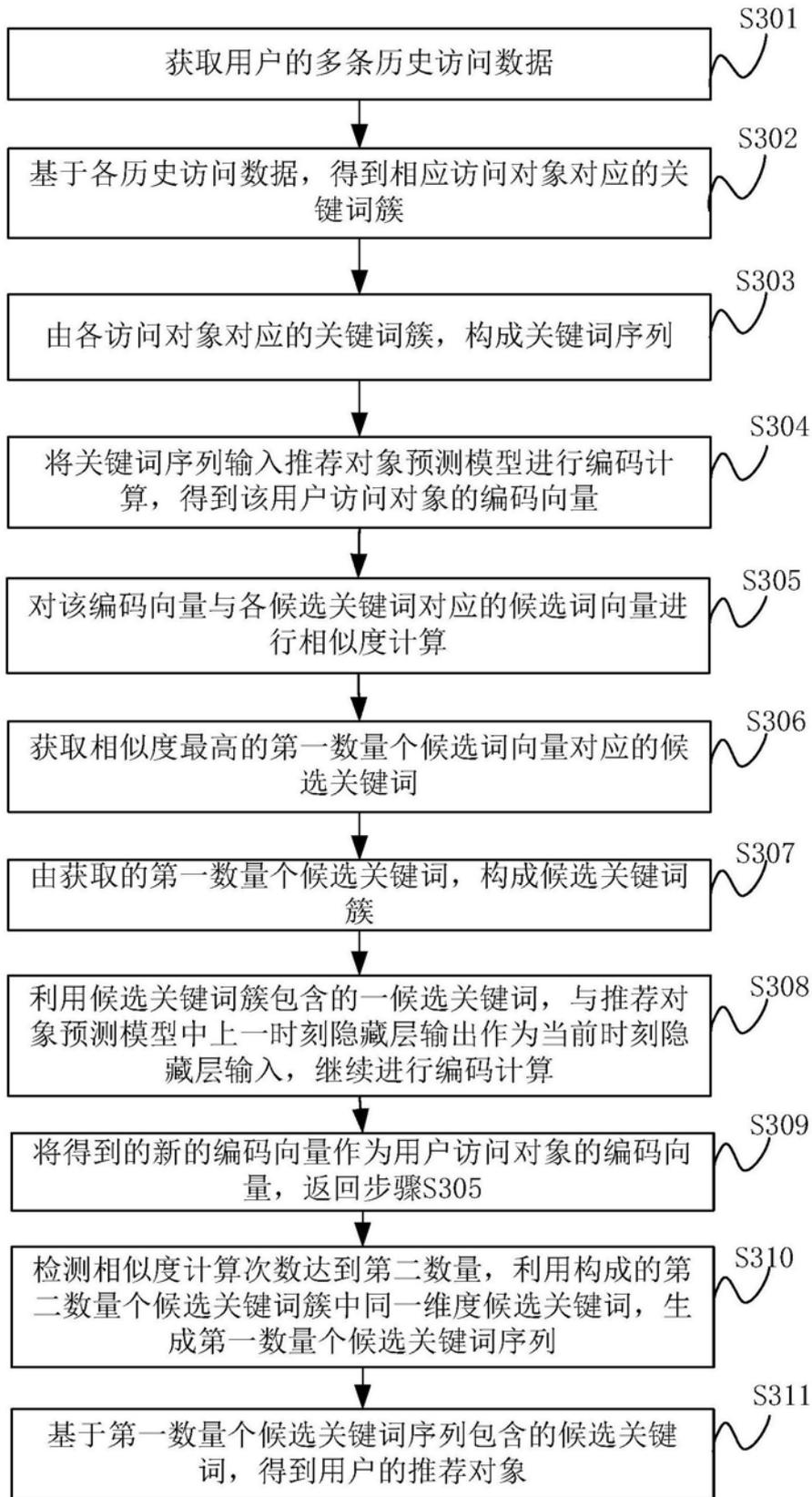


图5

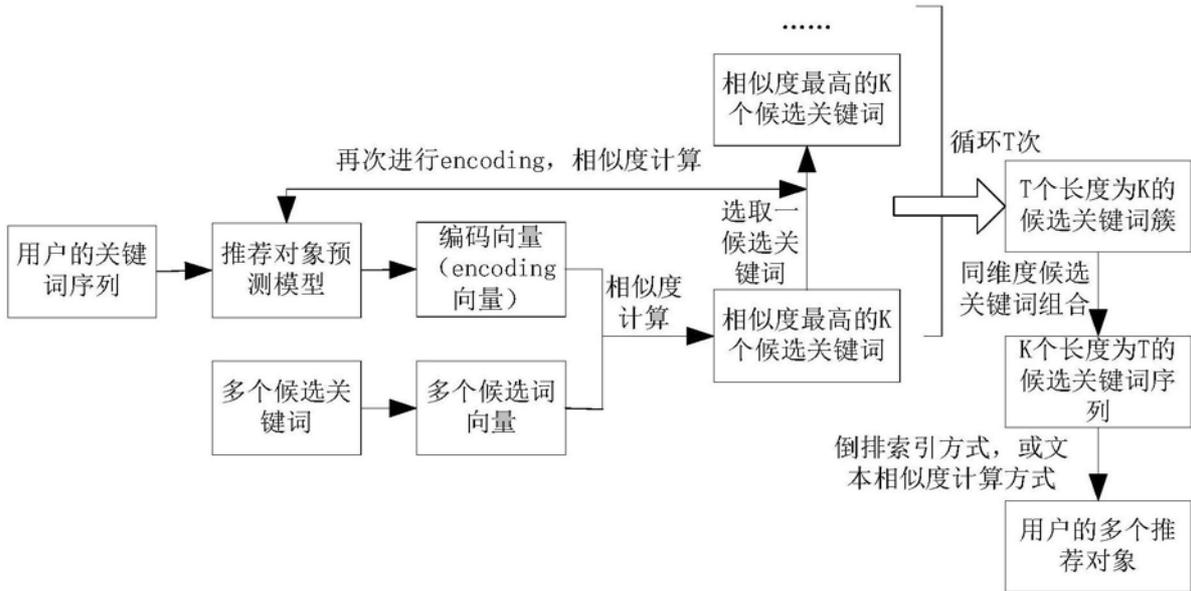


图6

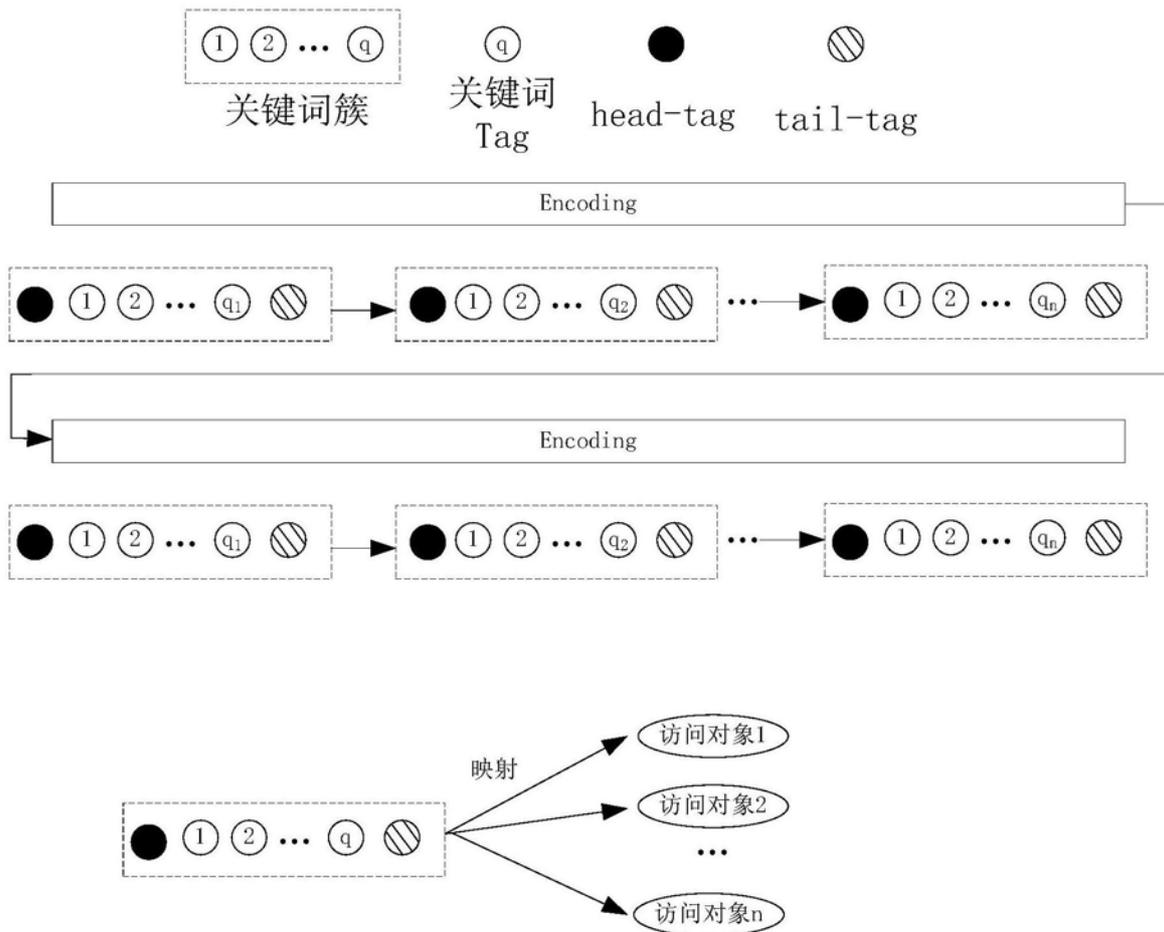


图7

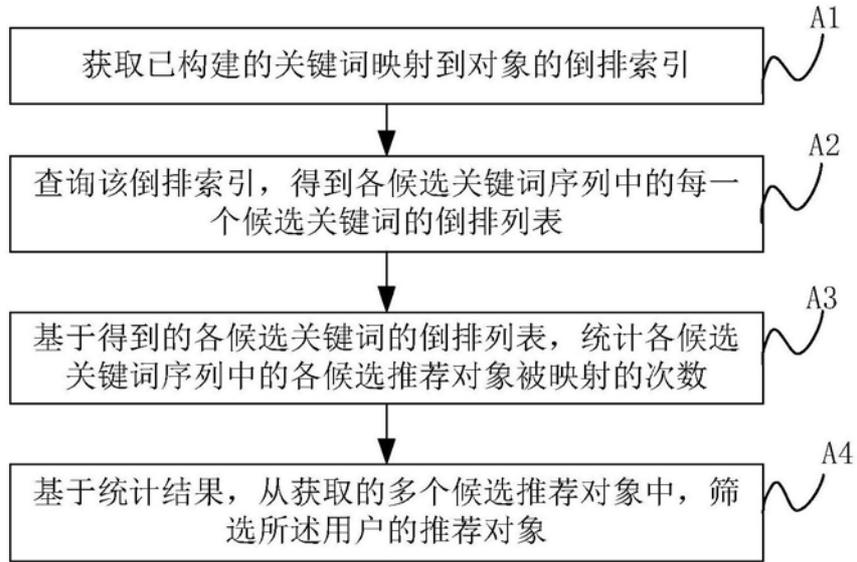


图8

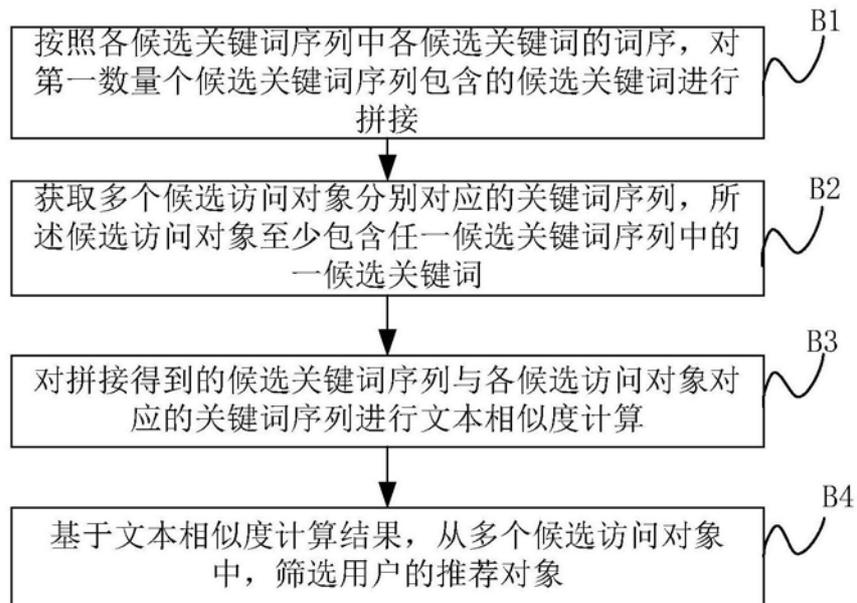


图9

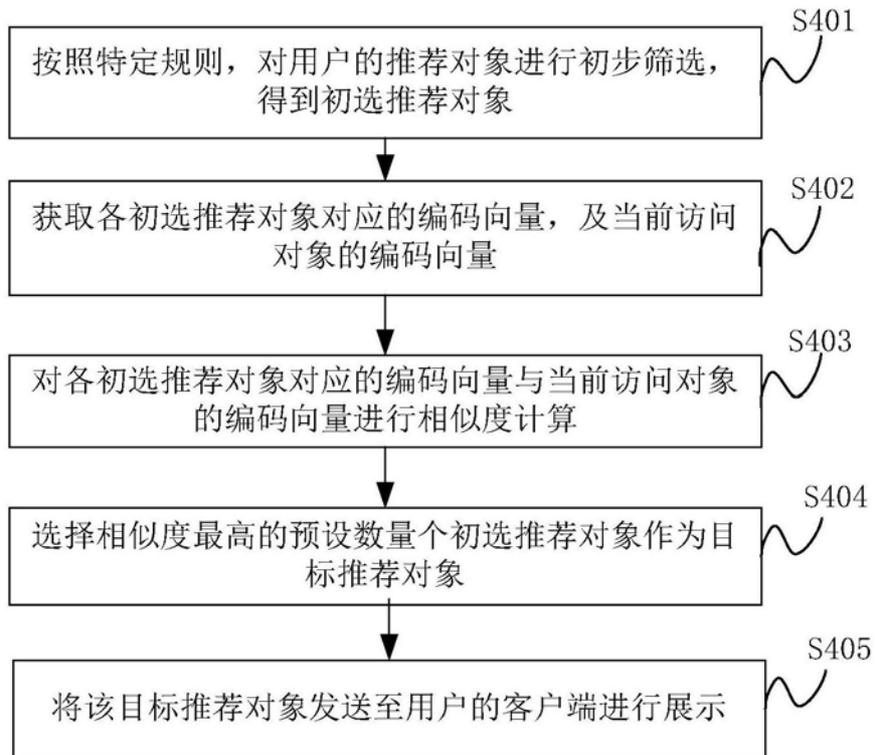


图10

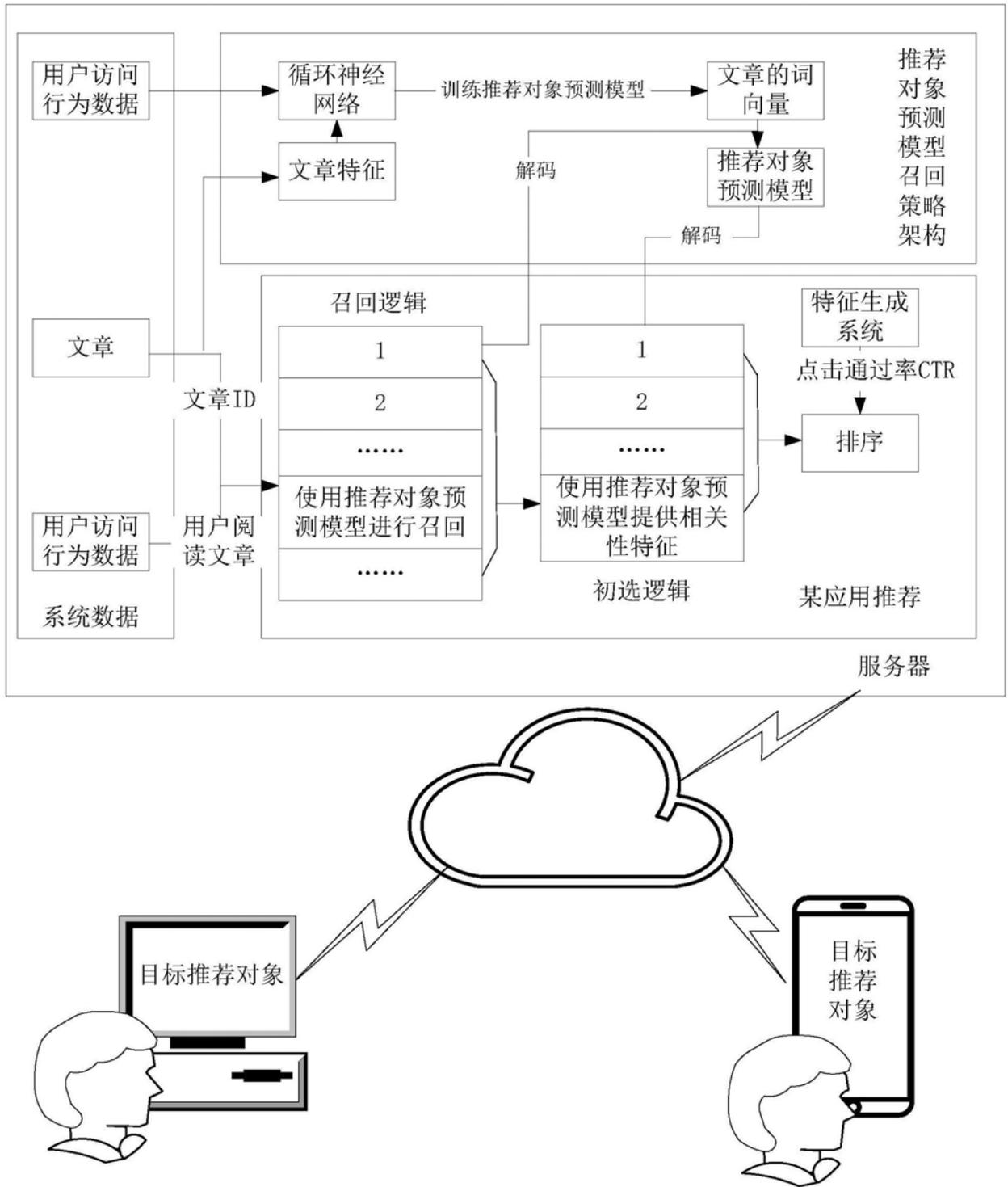


图11

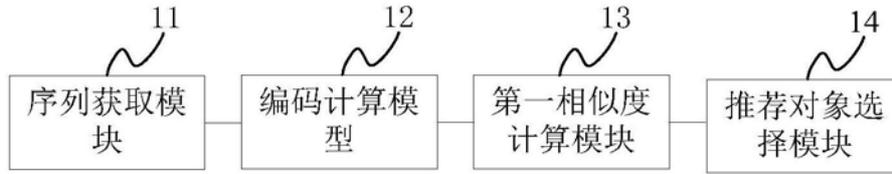


图12

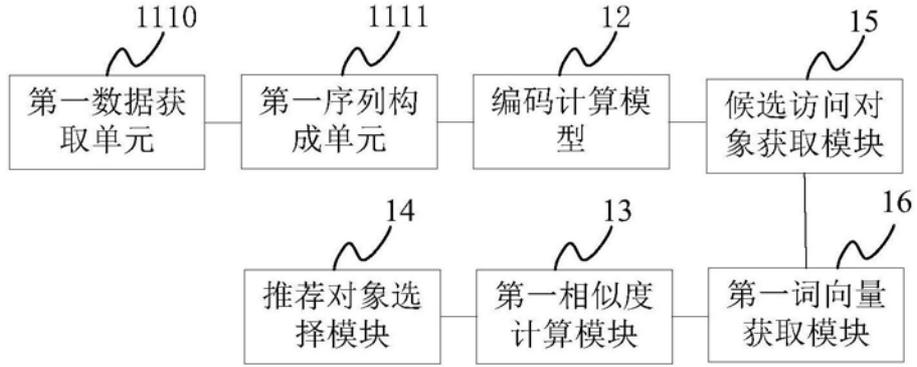


图13

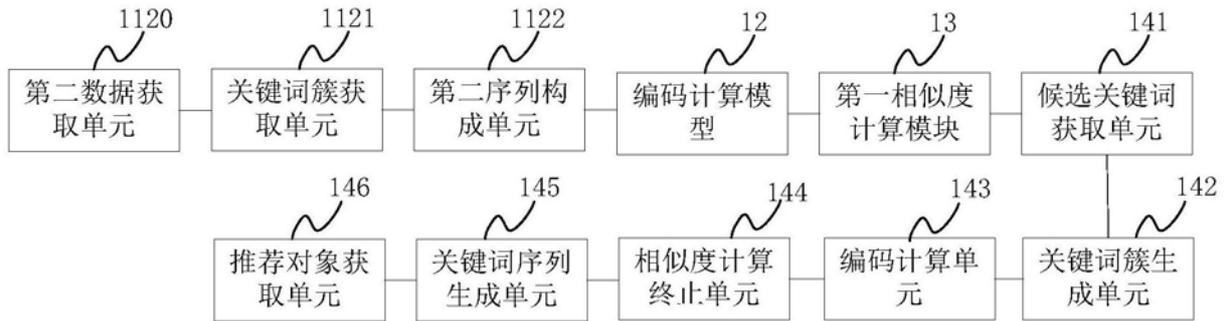


图14

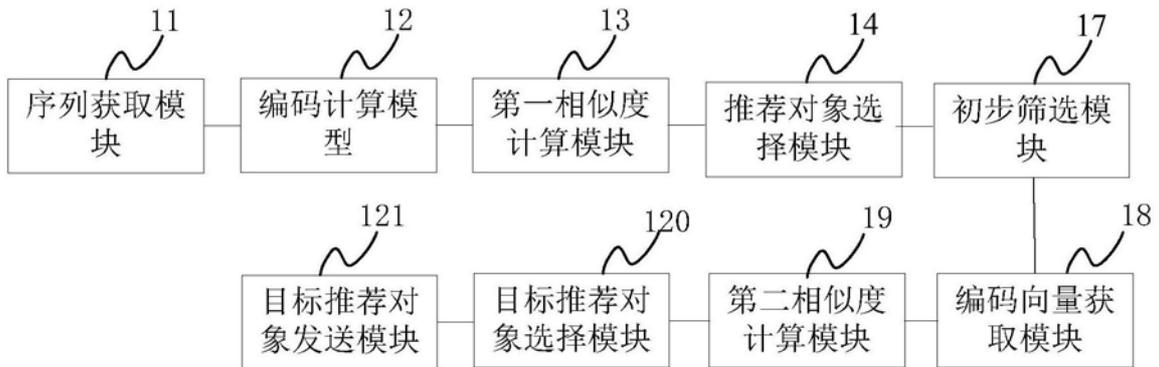


图15

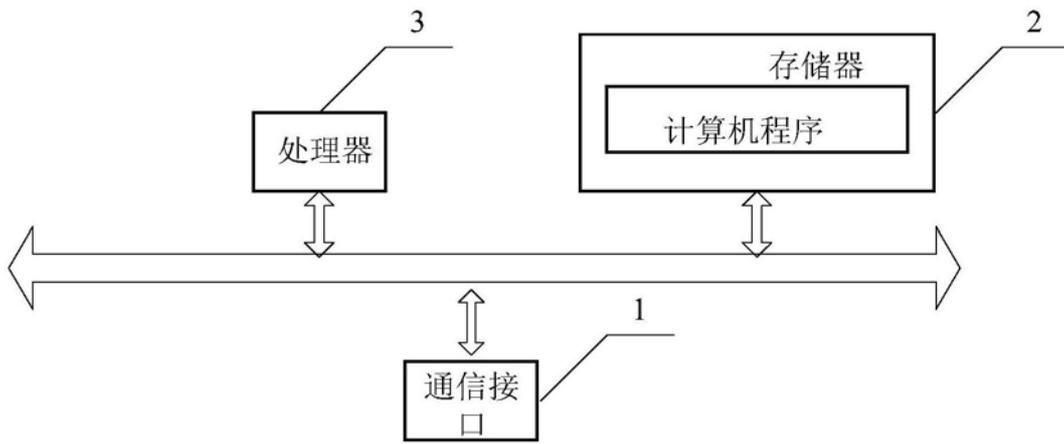


图16