

# 基于 DNN 算法的移动视频推荐策略

陈 亮 汪景福 王 娜 李 霞

(深圳大学信息工程学院 广东 深圳 518060)

**摘 要** 该文针对移动互联网在线视频服务个性化视频推荐问题,提出了一种基于深度学习(Deep learning)模型的内容推荐策略.深度学习在图像、语音和自然语言处理等领域获得的突破性进展,为在线视频服务的推荐策略研究提供了基础.该文工作通过传统基于内容推荐的基础上,引入深度神经网络(Deep Neural Network, DNN)词向量方法,根据视频的媒资信息和用户的历史行为,将视频特征和用户特征映射在高维向量空间.在构建用户正负行为与视频向量的余弦距离模型基础上,筛选过滤对所推荐视频感兴趣的用户群,并通过移动互联网应用的消息推送功能提示移动终端用户观看所推荐的内容.基于在大规模移动视频服务系统中的离线和在线实验,该文所提出的基于 DNN 算法的推荐策略,相比随机方法、ContentKNN 以及 ItemCF 等算法,在点击率方面平均分别获得 106%、41% 和 57% 的相对提升,在覆盖率方面一定程度上避免了推送活跃用户的偏颇问题,从整体上得到了较好的推荐效果.

**关键词** 移动互联网;深度学习;词向量;消息推送;基于内容推荐;机器学习

**中图法分类号** TP393 **DOI 号** 10.11897/SP.J.1016.2016.01626

## DNN Based Mobile Online Video Recommendation

CHEN Liang WANG Jing-Fu WANG Na LI Xia

(College of Information and Engineering, Shenzhen University, Shenzhen, Guangdong 518060)

**Abstract** We study the personalized video recommendation strategy in mobile online video services by introducing the deep learning method in the content-based recommendation. Deep learning algorithms have made a good progress in image recognition, speech recognition and natural language processing, which lay the foundation of the recommendation study in online video services. By mapping each video (according to its description information) and each user (according to her/his history behavior) into a high dimensional vector space with word embedding methods, we can filter the target user shaving potential interest in candidate videos based on the vector relationship model, and push notifications of recommended videos to them accordingly. With offline and online experiments in a large-scale mobile online video system, we evaluate our proposal and compare it with the random method, contentKNN algorithm and ItemCF algorithm. The experiment results show that our proposal's performance has been improved by as much as 106%, 41% and 57% respectively on the click-through rate in the mobile online video services.

**Keywords** mobile Internet; deep learning; word embedding; push notification; content-based recommendation; machine learning

收稿日期: 2015-02-09; 在线出版日期: 2015-10-14. 本课题得到国家自然科学基金(61502315, 61309030)、广东省自然科学基金(2015A030310366)、深圳大学科研启动项目(201558)资助. 陈 亮, 男, 1983 年生, 博士, 讲师, 中国计算机学会(CCF)会员, 主要研究方向为网络服务分析、数据科学与数据工程. E-mail: lchen@szu.edu.cn. 汪景福, 男, 1989 年生, 硕士研究生, 主要研究方向为数据分析、机器学习. 王 娜, 女, 1977 年生, 博士, 教授, 主要研究领域为机器学习、模式识别和算法优化. 李 霞(通信作者), 女, 1968 年生, 博士, 教授, 主要研究领域为机器学习、智能计算和信号处理. E-mail: lixia@cuhk.edu.cn.

## 1 引 言

视频推荐<sup>[1-4]</sup>是在线视频服务重要且基础的组成部分。国外的 Netflix、YouTube 等, 以及国内的优酷土豆、腾讯视频等在线视频服务提供商均投入大量资源探索这一领域, 因为优质的视频推荐可以提升用户使用在线视频服务的时长和增强用户的忠诚度。在这一领域, 研究人员提出各类推荐算法, 如协同过滤 (Collaborative Filtering, CF)、基于内容的过滤 (Content-Based Filtering, CBF)、基于社交关系链的过滤以及混合模型等<sup>[5-12]</sup>。在实际工业界, 在线视频推荐服务常采用 CBF+X 的混合模型, 如 YouTube 使用 CBF+ItemCF 的方案<sup>[13]</sup>。因为 CBF 有助于解决推荐新内容的冷启动问题, 所以它成为在线视频推荐服务不可或缺的一部分。

互联网信息的快速更新 (如新上线的商品、新闻、视频等) 突出了推荐系统中的冷启动问题。如何有效解决该问题, 成为当前研究的一个关注点。目前, 普遍采用混合算法或改进 CBF 方法<sup>[14-20]</sup>来应对冷启动问题。在以往应用 CBF 算法时, 通常会基于用户的注册信息、物品的文本描述关键词以及人工标签等来建立对用户和物品的特征描述<sup>[21]</sup>。然而, 这样的做法往往不能深入地把握用户和物品的特征, 从而导致推荐失效。为了提高 CBF 的效果, 研究人员考虑了通过对内容的进一步理解着手, 如 Gantner 等人<sup>[22]</sup>研究了利用视频的基因信息来解决冷启动问题。内容基因化方法可以相对较好地提高 CBF 的推荐效果, 但是其计算成本和人力成本投入也相对较大。

伴随智能手机、平板等的普及, 越来越多的用户使用移动设备访问互联网服务。对于用户而言, 面对互联网内容服务商 (Internet Content Provider, ICP) 提供的海量内容, 想要获取的是感兴趣的或有价值的信息。而移动智能终端设备相对于传统个人电脑更贴近使用者, 且具有接收消息推送 (Push Notification) 的功能, 为视频推荐服务提供新的场景和途径。通过对用户兴趣的挖掘定制化消息推送, 正在成为移动视频推荐的新方式。因此在移动在线视频服务中, 所面临的场景为: ICP 选定一组视频 (最新上线或需要推广的新内容), 计划通过消息推送提示用户, 希望找到对所推送内容感兴趣的, 用户点击消息并观看其中推荐的视频, 从而实现个性化移动视频推荐。相对传统推荐场景中用户请求

ICP 页面再获取推荐, 消息推送可被看作是在移动互联网环境下由 ICP 发起的一种主动推荐服务。

因待推送的新视频尚无观看历史记录, 属于物品冷启动问题。如使用传统 CBF 方法, 基于简单信息抽取特征建立视频和用户之间的关系, 在实际应用中发现推荐效果往往有限。在基于内容推荐算法发展的近几年中, 研究者们不再局限于以往直接利用物品的一些基本信息和常规方法对用户兴趣和物品建模, 如 Deng 与 Sahebi 等人<sup>[23-24]</sup>提出引用不同平台、不同领域的的数据, 来提高 CBF 推荐效果。通过结合视频自身属性, 研究人员尝试利用视频的图像特征来分析用户的喜好<sup>[25-26]</sup>, 如 Sánchez 等人<sup>[27]</sup>通过视频的视听特征给用户推荐体育类视频; Qu 等人<sup>[28]</sup>进一步考虑综合应用视频的文本、图像、音频三种特征信息来预测用户对视频的评分, 从而丰富评分行为过少用户的特征, 有效提高推荐效果。在线视频的文本及互动信息为构造新的推荐模式提供了基础, 如 Yoshida 等人<sup>[29]</sup>提出对视频进行语义分析和情感分析, 并结合用户的兴趣和心情来建立推荐; Zhang 等人<sup>[30]</sup>利用用户对视频的评论文本和表情符号来预测用户对已观看的视频的评分, 较有效地解决了用户评分稀疏的问题。同时, 自然语言处理 (Natural Language Processing, NLP) 的发展为在线视频推荐算法带来了新元素, 如 Zhu 等人<sup>[31]</sup>研究将 LDA 模型应用在 CBF 中从而缓解冷启动问题并获得较好的成效; Simonet 等人<sup>[32]</sup>利用语义分析对 YouTube 的视频频道进行分类; 特别地, Bendersky 等人<sup>[13]</sup>提出结合主题模型和 CBF 的视频检索方法来推荐视频, 该策略对 YouTube 新视频 (如更新较快的新闻类视频) 的实时推荐起到了明显的提升作用。然而, 结合深度学习 (Deep Learning) 模型在 CBF 推荐中的研究和应用尚处于探索阶段。随着深度学习在自然语言处理方面的快速发展, 以及大数据环境下描述信息变得越来越丰富, 基于深度学习词向量模型的个性化视频推荐成为值得探索的研究方向。

针对上述场景, 本文提出了面向移动互联网服务应用深度学习词向量模型设计个性化的视频推荐策略。在该方法中通过引入词向量模型, 建立用户正行为向量和负行为向量, 结合视频内容的语言描述信息过滤推荐用户群。本文在第 2 节介绍深度学习模型在词向量方面的应用; 第 3 节详细介绍 DNN 推荐方法; 第 4 节给出实验结果, 并对结果进行分析和讨论; 第 5 节总结全文。

## 2 词向量模型

深度学习模型在图像、语音和自然语言处理等领域已带来突破和重要影响. 研究发现, 词语的向量化表示, 可以更好地构建对语言内容的数学操作和机器理解. 因为一个词包含的意义往往与该词周围的其他词有关, 所以研究人员提出使用神经网络方法学习词语的向量化表示<sup>[33]</sup>. 在向量空间中, 将词映射为高维向量, 词与词之间的向量操作存在和语义及句法相对应的一些关系. 基于这样的想法, Mikolov 等人于 2013 年开放了基于深度学习模型开发的将词转化为词向量的工具 word2vec<sup>①②</sup>. 本文所提出的移动视频推荐策略应用了该工具中基于 Hierarchical Softmax (HS) 的 CBOW 模型 (Continuous Bags-of-Words Model), 本节下述部分简要介绍该模型的基本原理<sup>[34-37]</sup>.

### 2.1 基于 HS 的 CBOW 模型

CBOW 模型网络结构如图 1 所示.

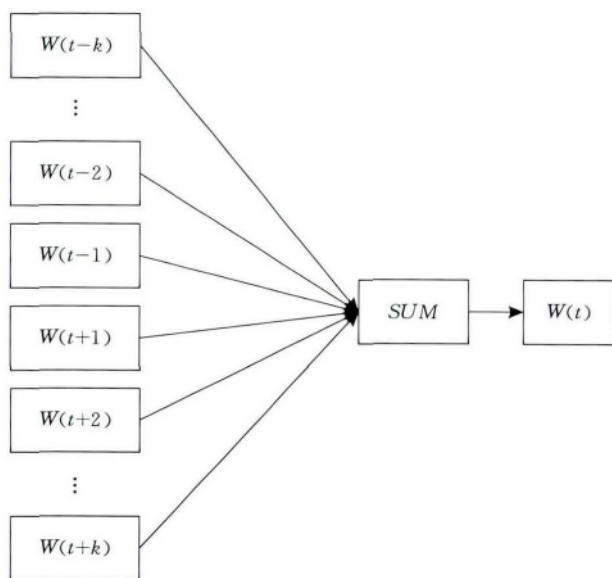


图 1 CBOW 模型网络结构

该网络结构主要分为 3 层: 输入层、映射层和输出层. CBOW 主要是通过一个词出现的上下文环境来判断该词出现的概率. 其数学表达式如下:

$$p(w_t | \ell(w_{t-k}, w_{t-k-1}, \dots, w_{t+k-1}, w_{t+k})) \quad (1)$$

其中  $w_t$  是要判断的词在  $w_{t-k}, w_{t-k-1}, \dots, w_{t+k-1}, w_{t+k}$  环境下出现的概率大小, 其中  $k$  是可调窗口参数, 通过调整  $k$  可以选择上下文环境的范围,  $\ell(w_{t-k}, w_{t-k-1}, \dots, w_{t+k-1}, w_{t+k})$  表示对输入  $w_t$  进行求和运

算. 在训练时, 整个网络的目标是使  $p(w_t | \ell(w_{t-k}, w_{t-k-1}, \dots, w_{t+k-1}, w_{t+k}))$  概率最大.

基于 HS 的 CBOW 模型的网络结构如图 2 所示.

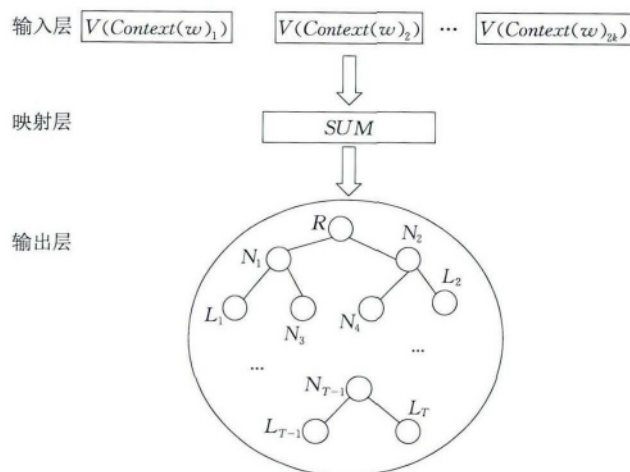


图 2 基于 HS 的 CBOW 模型的网络结构

图中的输入层为包含  $2k$  个词的词向量  $v(\text{Context}(w)_1), \dots, v(\text{Context}(w)_{2k}) \in R^m$ ,  $k$  为窗口大小, 即选择与中心词前后相距各  $k$  个词,  $m$  为词向量的长度. 在训练开始时系统会给整个语料库字典中的词随机初始化一个长度为  $m$  的向量. 映射层将输入层输入的词向量累加后得到结果  $X_w$ . 输出层为一个 Huffman 树, 其根节点记为  $R$ , 根节点的输入为  $X_w$ . 非叶子节点记为  $N_i, 1 \leq i \leq T-1$  ( $T$  为整个语料库对应的字典的大小), 非叶子节点对应的值记为  $\theta_i \in R^m$ . 叶子节点记为  $L_i, 1 \leq i \leq T$ , 其对应的值记为  $v_i \in R^m$ .  $v_i$  为整个语料库字典中词的词向量.

### 2.2 目标函数及最优化求解

对于图 2 中输出层的 Huffman 树, 每一条从根节点  $R$  到叶子节点  $L_i$  都对应唯一编码, 码元由 0, 1 组成. 码元 0, 1 可以类比成分类问题, 其中非叶子节点的左孩子对应码元 1, 被标为负类; 右孩子对应码元 0, 被标为正类. 用  $d_i$  对应第  $i$  个节点的码元. 分类时可采用 Sigmoid 函数, 即被分为正类的概率为

$$\sigma(X_w^T \theta) = \frac{1}{1 + e^{-X_w^T \theta}} \quad (2)$$

分为负类的概率为  $1 - \sigma(X_w^T \theta) = 1 - \frac{1}{1 + e^{-X_w^T \theta}}$ . 由此

得到从根节点到达  $w$  对应的叶子节点的概率为

① word2vec. <https://code.google.com/p/word2vec>

② Deep Learning in NLP(一)词向量和语言模型. <http://licstar.net/archives/328>

$$p(w | \mathbf{X}_w, \theta) = \prod_{i=2}^{l^w} p(d_i^w | \mathbf{X}_w, \theta_{i-1}^w) \\ = (1 - \sigma(\mathbf{X}_w^T \theta_1)) \cdot \sigma(\mathbf{X}_w^T \theta_2) \cdot \sigma(\mathbf{X}_w^T \theta_3) \quad (3)$$

其中:

$$p(d_i^w | \mathbf{X}_w, \theta_{i-1}^w) = \begin{cases} \sigma(\mathbf{X}_w^T \theta_{i-1}^w), & d_i^w = 0 \\ 1 - \sigma(\mathbf{X}_w^T \theta_{i-1}^w), & d_i^w = 1 \end{cases} \quad (4)$$

上述公式中:  $l^w$  为与中心词  $w$  前后相邻  $k$  个词组成的文本的大小(其值为  $2k$ ),  $d_i^w$  对应 Huffman 树中第  $i$  个节点的编码码元. 整个网络的目标就是使上述概率值最大. 将上式扩展到一般情形, 网络的目标函数可表示如下:

$$L = \sum_{w \in C} \prod_{i=2}^{l^w} \{ [\sigma(\mathbf{X}_w^T \theta_{i-1}^w)]^{1-d_i^w} [1 - \sigma(\mathbf{X}_w^T \theta_{i-1}^w)]^{d_i^w} \} \quad (5)$$

对式(5)取对数相应有

$$\sum_{w \in C} \sum_{i=2}^{l^w} \{ (1 - d_i^w) \cdot \log[\sigma(\mathbf{X}_w^T \theta_{i-1}^w)] + d_i^w \cdot \log[1 - \sigma(\mathbf{X}_w^T \theta_{i-1}^w)] \} \quad (6)$$

对于目标函数(6), 采用随机梯度算法来求目标函数的最大值. 目标函数中的参数为  $\theta_{i-1}^w$  和  $\mathbf{X}_w$ , 在梯度算法中, 求解的关键是目标函数对应的两个参数的梯度.

$$\frac{\partial \ell(w, i)}{\partial \theta_{i-1}^w} = \frac{\partial}{\partial \theta_{i-1}^w} \{ (1 - d_i^w) \cdot \log[\sigma(\mathbf{X}_w^T \theta_{i-1}^w)] + d_i^w \cdot \log[1 - \sigma(\mathbf{X}_w^T \theta_{i-1}^w)] \} \\ = (1 - d_i^w) [1 - \sigma(\mathbf{X}_w^T \theta_{i-1}^w)] \mathbf{X}_w - d_i^w \sigma(\mathbf{X}_w^T \theta_{i-1}^w) \mathbf{X}_w \\ = [1 - d_i^w - \sigma(\mathbf{X}_w^T \theta_{i-1}^w)] \mathbf{X}_w \quad (7)$$

根据式(7)计算的梯度, 可得参数  $\theta_{i-1}^w$  的更新公式为

$$\theta_{i-1}^w = \theta_{i-1}^w + \eta [1 - d_i^w - \sigma(\mathbf{X}_w^T \theta_{i-1}^w)] \mathbf{X}_w \quad (8)$$

同理, 可得参数  $\mathbf{X}_w$  的梯度公式:

$$\frac{\partial \ell(w, i)}{\partial \mathbf{X}_w} = [1 - d_i^w - \sigma(\mathbf{X}_w^T \theta_{i-1}^w)] \theta_{i-1}^w \quad (9)$$

于是, 对词的更新公式为

$$V(\text{Context}(w)_i) = V(\text{Context}(w)_i) + \eta \sum_{j=2}^{l^w} \frac{\partial \ell(w, j)}{\partial \mathbf{X}_w}, \\ i \in (0, 2k], l^w = 2k \quad (10)$$

上述即为 word2vec 中基于 HS 的 CBOW 模型的原理, 在模型中可以通过选择不同大小的词向量维度和文中中心词前后相邻词的个数(称为文本窗口长度)来获取不同的词向量表示效果.

### 3 基于 DNN 的移动视频推荐方案

对于推荐系统中的新物品、新用户都会不可避免地存在冷启动问题. 针对这样的情况, 目前大部分推荐系统会考虑使用混合模型和基于内容分析的推荐策略. 传统基于内容的推荐算法主要用到物品和用户 Tag 等描述信息, 信息越丰富, 推荐效果往往越好. 然而这些信息通常是由人工添加, 不同的人对同一事物有不同的看法, 也有不同的描述方式, 这样多多少少会造成数据的差别. 比如在描述时髦时, 有的人可能会用“时尚”, 有的人会用“潮流”, 这样就造成了差异. 在 NLP 领域为了处理类似的问题, 大部分情况会考虑使用话题模型(如 LSI、LDA), 或者使用基于深度学习模型的方法.

本文面向移动视频服务基于内容推荐的实际应用, 利用深度学习模型在词向量上的训练, 配合海量内容数据和用户行为数据, 设计移动视频推荐策略. 本节主要介绍使用词向量模型将视频向量化, 并建立用户的正行为向量和负行为向量以及利用向量之间关系来进行推荐分析.

#### 3.1 词向量模型训练

词向量模型训练过程如图 3 所示. 在对中文文本进行分析前, 一般要先对文本进行分词处理和提取关键词. 在本文的工作中, 使用了北京理工大学张华平等人开发的 NLP/ICTCLAS 分词系统<sup>[38]</sup>, 该系统对中文分词速率较快、效果较好、准确率相对较高. 通过分词并去停用词, 构建视频语料重要词汇库, 该词库是词向量工具的输入. 在获得语料重要词汇库后会对其提取关键词, 为构建视频向量和用户向量做准备.



图 3 词向量模型训练过程

在词向量模型训练中, 本实验使用了 word2vec 工具. 所训练的语料库包含腾讯视频提供的约 300 万部视频的媒资信息, 每条描述信息包括视频标题, 视频制作地区, 视频类型标签, 视频导演、主演、编剧, 视频简介, 视频看点介绍, 用户评论等信息. 为了减少噪声数据对训练词向量的影响, 本实验首先对数据进行过滤及合并处理: 过滤主要针对数据字段

不全的视频,此类视频在语料库中大部分属于 UGC (用户制作内容)类视频;合并主要针对具有相同信息的视频,此类视频多数为电视剧、动漫等剧集类视频.经数据清洗处理后,得到的有效数据约 100 万条视频文本数据.考虑训练速度和实现推荐的复杂度,本实验选用训练较快且工程上相对容易实现的 HS-CBOW 模型来建立视频语料词库的词向量.

在词向量维度的选择上,一般而言维度越高、文本窗口越大,词向量的特征表示效果相对会好,但同时词向量训练耗时越长,训练结果存储占用空间越大.本文工作所使用的数据集相对较大,应用 word2vec 训练词向量时,在维度设定为 200 维、文本窗口选为 5 的情况下,基于通用工作站(24 核 2.6 GHz CPU, 64 GB 内存)训练时长已超过 30 min.再使用训练得到的 200 维词向量对用户建模(基于大数据分布式计算平台),所产生的结果数据已超过 200 GB.综合考虑计算复杂度和所消耗的资源,同时为了保持较快的计算效率,本文提出的方案在大规模系统实验中设定的词向量维度保持为 200 维,文本窗口为 5.基于所使用的语料库,通过训练最后获得了约 20 万左右词汇的词向量.

### 3.2 视频向量化描述

词向量模型和传统的语义分析模型(如 LDA、LSI 等),在分析结果上的一个较大差别在于词向量模型是对文本中的词构建向量,而不是对整体文本构建向量.由这样的模型训练得到的词向量在向量空间上可以进行加减运算<sup>[37]</sup>,于是本文工作利用词向量在向量空间具有的加法运算性质构建视频文本向量.视频向量的计算公式如下:

$$\mathbf{V}_v = \frac{1}{n} \sum_{i=1}^n \mathbf{V}_{w_i} \quad (11)$$

其中: $\mathbf{V}_v$ 表示视频  $v$  的向量; $n$  为一个视频中抽取的  $n$  个关键词, $\frac{1}{n}$ 是归一化系数,为了防止不同视频提取的关键词个数不同而产生差异. $\mathbf{V}_{w_i}$ 是使用 word2vec 工具把第  $i$  个关键字表示为一个 200 维的向量.因此,可以使用上一小节中基于视频媒资信息作为语料库训练得到的词向量来构建将要被推送的视频的向量.

### 3.3 用户向量化描述

在本文所提出的推荐方案中,我们首先定义用户点击已推送的消息(Push Notification)为正行为,用户未点击已推送的消息为负行为,并分别根据用户的正行为和负行为建立向量模型.基于这样的设

计,我们可以从两个方面表征用户行为和兴趣.其中正行为代表了用户对推荐的视频产生了点击行为,一定程度上表达了用户接受了所推荐的内容.负行为代表用户对推荐的视频没有产生点击行为,但采用负行为表达用户不感兴趣时需要更多讨论,因为用户没有点击并不表示用户一定不感兴趣(不点击的原因也有可能是因推送时间、场景并不适合用户).在对实际数据的观察中,发现多数用户,负行为比正行为多,面对正负行为数据不均衡问题,我们目前较工程化地采取在数量级上从用户负行为数据中随机抽取 1.7 倍于正行为的负行为数据,来构建用户正负行为向量.为讨论方便,我们定义选中推送的用户为推送目标用户,用来选择推送目标用户的用户群称为备选用户群.

基于词向量所具有的加法运算特性,并结合 TF-IDF 方法,我们设计了用户的正行为向量化公式为

$$\mathbf{V}_u^+ = \frac{1}{m^+} \sum_{j=1}^{m^+} b_j \left( \frac{1}{n^+} \sum_{i=1}^{n^+} a_i^+ \cdot \mathbf{V}_{w_i}^+ \right) \quad (12)$$

其中: $\mathbf{V}_u^+$ 为用户  $u$  的正行为向量; $m^+$ 为用户  $u$  点击 Push 的视频个数; $n^+$ 为视频  $V^+$  的关键词个数; $\frac{1}{m^+}, \frac{1}{n^+}$ 均为归一化系数,目的是为了防止不同用户点击视频个数不同以及不同视频提取的关键词个数不同而产生差异. $a_i^+$ 为视频  $V^+$  中第  $i$  个词的 TF-IDF 权重. $\mathbf{V}_{w_i}^+$ 为用户点击的视频  $V^+$  中的第  $i$  个关键字,公式中系数  $b_j = \frac{1}{\log(1+N)}$ ,  $N$  为该视频在腾讯视频的播放量,该系数用于降低热门视频的影响力.

同理,用户的负行为向量化公式设计为

$$\mathbf{V}_u^- = \frac{1}{m^-} \sum_{j=1}^{m^-} b_j \left( \frac{1}{n^-} \sum_{i=1}^{n^-} a_i^- \cdot \mathbf{V}_{w_i}^- \right) \quad (13)$$

其中: $\mathbf{V}_u^-$ 为用户  $u$  的负行为向量; $m^-$ 为用户  $u$  没有点击 Push 视频的个数; $n^-$ 为视频  $V^-$  的关键词个数; $\frac{1}{m^-}, \frac{1}{n^-}$ 为归一化系数,以应对不同用户没有点击视频个数不同以及不同视频提取的关键词个数不同而产生差异. $a_i^-$ 为视频  $V^-$  中第  $i$  个词的 TF-IDF 权重. $\mathbf{V}_{w_i}^-$ 为用户对推送视频中没有点击的视频  $V^-$  的第  $i$  个关键字,其余变量含义同前.

### 3.4 结合用户正、负行为向量的综合推荐

结合用户正行为和负行为的综合推荐可以有多种方法,为了便于实现和实际应用,我们采用两组向



量与待推送视频向量之间距离的比值关系. 在计算高维向量之间距离的方法上, 我们在实验中尝试了欧式距离、皮尔逊距离、余弦距离以及余弦距离和欧式距离的混合距离, 实验结果见本文第 4 节详细讨论. 其中欧式距离的实验点击率 (Click Through Rate, CTR) 稍逊于其他 3 种距离; 余弦距离 CTR 平均值效果最优, 但其 CTR 波动大于其他 3 种距离. 为了最大限度提升推荐系统 CTR, 我们选用余弦距离作为高维向量间的距离计算方法.

在建立用户正行为向量和负行为向量后, 我们分别计算待推送视频的向量和全体用户的正负行为向量之间的余弦距离, 并计算它们之间的比值. 在此步骤中, 所建立的正负行为向量和视频向量余弦值的比值公式为

$$P = \frac{\cos\langle \mathbf{V}_v, \mathbf{V}_{u^+} \rangle}{\cos\langle \mathbf{V}_v, \mathbf{V}_{u^-} \rangle} = \frac{\frac{\mathbf{V}_v \cdot \mathbf{V}_{u^+}}{\|\mathbf{V}_v\| \cdot \|\mathbf{V}_{u^+}\|}}{\frac{\mathbf{V}_v \cdot \mathbf{V}_{u^-}}{\|\mathbf{V}_v\| \cdot \|\mathbf{V}_{u^-}\|}} \quad (14)$$

令  $x = \frac{\mathbf{V}_v \cdot \mathbf{V}_{u^+}}{\|\mathbf{V}_v\| \cdot \|\mathbf{V}_{u^+}\|}$ ,  $-1 \leq x \leq 1$ , 当  $x$  越接近 1, 表示用户的正行为向量和推送视频向量越相关, 反映了用户越可能对该推送视频感兴趣; 当  $x$  越接近 -1, 表示用户正行为向量和该推送视频向量越不相关, 反映了用户越可能对该推送视频不感兴趣. 令  $y =$

$\frac{\mathbf{V}_v \cdot \mathbf{V}_{u^-}}{\|\mathbf{V}_v\| \cdot \|\mathbf{V}_{u^-}\|}$ ,  $-1 \leq y \leq 1$ , 当  $y$  越接近 1 时, 表示用户的负行为向量和该推送的视频向量越相关, 但反映了用户越有可能对该视频不感兴趣. 而当  $y$  越接近 -1 时, 表示用户的负行为向量和该推送向量越不相关, 但反映了用户越可能对该视频感兴趣. 根据  $x, y$  取值的实际意义, 在推送一个视频时, 最理想的目标用户是  $x=1, y=-1$ . 在实际情况下, 用户的  $x, y$  值是分布于  $[-1, 1]$  之间. 为了建立数学模型来量化选定推送目标用户, 我们考虑把  $-1 \leq x \leq 1$ ,  $-1 \leq y \leq 1$  分为图 4 中  $C_1, C_2, C_3, C_4$  这 4 个区域, 并且每个用户的  $x, y$  值对应坐标中的一个点.

考虑  $x, y$  取值的实际意义, 那么  $(x, y)$  值属于  $C_2$  区域的用户将直接从备选用户群中剔除, 而属于  $C_4$  区域的用户将保留在备选用户群中. 对于  $C_1$  区域中的用户, 我们保留  $\frac{x}{y} \geq p, p \geq 1$  (对应  $C_1$  中的区域①) 的用户在备选用户群中, 过滤掉落入区域②中的用户. 这样我们保留了  $C_1$  区域中对推送视频感兴趣的可能性大于等于不感兴趣的可能性的用户.

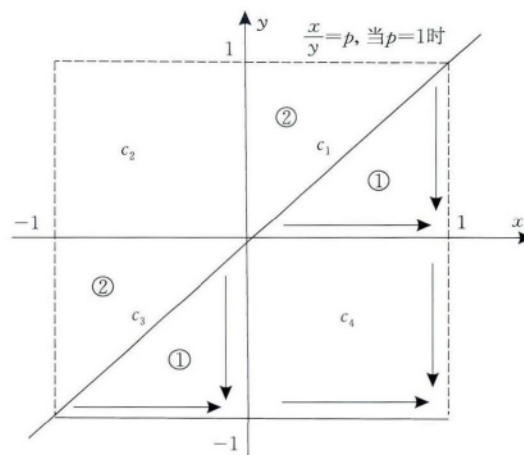


图 4 备选用户群与用户  $(x, y)$  值在坐标平面的分布关系

同样, 我们保留  $C_3$  中  $\frac{x}{y} \leq p, p \leq 1$  (对应  $C_3$  中的区域①区域) 的用户在备选用户群中, 过滤掉落入区域②中的用户.  $C_1, C_3$  中的子区域①, ②以直线  $\frac{x}{y} = p, p=1$  为边界. 图 4 中不同区域的箭头指向了不同区域内的备选用户成为推送目标用户的  $x, y$  取值趋势. 从  $x, y$  取值趋势, 我们能得到  $|p|$  值的大小趋势, 根据  $|p|$  值大小, 可以从备选用户群中选出推送目标用户. 表 1 简要总结了不同区域中的目标用户的  $|p|$  值大小.

表 1 不同区域中目标用户  $|P|$  值大小分布

区域	$x$	$y$	目标用户 $ P $ 值大小
$C_1$ -①	+	+	大
$C_3$ -①	-	-	小
$C_4$	+	-	小

$C_1$ -①表示  $C_1$  区域中①子区域,  $C_3$ -①表示  $C_3$  区域中①子区域. 在表 1 中可以看出在不同的区域中目标用户  $|P|$  值大小趋势存在不一致, 在  $C_1$ -①区域中, 要选择  $|P|$  较大的用户为推送目标用户, 其余两种情况要选择  $|P|$  较小的用户为推送目标用户. 因为存在这样的不确定情况, 我们不能有效的根据  $|P|$  值大小来选定目标用户. 但在此基础上, 我们观察到不同区域下, 目标用户的  $x, y$  的取值趋势通常均偏离  $\frac{x}{y} = p$  直线, 且  $C_4$  区域内的最理想目标用户  $(x=1, y=-1)$  偏离  $\frac{x}{y} = p$  直线最远. 由此我们考虑通过计算备选用户  $(x, y)$  到直线  $\frac{x}{y} = p$  的距离, 并对该距离按降序排序后取前 1000 万用户为推送目标

用户.

初始情况下,由于用户反馈数据尚不丰富,我们定义初始参考直线 $\frac{x}{y}=1$ ,即使用用户正负行为向量和视频向量关系相等为基准直线,并按公式 $d=\frac{|x_i-y_i|}{\sqrt{2}}$ 计算 $(x_i, y_i)$ 到直线 $\frac{x}{y}=1$ 的距离,根据该距离值降序排序后取前 1000 万用户为备选用户群.

## 4 实验及结果分析

本节简要介绍所提出的推荐方案在实验中所使用的数据以及实验环境和评价指标,然后比较离线环境下 DNN 算法和 Random 随机方法、Content-KNN 算法、ItemCF 算法的推荐效果,并考察在线环境下 DNN 推荐方案的推送结果,同时对数据结果进行分析.

在实验方案中,为了比较 DNN 方案和其他常用方法的推荐效果,我们增加了对 ContentKNN 算法的讨论. ContentKNN 是针对个性化推送设计的一种算法,该算法首先将视频库中所有视频向量化,在推送一个视频 $v$ 时:①先计算视频 $v$ 和视频库中其他视频的相似度,按相似度取 $k$ 个相似视频;②再基于 userCF 的思想,统计所有用户观看①中选取的 $k$ 个视频的数量情况;③然后根据②中所有用户观看数量取 TopN 个用户作为待推送视频的筛选用户群.

### 4.1 实验数据和实验环境

本实验基于腾讯视频的视频描述数据以及腾讯视频移动端用户点击推送消息中所推荐视频的历史数据,实验环境主要分为离线测试和在线测试.

腾讯视频是国内较有影响力的在线视频服务商之一,目前移动端拥有超过 4 亿用户,月度活跃用户超过 2 亿,每日高峰时段有超过 4000 万的用户同时在线<sup>①</sup>.在腾讯视频的视频库中,提供有数百万部视频,用户每天产生近十亿播放量,日播放记录数据达数百 GB.由于数据量大,本实验中只使用周期长度为 30 天的数据.在基于 30 天内推送视频的历史记录数据中,有接收 Push 推送消息的用户数数亿.在词向量训练方面,使用的视频描述数据(媒资信息)训练得到约近百万级别关键词的词向量.在进行向量化后,获得视频向量数据和用户向量数据在 TB 级别.

面对这样较大数据量分析的问题,本实验使用

了基于 Hadoop 大数据处理平台,机器集群由多 cluster 组成,每个 cluster 计算能力可达 60 GHz,存储能力 12 TB,网络传输能力 10 Gbps<sup>②</sup>.

### 4.2 评价指标

本实验的推荐目标主要针对待推送的视频过滤对该视频可能感兴趣的移动终端用户,因此需要考虑从视频角度和用户角度两方面综合评价所实验的推荐策略.

#### 4.2.1 视频角度

从视频角度出发主要考察视频点击率(CTR)、视频覆盖率(Coverage)、推送用户接受推送视频数量和点击视频数量的基尼系数(Gini coefficient, Gini).各个指标的定义公式如下:

$$CTR = \frac{N_c}{I_s} \quad (15)$$

其中: $N_c$ 是一个推送视频的点击用户数; $I_s$ 是一个推送视频下发的用户数.而覆盖率为

$$Coverage = \frac{N_s}{U} \quad (16)$$

其中: $N_s$ 为 $n$ 次推送总共下发的用户数; $U$ 为可以接受推送消息的用户总量.

Gini 系数原是经济学中衡量贫富差距的指标,推荐系统方面的研究主要借用该指标来衡量推送下发数量和推送点击量在用户之间的差距.针对本文中数据量较大的情况,采用较为简化的计算 Gini 的公式,如式(17)所示:

$$Gini = 1 - \frac{1}{N} \left( 2 \cdot \sum_{i=1}^N w_i - 1 \right) \quad (17)$$

其中, $w_i$ 为从对用户按接收或点击 Push 视频个数排序后从第 1 个用户到第  $N$  个用户接收或点击推送视频的累计个数占总的用户接收或点击视频个数的比例.

#### 4.2.2 用户角度

从用户角度出发主要需要考察预测准确度(Precision),计算式为

$$Precision = \frac{\sum_{i \in I} |R(i) \cap T(i)|}{\sum_{i \in I} |R(i)|} \quad (18)$$

其中: $R(i)$ 为使用推荐算法下发用户 Push 的视频集合; $T(i)$ 为用户实际点击的 Push 视频集合.准确度衡量各个算法过滤得到的用户群是否有效命中用

① 艾瑞:腾讯视频加冕视频移动端新王者. <http://www.cnetnews.com.cn/2015/0316/3048055.shtml>

② 本节中涉及的数据和数据处理平台,由腾讯公司合作提供.

户兴趣。

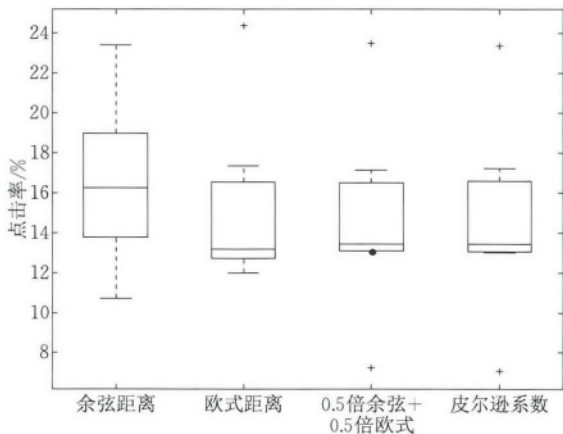
#### 4.3 分 析

在实验初期,我们尝试了传统基于内容的推荐方法,利用视频媒资信息关键词、视频分类标签、演员导演等信息建立向量,再依据用户观看过的历史记录建立行为模型,匹配两者。但在实际实验中发现这样建立出来的模型计算的视频和用户之间的相似度往往会接近于 0,结果不能直接用来做推荐。其原因主要是简单的关键词过多或过少都不能准确描述

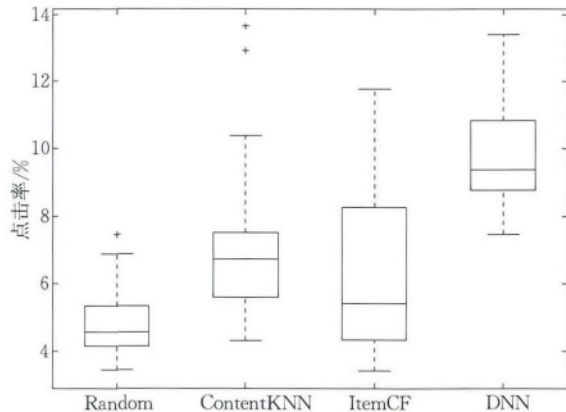
内容和用户,使得在稀疏矩阵上的距离计算失效。而在基于深度学习词向量模型建立的视频向量和用户行为向量上,则不会出现上述问题及现象。所以在实际系统的实验中,本文将主要比较 DNN 算法、ContentKNN 算法、ItemCF 算法以及 Random 方法的推荐结果。

##### 4.3.1 CTR 指标结果

在评估 CTR 指标时,我们分为离线测试和在线测试两种环境。其结果分别对应图 5 和图 6。



(a) 向量间不同距离计算方式离线测试 CTR 结果



(b) 离线实验 CTR 结果比较

图 5 离线测试 CTR 结果

图 5(a)主要对比了余弦距离、欧式距离、皮尔逊相关系数以及余弦距离和欧式距离的混合距离来衡量向量间距离的离线测试结果。我们使用箱线图描述统计结果。从结果图中可以较明显看出选用余弦距离,在提升推荐系统 CTR 上效果最优,因此本文中的实验选用余弦距离作为高维向量间距离的衡量方法。

图 5(b)对比了离线环境下的 ContentKNN、ItemCF、DNN 算法以及 Random 方法对 35 次推送的移动视频的 CTR 指标结果。从图中的统计描述结果可以看到,在离线实验中,ContentKNN、ItemCF 和 DNN 算法在 CTR 指标的上四分位值,中位值,下四分位值上均要高于 Random 方法,而 DNN 算法明显优于其他算法。同时,DNN 算法和 ContentKNN 算法的 CTR 结果稳定性相比 ItemCF 算法要好。4 种算法在对 35 次不同的移动视频的推荐(基于消息推送方法)中 CTR 的平均值分别为 4.8%、7.0%、6.3%、9.9% (分别对应 Random、ContentKNN、ItemCF、DNN 算法)。

鉴于上述离线结果,我们将 DNN 算法应用在在线环境。测试 DNN 算法在线上情况下的效果并对比同一时期在线 Random 算法的实验结果。在线

环境共测试了 35 次,CTR 对比结果如图 6(a)。从图中可以看到在在线环境下,DNN 算法在 CTR 指标上仍然保持着较大的优势。

此外,图 6(b)、图 6(c)中对比了在线环境下 DNN 算法在国产剧、国外剧、长视频、短视频上的推荐效果,测试次数各为 20 次。从结果上可以观察到 DNN 算法对国产剧的推荐效果要稍好于国外剧,而对长视频的推荐效果稍好于短视频。国产剧的推荐效果相对好于国外剧,其主要原因是在本实验中所使用的视频语料库信息中大部分是关于国产视频,而对于占相对少数的国外剧在整个由媒资信息组成的语料库中并没能很好的建立词向量模型。同时,观看国外剧的用户群相对少于国产剧,从而影响用户向量的生成。另一方面,长视频的推荐效果要稍好于短视频,因为长视频(电影、电视剧、综艺节目等)的描述信息比短视频(搞笑、生活、用户上传视频等)的描述信息更加丰富。

##### 4.3.2 Coverage 和 Gini 指标结果

本小节主要讨论 DNN 算法在离线和在线环境下的 Coverage 和 Gini 指标,并对比了同一时期在线 Random 推送的 Coverage 和 Gini 指标。其实验阶段小规模测试结果如表 2 所示。



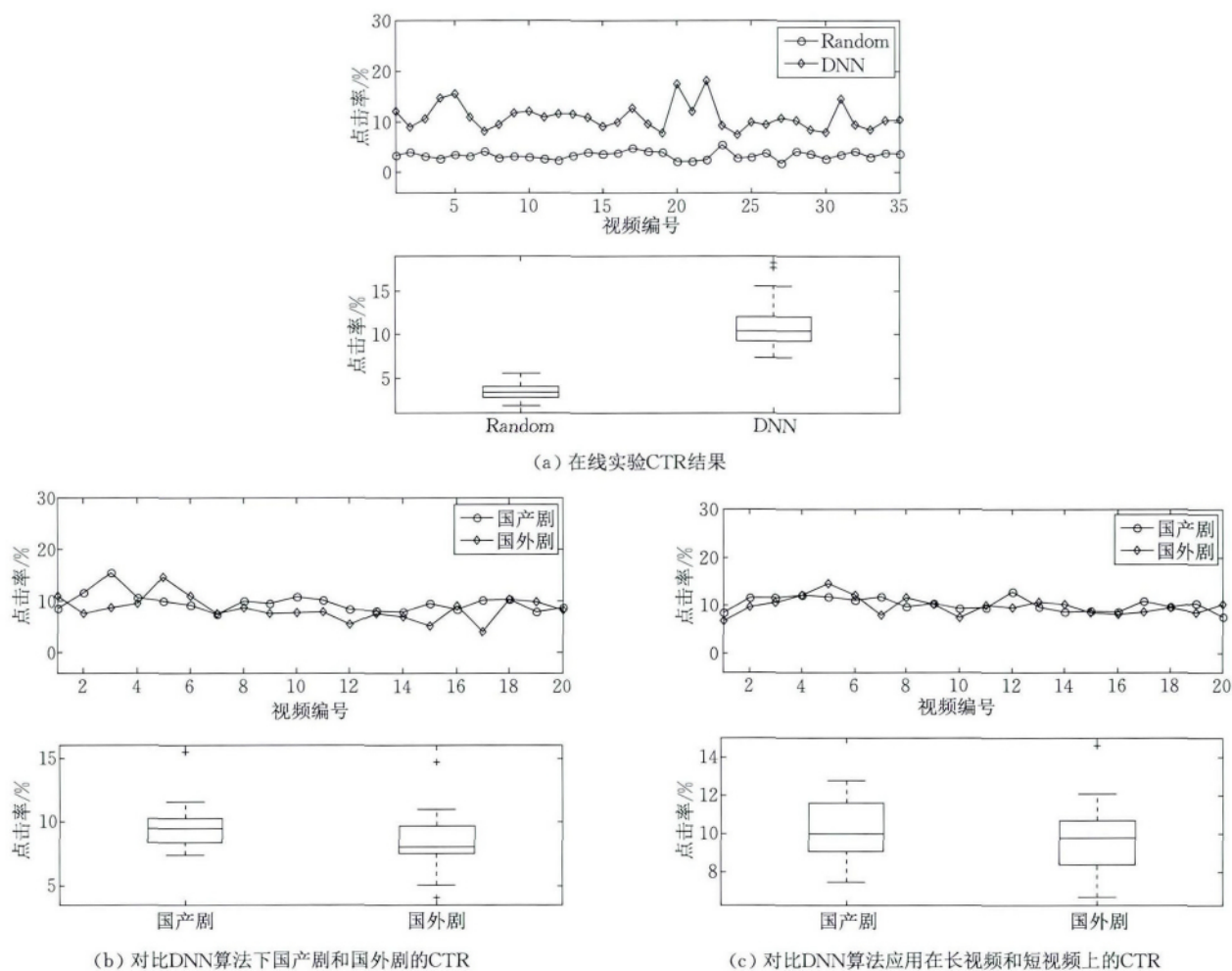


图 6 在线测试 CTR 结果

表 2 各种算法在离线或在线情况下的下发量、下发用户、点击量和 Coverage

算法	推送次数	下发量	下发用户数	点击量	Coverage/%
离线 DNN	23	902 997	398 876	67 211	19.5
离线 ContentKNN	23	704 449	485 709	38 154	23.8
离线 ItemCF	23	715 461	572 752	33 794	28.0
在线 Random	23	5 914 213	1 371 011	203 949	67.1
在线 DNN	23	1 070 975	266 251	91 298	23.0

从表 2 中可以看到在离线测试情况下, DNN 算法在 Coverage 指标和下发用户数方面相对其余两种离线算法表现要弱, 但是在下发量, 点击量, 点击用户数上效果要好. 而 Random 算法的下发量、下发用户数和点击量这些指标的绝对值都大于其他算法, 主要因为随机情况下考虑的是全体用户, 其用户基数大, 在随机抽取的时候量级依然较大. 然而, 从相对量角度观察, 3 种算法的表现, 都普遍要优于 Random 算法. 只是在 Coverage 覆盖率指标上, DNN 算法不存在优势, 这与计算用户行为向量的方法以及基于用户向量过滤筛选用户的策略有关.

在图 7 和图 8 中, 我们考察了不同推荐方法下发给用户的次数和用户点击次数的基尼系数. 从结果中可以观察到, DNN 下发给用户次数相比 Random 方法更加均衡, 点击用户要比 Random 方法集中.

对于 Random 算法中下发消息不均衡的问题, 我们分别从用户在 1 天、3 天、5 天和 7 天里观看次数的累积概率分布 (Cumulative Distribution Function, CDF) 观察移动视频服务用户的活跃度, 结果如图 9 所示.

从图 9 中可以观察到, 用户在活跃度上存在较大差异.

基于上述分析, Random 算法往往跟用户观看活跃程度有关. 对这种情况, 我们将移动视频用户分为接收推送消息量多的用户和接收推送消息量少的用户 (以用户接收推送消息数量衡量), 对比分析了这两类用户的观看活跃度 (以观看视频量衡量) 的 CDF 图, 其结果如图 10.

从图 10 可以观察到接收推送消息量多的用户

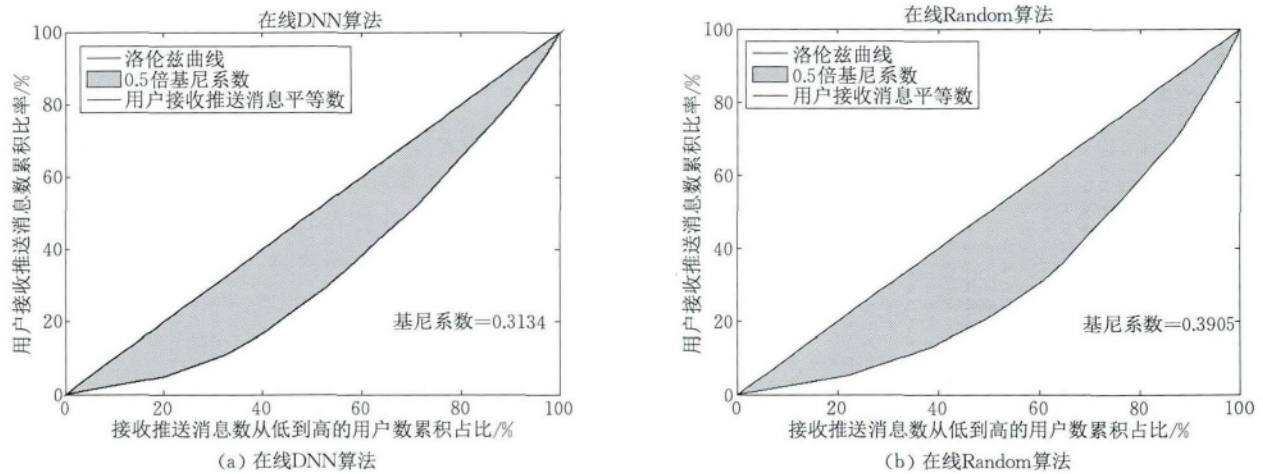


图 7 用户接收推送消息次数的基尼系数

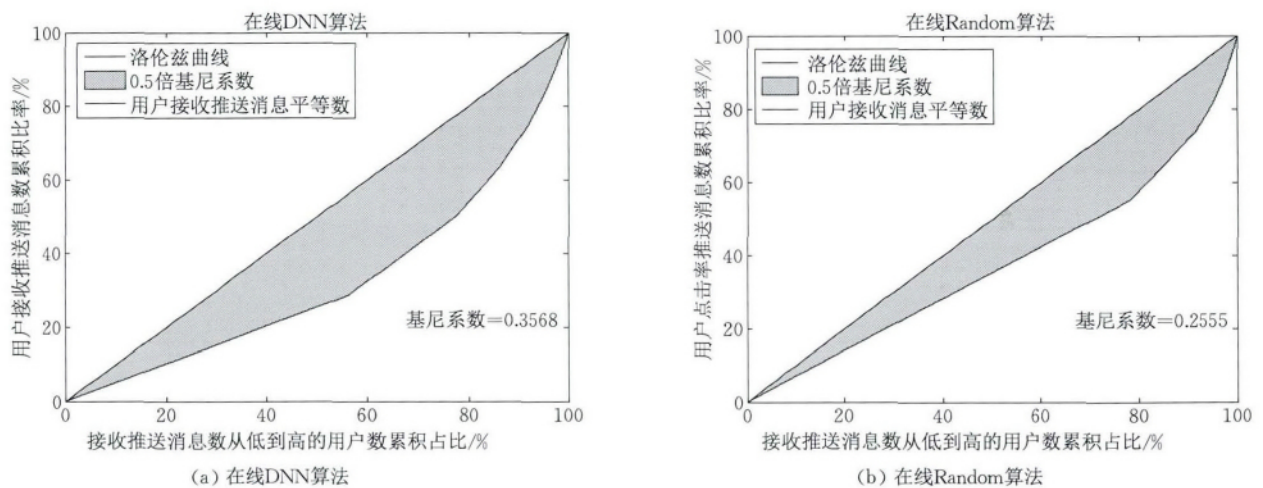


图 8 用户点击推送消息次数基尼系数

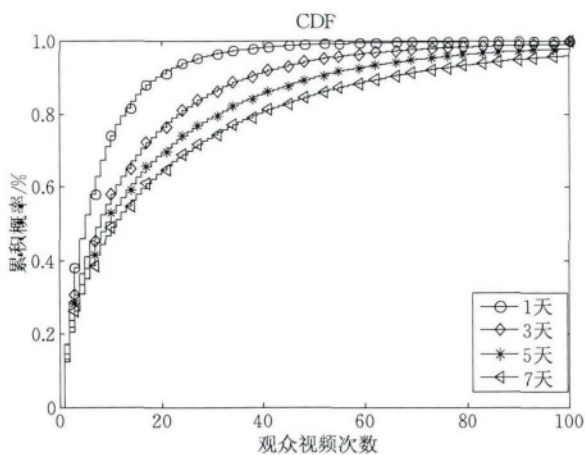


图 9 用户分别在 1、3、5、7 天中观看视频次数累积概率分布的观看活跃度要比接收推送消息量少的用户高,说明 Random 算法存在偏离缺陷。

最后,我们计算了 Random 和 DNN 算法各自推送用户接收推送消息个数与其观看活跃的相关度,分别为 0.40 和 0.31。从而可知 Random 算法偏

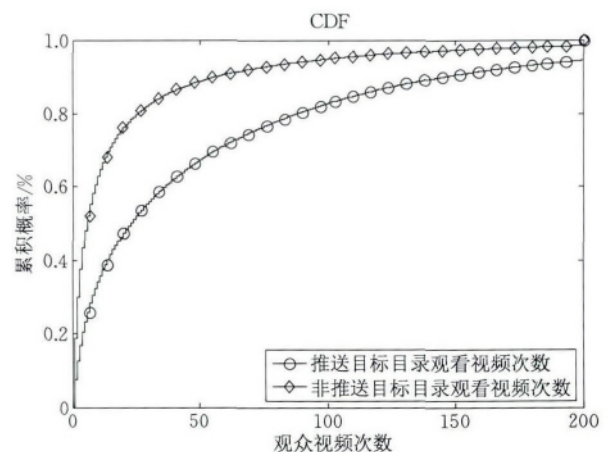


图 10 两类用户观看次数的累积概率分布

向选择观看活跃的用户。而 DNN 算法的点击用户比 Random 方法集中,说明了 DNN 算法在实际中提高了用户点击的可能性。

#### 4.3.3 预测准确度

对于预测准确度指标 (Precision), 我们主要对

比了线上 DNN 算法和同一时期线上 Random 方法的结果,以两种算法推送的 23 次推送的数据来对比分析. 两种情况下用户的预测准确度 CDF 分布如图 11 所示.

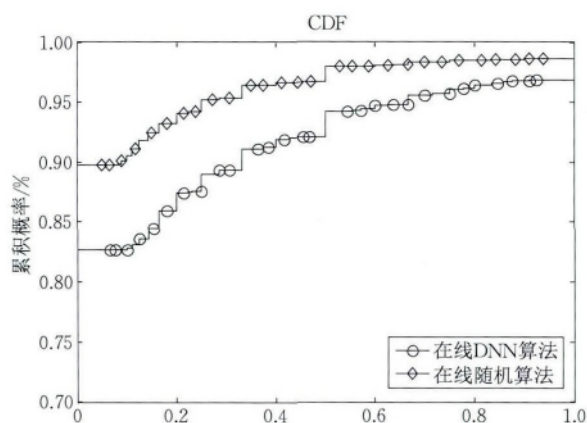


图 11 DNN 算法对用户预测准确度的影响

基于图中观察的结果可知,DNN 算法的预测准确度分布要高于 Random 方法. 基于在线环境下的实际测试结果来观察,DNN 算法有效提高了用户对推送视频的点击率.

## 5 结 论

针对信息过载现状和移动智能终端设备使用越来越广泛的情况,本文结合推荐算法和移动设备的推送功能应对信息过载问题,提出了一种解决冷启动问题的个性化移动视频推荐策略. 该方法主要利用了基于深度学习词向量模型建立视频向量和用户正负行为向量,把用户正负行为向量与视频向量的相似度映射到二维平面上,将选择推送用户的问题转化为计算点到直线的距离. 在检验基于深度学习模型的推荐策略的性能表现时,本文基于实际系统数据主要对比了其与离线 ContentKNN 算法、ItemCF 算法以及 Random 方法的推荐效果. 从实验结果比较可知,本文所提出的基于 DNN 算法的推荐方案在提升视频的 CTR、预测准确度等方面均获得较明显的效果.

本文所提出的推荐策略主要面对移动互联网用户个性化推送场景,在实际系统的工程应用中发现还需要综合考虑用户移动设备的网络在线状况(如 3G/4G 网络或 Wifi 环境对用户点击推荐视频的影响等),以及用户的观看行为等因素(分析用户观看视频的行为特征). 这些将作为本文后续工作探索的研究方向,而如何利用用户接收消息推送后的反馈

信息动态优化推荐策略,也将是本文后续研究的重要工作之一.

致 谢 腾讯(深圳)公司为本论文研究提供了数据、数据处理平台和大规模在线服务实验环境,本研究获得了腾讯公司的数据支持和实践应用. 审稿人也给予了宝贵的意见和建议. 在此一并表示感谢!

## 参 考 文 献

- [1] Jannach D, Zanker M, Felfernig A, et al. Recommender systems: An introduction. Cambridge, UK: Cambridge University Press, 2010
- [2] Xiang Liang. Recommended System in Action. Beijing: Posts & Telecom Press, 2012(in Chinese)  
(项亮. 推荐系统实战. 北京: 人民邮电出版社, 2012)
- [3] Davidson J, Liebal B, Liu J, et al. The YouTube video recommendation system//Proceedings of the 4th ACM Conference on Recommender Systems. Barcelona, Spain, 2010: 293-296
- [4] Liu Jian-Guo, Zhou Tao, Wang Bing-Hong. The research progress of personalized recommendation system. Progress in Natural Science, 2009, 19(1): 1-15(in Chinese)  
(刘建国, 周涛, 汪秉宏. 个性化推荐系统的研究进展. 自然科学进展, 2009, 19(1): 1-15)
- [5] Breese J S, Heckerman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering//Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence. Madison, USA, 1998: 43-52
- [6] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms//Proceedings of the 10th International Conference on World Wide Web. Hong Kong, China, 2001: 285-295
- [7] Linden G, Smith B, York J. Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 2003, 7(1): 76-80
- [8] Herlocker J L, Konstan J A, Terveen L G, et al. Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems (TOIS), 2004, 22(1): 5-53
- [9] Pazzani M J. A framework for collaborative, content-based and demographic filtering. Artificial Intelligence Review, 1999, 13(5/6): 393-408
- [10] Balabanović M, Shoham Y. Fab: Content-based, collaborative recommendation. Communications of the ACM, 1997, 40(3): 66-72
- [11] Walter F E, Battiston S, Schweitzer F. A model of a trust-based recommendation system on a social network. Autonomous Agents and Multi-Agent Systems, 2008, 16(1): 57-74
- [12] Kautz H, Selman B, Shah M. Referral Web: Combining social networks and collaborative filtering. Communications of the ACM, 1997, 40(3): 63-65

- [13] Bendersky M, Garcia-Pueyo L, Harmsen J, et al. Up next: retrieval methods for large scale related video suggestion//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA, 2014: 1769-1778
- [14] Guo Yan-Hong, Deng Gui-Shi. Hybrid recommendation algorithm of item cold-start in collaborative filtering system. Computer Engineering, 2008, 34(23): 11-13(in Chinese)  
(郭艳红, 邓贵仕. 协同过滤系统项目冷启动的混合推荐算法. 计算机工程, 2008, 34(23): 11-13)
- [15] Lam X N, Vu T, Le T D, et al. Addressing cold-start problem in recommendation systems//Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication. Suwon, Korea, 2008: 208-211
- [16] Park S-T, Chu W. Pairwise preference regression for cold-start recommendation//Proceedings of the 3rd ACM Conference on Recommender Systems. New York, USA, 2009: 21-28
- [17] Zhang Y, Shen H, Xiong Z, et al. Collaborative filtering algorithm based on item semantic and user characteristics. Journal of Information & Computational Science, 2015, 12(10): 4059-4067
- [18] Chen C C, Wan Y-H, Chung M-C, et al. An effective recommendation method for cold start new users using trust and distrust networks. Information Sciences, 2013, 224: 19-36
- [19] Nikolakopoulos A N, Garofalakis J D. Top-N recommendations in the presence of sparsity: An NCD-based approach. Web Intelligence, 2015, 13(4): 247-265
- [20] Maiti S, Chandrasekhar A, Pathan S, et al. Designing a parallel hybrid and commercial movie recommendation system. International Journal of Current Engineering and Technology, 2015, 5(3): 1694-1697
- [21] Pazzani M J, Billsus D. Content-Based Recommendation Systems//Brusilovsky P, Kobsa A, Nejdl W eds. The Adaptive Web. Berlin: Springer, 2007: 325-341
- [22] Gantner Z, Drumond L, Freudenthaler C, et al. Learning attribute-to-feature mappings for cold-start recommendations //Proceedings of the 2010 IEEE 10th International Conference on Data Mining (ICDM). Sydney, Australia, 2010: 176-185
- [23] Deng Z, Sang J, Xu C. Personalized video recommendation based on cross-platform user modeling//Proceedings of the 2013 IEEE International Conference on Multimedia and Expo (ICME). California, USA, 2013: 1-6
- [24] Sahebi S, Walker T. Content-based cross-domain recommendations using segmented models//Proceedings of the Workshop Proceedings (CBRecSys). Silicon Valley, USA, 2014: 57-64
- [25] Yamamoto K, Togashi R, Nakayama H. Content-based viewer estimation using image features for recommendation of video clips//Proceedings of the 1st Workshop on Recommender Systems for Television and Online Video (RecSysTV). Silicon Valley, USA, 2014
- [26] Azodinia M R, Hajdu A. A recommender system that deals with items having an image as well as quantitative features//Proceedings of the 2015 IEEE 9th International Symposium on Intelligent Signal Processing (WISP). Siena, Italy, 2015: 1-6
- [27] Sánchez F, Alduán M, Alvarez F, et al. Recommender system for sport videos based on user audiovisual consumption. IEEE Transactions on Multimedia, 2012, 14(6): 1546-1557
- [28] Qu W, Song K-S, Zhang Y-F, et al. A novel approach based on multi-view content analysis and semi-supervised enrichment for movie recommendation. Journal of Computer Science and Technology, 2013, 28(5): 776-787
- [29] Yoshida T, Irie G, Arai H, et al. Towards semantic and affective content-based video recommendation//Proceedings of the 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW). California, USA, 2015: 1-6
- [30] Zhang Weishi, Ding Guiguang, Chen Li, et al. Augmenting Chinese Online Video recommendations by using virtual ratings predicted by review sentiment classification//Proceedings of the 2010 IEEE International Conference on Data Mining Workshops (ICDMW). Sydney, Australia, 2010: 1143-1150
- [31] Zhu Qiusha, Shyu Mei-Ling, Wang Haohong. VideoTopic: Content-based video recommendation using a topic model//Proceedings of the 2013 IEEE International Symposium on Multimedia (ISM). California, USA, 2013: 219-222
- [32] Simonet V. Classifying youtube channels: A practical system//Proceedings of the 22nd International Conference on World Wide Web Companion. Rio de Janeiro, Brazil, 2013: 1295-1304
- [33] Collobert R, Weston J, Bottou L, et al. Natural Language Processing (Almost) from Scratch. The Journal of Machine Learning Research, 2011, 12: 2493-2537
- [34] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space//Proceedings of the International Conference on Learning Representations (ICLR). Scottsdale, USA, 2013
- [35] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality//Burgess C, Bottou L, Welling M, et al, eds. Advances in Neural Information Processing Systems. Lake Tahoe, USA, 2013: 3111-3119
- [36] Mikolov T, Le Q V, Sutskever I. Exploiting similarities among languages for machine translation. arXiv preprint arXiv:1309.4168, 2013
- [37] Le Q V, Mikolov T. Distributed representations of sentences and documents//Proceedings of the International Conference on Machine Learning (ICML). Beijing, China, 2014
- [38] Zhang H-P, Yu H-K, Xiong D-Y, et al. HHMM-based Chinese lexical analyzer ICTCLAS//Proceedings of the 2nd SIGHAN Workshop on Chinese Language Processing-Volume 17. Sapporo, Japan, 2003: 184-187



**CHEN Liang**, born in 1983, Ph. D., assistant professor. His main research interests include network services analytics, data science and data engineering.

**WANG Jing-Fu**, born in 1989, M. S. candidate. His research interests include machine learning and data analysis.

**WANG Na**, born in 1977, Ph. D., professor. Her main research interests include machine learning, pattern recognition and algorithm optimization.

**LI Xia**, born in 1968, Ph. D., professor. Her main research interests include machine learning, intelligent computing and signal processing.

## Background

The issue of personalized video recommendation has aroused interest of people in mobile online video services. Our work introduces the deep learning based method in the content-based recommendation to design a smart mobile online video recommendation strategy. Along with the development of mobile Internet, its convenience and immediacy enable mobile applications to serve users comprehensive promptly. Meanwhile, the explosion of information on Internet makes the recommendation become increasingly important. Currently there exist developed studies for content-based recommendation, including extracting topics with Latent Dirichlet allocation (LDA) model, analyzing the internal features of recommended items, and so on. Accompanying the development of deep learning in the area of natural language processing (NLP) and the abundant text descriptions in current big data environment, it is promising to explore the deep learning based recommendation strategy for mobile online videos.

This work proposed an improved content-based recommendation strategy, based on a deep learning related model,

and the push notifications enabled by mobile applications. There are three major steps involved. Firstly, we train the word vectors based on rich corpus of online videos with the word2vec tool. Secondly, we build up the video vectors and user vectors respectively via the video descriptions and user behavior history. Thirdly, we filter the target users who have potential interest in the candidate videos based on a proposed vector relationship model, and then push notifications of recommended videos to them accordingly. With abundant experiments in a large-scale online video system, we evaluate our proposal and compare it with random method, content-KNN algorithm and ItemCF algorithm. The experiment results show that our proposal gets significant improvement on the recommendation performance.

The project was supported by the National Natural Science Foundations of China (Grant Nos. 61502315, 61309030), the Natural Science Foundation of Guangdong Province, China (Grant No. 2015A030310366), and the Natural Science Foundation of SZU (Grant No. 201558).