

一种面向 Chrome 浏览器的视频云笔记插件

辛超 乔子健 孙艳春

(北京大学信息科学技术学院软件研究所 北京 100871)

(高可信软件技术教育部重点实验室 北京 100871)

摘要 随着互联网的发展以及 edX, Coursera 和 Udacity 等开放式网络教学平台的推广, 互联网教学逐渐兴起并广泛普及。互联网教学中学生获取知识的一个主要媒介是互联网课程中的教学视频资源。然而, 当前的互联网教学辅助平台中存在一个明显的不足: 学生无法及时针对教学视频细节在云端记录并分享笔记或寻求帮助。这将会影响学生对视频形式的教学内容的理解, 也不利于形成优质的学习生态系统。由此, 设计了一种面向 Chrome 浏览器的视频云笔记工具, 并采用了 HTML5, Node.js, MongoDB 等关键技术对其进行了实现。该工具能够针对互联网教学视频资源的细节内容记录、发布笔记并在云端共享, 有利于学生对视频教学资源中细节的整理、理解、讨论和最终掌握; 同时该工具可以根据学生人群在使用工具的过程中产生的数据对视频内容进行一定程度的解析, 挖掘视频教学资源中的关键点, 从而降低学生获取关键知识的难度, 辅助学生的学习过程。最后, 进行了相关的实例研究, 结果证明了所提出的分析设计方案的可行性和有效性。

关键词 开放式网络教学平台, 互联网教学, 视频教学资源, 视频云笔记, Chrome 插件

中图分类号 TP311 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.04.013

Cloud-based Notes Plugin for Video on Google Chrome

XIN Chao QIAO Zi-jian SUN Yan-chun

(Institute of Software, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

(Key Laboratory of High Confidence Software Technologies (Peking University), Ministry of Education, Beijing 100871, China)

Abstract With the development of Internet and the promotion of massive open online courses, online teaching has been widely used. In the environment of online teaching, most students acquire knowledge from the videos in online classes. However, there is an obvious shortcoming existing in most current online learning platforms that students can't take and share notes or ask for help conveniently while learning from videos. As a result, it may be difficult for students to have a deep understanding of the video lectures. So this paper designed and implemented a cloud-based note plugin for video on Google Chrome and implemented it with HTML5, Node.js, MongoDB and other key technologies. Students can take notes on videos lectures and share them with others. At the same time, the plugin can analyze the lecture content and find the key points in some degree according to students' using data so that it can support much help in both interactive discussion and knowledge understanding. Finally, this paper gave case studies to verify the feasibility and effectiveness of the solution.

Keywords Massive open online courses (MOOC), Internet education, Video teaching resources, Cloud-based notes plugin for video, Chrome plugin

1 引言

近年来, 在传统的课堂教育的基础上, 互联网教育逐渐兴起并且得到了迅猛发展。以 MOOC^[1] 为代表, 互联网教育实现了全球范围内的资源共享和平台开放, 使得任何人可以在任何时间、任何地点学习平台上的教学资源, 给教育行业带来

了新的发展动力。与此同时, 新的挑战也随之诞生。区别于之前的文本形式, 互联网中的教学资源多以视频为主, 这在考验学生自主学习能力的同时, 也对他们理解学习内容、区分视频中的重难点也造成了极大的困难。

针对上述问题, 本文研究并提出一种面向 Chrome 浏览器的视频云笔记设计方案, 并采用 HTML5 和 Node.js 等关

到稿日期: 2015-11-30 返修日期: 2016-03-01 本文受国家重点基础研究发展规划 973 项目(2011CB302604), 国家自然科学基金联合基金项目(U1201252), 国家自然科学基金创新研究群体科学基金(61421091)资助。

辛超(1992-), 男, 硕士生, 主要研究方向为软件工程; 乔子健(1992-), 男, 硕士生, 主要研究方向为软件工程; 孙艳春(1970-), 女, 副教授, 主要研究方向为软件工程、服务计算、计算机支持的协同工作, E-mail: sunyc@pku.edu.cn(通信作者)。

键技术设计并实现了该云笔记。在该云笔记的辅助下,学生可以在观看视频教学的同时及时有效地记录自己的想法和问题,并在云端共享,同时还可以在平台上查阅其他人对相关视频内容的理解,并与之交流和分享看法,从中得到启发,进而有效地提高了学生交流学习视频课件和知识分享的效率。同时根据对该云笔记的记录数据和用户的行为数据的统计和分析,该笔记可以对视频内容进行一定程度的解析,捕获视频中的关键点,从而帮助学生更好地学习视频课件,降低学生的学习成本。

本文第2节介绍了相关工作;第3节介绍了视频云笔记的系统分析和设计;第4节具体说明了视频云笔记的系统架构和关键技术;第5节通过实例研究介绍了该视频云笔记的使用场景,并验证了系统设计方案的可行性和有效性;最后进行总结全文并展望未来。

2 相关工作

本文的相关工作主要有3个部分:1)MOOC上的视频学习研究;2)视频云笔记产品;3)视频展示技术。

2.1 MOOC上的视频学习研究的相关工作

通过调研发现,目前学术界对于MOOC上的视频学习的研究大部分集中于以下几个方面。

(1)研究学生学习过程中的行为模式和中途停止学习的影响因素。如文献[2]通过对edx上4门课程的862个视频学习资源进行以秒为单位的用户行为数据的记录和分析,发现在长视频、复习视频和辅导视频中发生学生停止学习的情况的频率较高,但后两类视频会导致频繁和尖锐的行为交互峰值。为了探究原因,文章通过数据采样分析观察到61%的交互高峰伴随着学习过程中的视觉转换,并总结出了导致交互高峰的5种主要行为模式:刚开始学习新的视频资源、重回退到错过的内容看视频、跟随视频中的一个辅导教程、回放一个简短的视频片段、重复看视频中的一个非视觉性的知识内容。

(2)研究视频学习资源和文本资源对学习效果的不同影响。通常情况下,人们认为,在现数字时代中视频学习相对于文本学习应当更受欢迎,因为视频资源的表达力更强。但是文献[3]却通过一个检验在线视频和在线文本在概念学习、任务完成时间和学生印象的不同之处的实验,指出两者的表现几乎是一样的。

(3)研究视频内容的组织形式对于学习参与度的影响。文献[4]通过edx上4门课程的690多万条视频学习记录数据以及基于学生观看视频的时长和是否认真回答视频评估问题的关于学习参与度的评估数据的分析研究,指出短视频都非常具有吸引力,非正式谈话开头的视频和Khan风格的绘图板式教学视频也非常受欢迎,而质量很高的课堂预录视频则不一定能得到学生的欢迎。文章还指出学生对于教学视频和辅导视频的参与度是不一样的。

(4)研究通过用户行为来分析视频内容的潜力。文献[5]经过实证研究发现,学生经常反复观看包含有课后习题答案的视频段,从而认为通过对用户行为进行记录和分析可以完成一定程度的视频分析工作。

(5)研究如何提高用户的学习效率和积极性。文献[6-7]分别研究视频注释对自主学习的影响以及如何获取有效的视频注释,从而改善学习效果。文献[8-10]分别提出了不同的更适合进行MOOC学习的方法,如文献[10]提出了从视频内容衍生出来的多维非线性导航式学习法。

与这些研究相比,本文首次提出将笔记记录与视频学习、在线知识交流和资源共享相结合,设计并实现了针对一般视频学习资源的面向Chrome浏览器的视频云笔记工具,并提出了一种比较完善的通过笔记数据统计和用户行为分析来分析视频内容的方法。

2.2 视频云笔记产品的相关工作

随着互联网教学的发展,视频学习辅助工具的需求也逐渐增加。然而本文通过调研发现,目前已有的视频学习辅助工具大多功能欠缺或不完善。以较有代表性的两个工具国外的VideoNot.es^[11]和国内的网易云课堂^[12]为例,其都提供了在观看学习视频时自由地记录笔记并在云端共享的基本功能,但又各自存在着缺陷。VideoNot.es需要用户手动输入视频的链接地址,资源的迁移成本非常高,且没有提供有效的知识交流机制;而网易云课堂则只允许用户使用其提供的服务学习云课堂平台上的有限的视频资源。

通过以上分析,本文认为视频云笔记应当具备以下几个特点:1)普适性:视频云笔记应当不局限于某一个MOOC性质的网站,而可以为所有系统支持的视频格式提供笔记记录和分享的服务;2)便捷性:学生在进行视频学习的过程中可以方便地将突发灵感或疑问记录下来而不影响视频的学习体验;3)开放性:学生可以在平台上共享自己的笔记,同时可以查看并评论他人的笔记或对自己笔记的回复,形成良好的知识交流和分享机制;4)实用性:工具可以对视频的内容提供一定的辅助解析,降低学生获取视频核心内容的成本,体现工具的价值。

2.3 视频展示技术的相关工作

由于视频是目前互联网教学资源的主要形式,也是视频云笔记系统功能的起点,因此如何在扩展程序中进行视频资源的展示是本文需要解决的问题之一。Video.js是一个功能强大的在网页上嵌入视频播放器的通用的开源JS库,它提供了一整套JavaScript的API可供开发者调用,从而允许操作视频对象实现各种功能。本文在其基础上进行了二次开发来完成视频的展示工作。但是由于Flash视频和Canvas格式视频的内部的DOM结构不可见,无法对其进行解析和操作,因此本文介绍的视频云笔记目前只支持HTML5格式视频的展示和其他功能。

3 系统分析及设计

在分析和设计阶段,本文首先通过绘制用况图完成了系统需求模型的建立,之后采用面向对象的方法分别完成了系统分析和系统设计。

3.1 建立系统的需求模型——用况图

视频云笔记的主要参与者为普通学生用户,用户的需求可以规约出系统的用况描述表。以查看视频为例,其用况描述表如表1所列。

表1 查看视频用况描述表

名称	参与者	行为陈述(用户动作左对齐,系统动作右缩进)
查看视频	用户	提交查看视频请求; 在视频列表中匹配视频链接; If 匹配失败 then 将新的视频信息存入数据库; End if;
		显示视频及相关信息; 提交更改视频源请求; If 是正确的视频源 then 切换视频源; Else 提示视频源错误; End if;

根据系统的用况描述表可以得到系统的用况图,如图1所示。

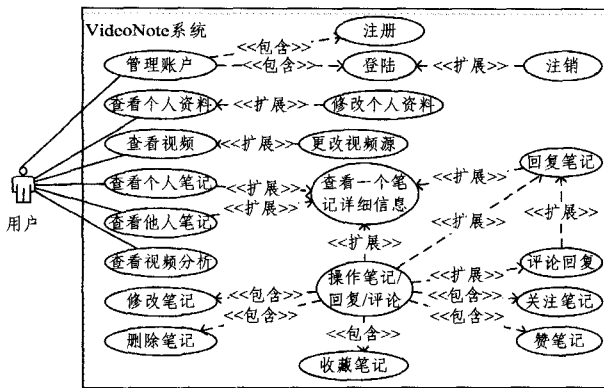


图1 系统用况图

3.2 系统分析

通过对问题域和系统责任进行分析,本文分别建立了类图和顺序图来完成系统的分析。

系统类图如图2所示,其规约如表2所列。

表2 系统类图规约

类名	说明
用户	记录用户的个人信息,提供用户的基本功能
视频	记录视频的基本信息
一段视频	记录一段时间视频上的信息,并提供查询此视频中所有笔记信息的功能
笔记	记录一个笔记的基本信息,并提供发布、回复等操作
回复	记录一个笔记的基本信息,并提供发布、回复等操作
评论	记录一个笔记的基本信息,并提供发布、回复等操作

在类图的基础上,本文还建立了顺序图作为辅助说明模型。图3示出了在一次笔记记录的过程中用户与浏览器端和服务器端交互的过程,其规约如下:

- 1)用户发起记录笔记的请求,并等待响应;
- 2)浏览器接收请求,并启动展示笔记记录界面;
- 3)浏览器向服务端请求记录此笔记需要用到的信息;
- 4)服务端接收请求并进行相应处理,然后返回处理结果;
- 5)浏览器端收到服务器信息,完成记录笔记页面展示,显示给用户;
- 6)用户看到浏览器响应后,在记录笔记页面输入笔记内容,并请求发布;
- 7)浏览器将笔记内容传输给服务器,并请求服务器处理;
- 8)服务器处理完毕,将处理结果返回给浏览器;

- 9)浏览器显示发布结果信息;
- 10)用户确认发布结果;
- 11)浏览器关闭笔记记录页面。

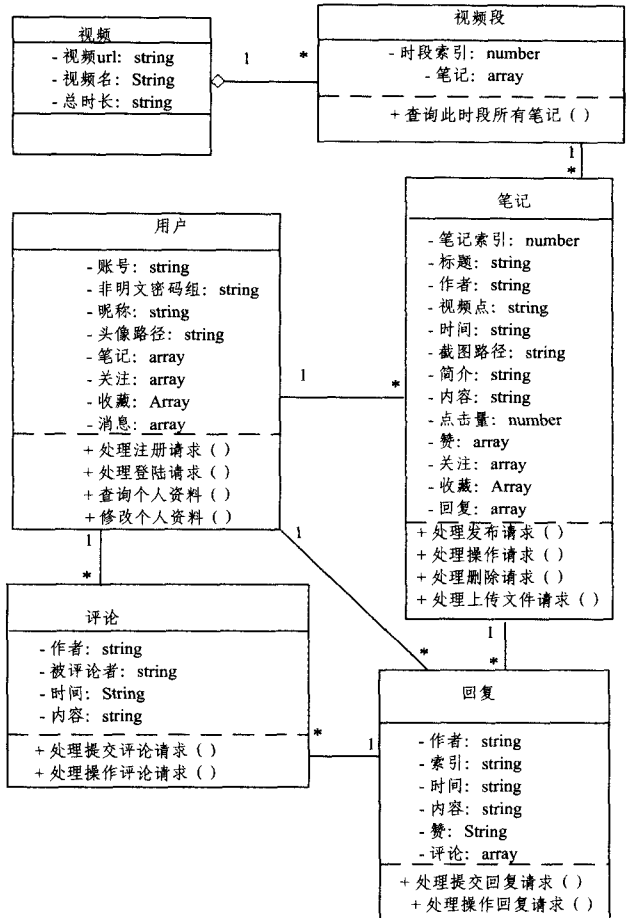


图2 系统类图

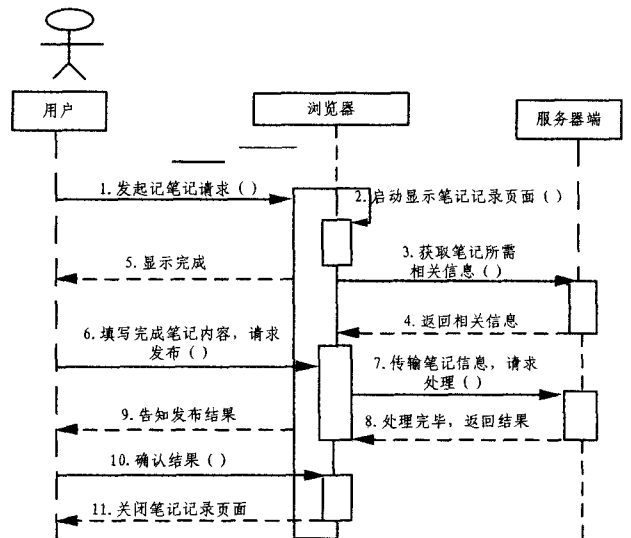


图3 记录笔记顺序图

3.3 系统设计

在系统设计的基础上,本文主要从问题域、人机交互、数据接口3个部分进行面向对象的设计。其中数据接口部分将在第4节中进行详细介绍。

在考虑了系统实现语言特性和运行平台之后,本阶段对问题域进行了以下改进:1)进一步明确类的属性和操作;2)确定编程语言为 JavaScript,存在基类 Object 可以被其他类复用;3)某些类的属性的数据是其他类的数组形式,如笔记类的回复属性是回复类的数组,需要进一步明确定义。

人机交互部分主要是对系统用户界面的设计。图 4 示出了人机交互部分的类图。

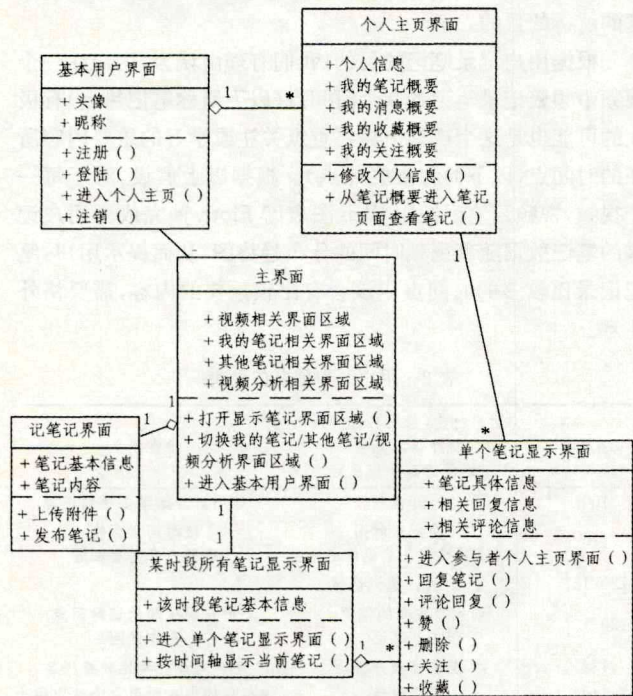


图 4 人机交互类图

4 系统架构及关键实现技术

本文设计的视频云笔记系统是一个 Node.js Web 应用,服务器端使用 Node.js 的 Web 开发框架——Express;数据库使用面向文档型的数据库——MongoDB;浏览器端使用 Bootstrap 和 Flat UI 前端开发框架来进行界面的设计,并以 Chrome 扩展程序的形式来实现系统功能。系统的整体架构如图 5 所示。

从结构上来看,视频云笔记系统主要分为服务器端和浏览器端。服务器端主要负责云端存储和后台运算,包括视频和笔记相关数据的存储、计算和分析,用户行为数据的统计等。浏览器端则负责与用户的交互,包括视频资源的获取和展示,笔记的记录、展示和分享,视频分析结果的展示等。服务器端和浏览器端之间通过 http 请求互相通信,共同协作。

从功能上来看,视频云笔记系统主要包括以下几部分:1)对视频的处理,包括视频的获取和展示。视频云笔记提供了两种获取视频资源的方式,第一种是在用户浏览嵌有视频的 Web 页面时,在视频对象的右键菜单中添加进入扩展程序的入口,并在用户进入扩展程序时传递视频相关的信息,如图 6 所示;第二种是在扩展程序的主界面提供用户手动输入视频链接的表单。将这两种视频获取方式相结合极大地降低了视频资源的迁移成本。当扩展程序获得视频的信息之后,就可以通过 Video.js 构建一个视频播放器来进行视频资源的

展示。2)对笔记进行处理。根据笔记的来源,可以将笔记分为用户自己记录的笔记和他人分享的笔记,视频云笔记系统将这两部分笔记区分开来,方便用户对笔记进行查看和管理。对于需要在一个页面上展示的笔记,系统首先以时间轴的形式展示所有笔记的简略信息,用户可以通过切换界面来显示某一个笔记的详细信息,包括该笔记的作者、创建时间、具体内容等,并且在笔记详细信息界面,用户可以完成对于某一个笔记的一些操作,如点赞、收藏、回复、评论等。笔记的记录是由富文本编辑器 Redactor.js 来完成的。3)视频辅助分析功能。视频云笔记提供了一种根据对笔记数据的统计分析来进行视频关键点挖掘的辅助功能,并提出了通过用户行为分析来进一步完善视频辅助分析功能的设想。

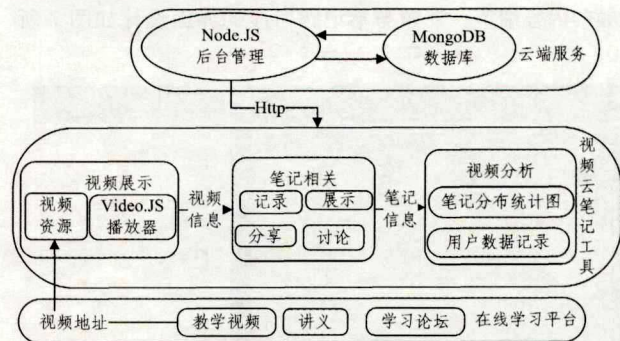


图 5 系统架构图



图 6 在视频对象的右键菜单添加扩展程序入口

在具体的实现过程中,本文主要遭遇并克服了以下 3 个问题:如何设计视频与笔记相对应的界面,如何构建数据模型,如何进行有效的视频辅助分析。

4.1 视频与笔记相对应的界面设计

与普通的文字资源相比,视频资源最大的不同点在于外界对资源的内容细节不易解析,视频资源内容无法在一个静态页面中全部展示出来,导致笔记不可能简单地与视频的某一部分内容直接对应,并且一个视频可能会对应非常多的笔记,因此需要设计一个简洁有效的对应方式。视频云笔记引入了一个视频段的概念,即将一个视频分为多个视频段,每个视频段与该时间段内的多个笔记相对应,且将笔记按对应视频内容所在的时间点排序以进行组织管理。每次在笔记展示界面展示笔记的简略信息时也仅展示当前视频段内的笔记。

当用户进入扩展程序的主界面时,系统首先根据获取到

的视频链接利用 Video.js 生成二次开发过后的视频播放器，从网上下载、解析并展示视频资源，同时将视频的信息上传至服务器端，经过服务器端的处理之后从数据库里查询出与该视频中当前视频段相对应的笔记信息，并由浏览器端根据笔记的信息在界面上以时间轴的形式显示其简略信息。随着视频的播放及间条的增长，每到达一个相关的笔记的时间点，与之对应的笔记都会高亮显示。当视频播放到达一个新的视频段时，系统会再次与服务器端交互，更新与之对应的笔记信息和界面。当用户在进行笔记记录操作时，系统会自动生成一个当前页面的截图，连同当前的系统时间、视频信息以及用户记录的笔记标题、内容等一起作为笔记信息的一部分上传至服务器。页面截图的存在使得用户能够根据笔记快速回想起视频的内容细节。视频与笔记相对应的界面设计如图 7 所示。

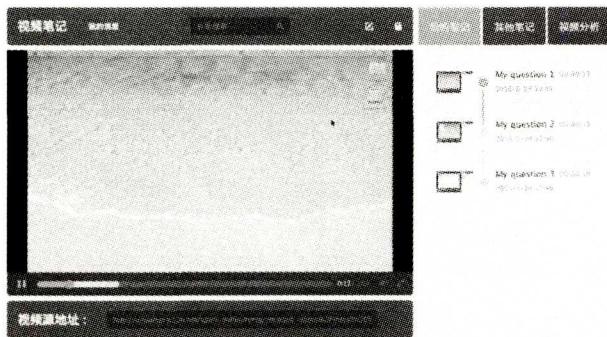


图 7 视频与笔记的对应

4.2 数据模型的构建

对于视频云笔记而言，用户是系统功能的参与者，视频是系统功能的源点，而笔记则是系统功能的主要载体。这三类对象中每个对象的数据资源都是十分庞大的，而且与其他对象之间的关系也非常复杂，所以如何构建简洁有效的数据模型成为本文遇到的一大挑战。

视频云笔记系统采用了多层嵌套的数据模型，即每个视频又包含有多个视频段，每个视频段又包含有多个笔记。同时系统定义其他用户对于笔记的文字回答为回复，而对于回复的文字回答为评论，对于评论的文字回答仍为同一级的评论，于是又有了每个笔记包含多个回复，而每个回复又包含有多个评论的关系。这样的多层嵌套特征也是系统使用面向文档型的数据库 MongoDB 来进行数据存储的一大原因。回复和评论实际就是特殊类型的“笔记”，他们和笔记有着几乎相同的属性(比如都记录了自己所属的视频的信息、所在时间点、创建者等)，但是在从属关系上互相区别。

视频云笔记系统中提供了很多用户对笔记的操作方式，如赞、关注、收藏、评论、回复等，每一项对笔记的操作都同时关联了用户、视频和笔记 3 个对象。故在系统构建的数据模型中对于笔记的检索提供了两种方式：1)根据笔记所在视频的位置信息，包括视频链接、视频段编号、笔记编号，如果该笔记是回复或评论，则需要回复编号和评论编号来进行笔记的检索；2)根据用户对笔记的操作类别来检索。用户每当对某个笔记进行了某项操作，都会将该笔记的检索信息记录到用户数据中相关操作的笔记列表里，因此根据用户 ID 和操作类

别也可以获得笔记所在视频的位置信息，从而检索到笔记。

由此可见，视频云笔记系统实际是将笔记对象作为媒介来连接视频对象和用户对象，从而完成数据模型的构建。

4.3 视频辅助分析

视频辅助分析是视频云笔记辅助用户学习的一个强有力的功能。通过对历史笔记数据以及用户使用数据等的统计分析来进行一定的研究，以达到对视频的关键内容进行一定程度的解析的目的。

根据用户记录笔记的目的，我们有理由认为：用户在一个视频中频繁记录笔记的某个时间点(以下简称笔记热点)有极大的可能也是这个视频中需要重点关注或学习的重要内容所在的时间点(以下简称视频热点)。根据以上假设，对于每一个视频，视频云笔记使用 Web 图表库 Flotr.js 完成了用户记录的笔记数据随着视频时间的分布趋势图，从而提示用户：笔记记录比较多的时间点应该会有比较重要的内容，需要格外注意。

表 3 用户行为信息记录表

编号	用户行为(除用户、时间、视频信息外需记录的额外信息)	潜在意义
100	打开视频	该用户开始学习视频资源
110	切换视频时间 (该时段相关笔记信息)	该时间点存在 重要内容需要回顾
120/121	显示/关闭笔记区域	
130/131	切换查看我的笔记/ 其他笔记	用户对视频的该时间点 存在疑问
132	查看视频分析	用户对该视频比较感兴趣
140	下载视频	该视频的内容对用户价值比较大
150	切换视频	用户切换学习资源
160	暂停视频	视频的该时间点存在疑问或 关键内容需要思考
200	查看某个笔记(该笔记相关信息)	
210/211	假/真回复(该笔记相关信息)	
220/221	赞/取消赞(该笔记相关信息)	
230/231	收藏/取消收藏(该笔记相关信息)	该笔记对用户有特殊意义/ 用户对该笔记有一定思考
240/241	关注/取消关注(该笔记相关信息)	
250	编辑某个笔记 (编辑后的笔记相关信息)	
260	删除某个笔记(该笔记相关信息)	该笔记对用户失去价值
300	查阅某人资料(该用户信息)	用户对被查阅用户比较感兴趣
400/401	假/真发布笔记(相关信息)	用户在该时间点存在思考

除此之外，用户在观看视频时还会存在一系列有意义的行为，这些行为对于视频的辅助分析也有一定的帮助，如暂停视频，可能是用户需要对某一时间点的重要内容进行思考；回退重看视频，可能是某一时间段的视频内容需要再次观看进行理解等。本文的相关工作中有研究表明，在 MOOC 上用户会对包含课后测验答案的视频段反复观看，从而证明了通过用户行为分析来进行视频辅助分析的潜力。但是用户行为分析需要一定规模的数据，而视频云笔记系统尚未推广，所以视频云笔记目前只是完成了在系统中对用户不同行为的检测分析和后台记录，尚未进行有效的行为数据分析研究。目前视频云笔记已实现的用户行为信息记录如表 3 所列。

5 实例研究

在实例研究中，本文针对 Coursea 上的一门课程设计了

一个用户通过视频云笔记系统来学习该课程的场景,从而证明视频云笔记系统可以帮助用户针对视频的任意细节记录并发布笔记且在云端与其他用户共享,相互讨论、启发,从而加强用户对于视频学习资源的理解和掌握,进而也验证了本文的云笔记的设计方案是可行有效的。

视频云笔记的主界面如图 7 所示,用户学习过程如下:

(1)用户 A 登陆视频云笔记系统并从 Coursea 上导入某个学习视频进而进入到系统的主界面进行学习,在视频播放到第 21 秒时用户 A 产生了疑问并准备将其记录下来,如图 8 所示。

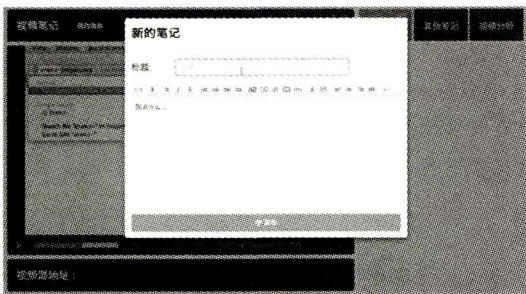


图 8 用户 A 记录笔记

(2)用户 A 记录完笔记之后将之发布到云端进行共享,希望自己的问题能得到其他人的解答,此时用户 B 观看到视频的同一段时间,在其他笔记一栏中发现了用户 A 的笔记,如图 9 所示。



图 9 用户 B 发现用户 A 的笔记

(3)用户 B 对用户 A 的问题进行了解答。之后用户 A 在自己的用户界面中发现有一条关于该笔记的新的消息,如图 10 所示。



图 10 用户 A 发现一条新消息

(4)用户 A 通过该消息发现了用户 B 的回答,在确认问题得到解决之后给用户 B 的回答给予了正面的评论,同时对该回答点了赞。如图 11 所示。

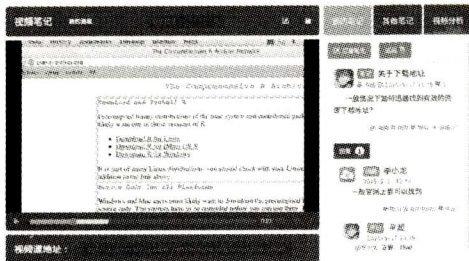


图 11 用户 A 评论了用户 B 的回答

(5)用户 A 查看了系统提供的该视频中笔记随视频时间的分布趋势图,得知视频在 50s~60s 的这个时间段可能存在关键内容,从而着重关注和学习。如图 12 所示。

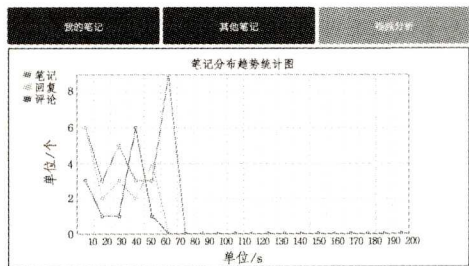


图 12 笔记随视频的分布趋势图

通过以上实例研究可以看出,视频云笔记给视频资源的学习者提供了一个简便的记录笔记、分享笔记和辅助分析视频内容的平台,有效地帮助用户对视频资源的学习过程,从而证明了本文提出的分析设计方案的可行性和有效性。

结束语 本文首先分析了随着互联网教育的发展存在的对视频学习辅助工具的迫切需求,接着通过对现有的 MOOC 视频学习研究和视频云笔记工具的调研,提出了完善的视频云笔记工具应当具备的特点。在分析相关工作的基础上,本文设计并实现了一个比较完善的视频云笔记的 Chrome 插件,其可以在帮助用户在观看学习视频的同时自由地记录并分享笔记,形成良好的知识交流和资源共享的机制。同时该工具还提供了一定的视频辅助分析的功能,可以帮助用户获取视频的核心内容,从而降低用户的学习成本。最后,本文进行了实例研究,验证了本文提出的分析设计方案是可行且有效的。

参 考 文 献

[1] SAHAMI M, GUZDIAL M, MARTIN F G, et al. The revolution will be televised: perspectives on massive open online education[C]//Proceeding of the 44th ACM Technical Symposium on Computer Science Education, 2013: 457-458.

[2] KIM J, GUO P J, SEATON D T, et al. Understanding in-video dropouts and interaction peaks in online lecture videos[C]// Proceedings of the First ACM Conference on Learning @ Scale Conference, 2014: 31-40.

[3] BREIMER E, COTLER J, YODER R. Video vs. text for lab instruction and concept learning[J]. Journal of Computing Sciences in Colleges, 2012, 28: 42-48.

副本存储位置的影响,可以看出,改进策略使低性能节点分担了高性能节点的副本存储数量,但是由于参数限制的原因,网络距离较远的节点存储的副本会少一些。随机策略主要以高性能节点存储文件,改进策略主要利用低配置节点存储访问率较低的副本,而高配置节点存储高访问率的副本。

系统会根据初期设定的冷热阈值在高低性能节点之间调动副本,从高配置节点中移出访问率低于冷阈值的副本到低配置节点,反之同理,从而达到根据设备级别来存储不同访问级别的数据的效果。

经过一段时间的访问,节点上的副本数的变动如图5所示。

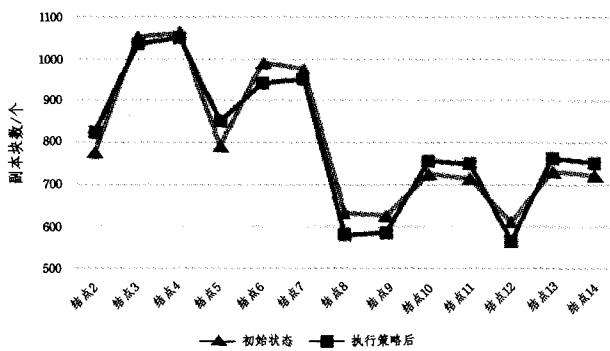


图5 副本数变化图

结束语 本文针对集群的异构性提出了一种改进方法,根据节点性能、内存负载、网络负载、存储负载和网络距离给出节点评价,基于评价结果优先选择存储节点。由于集群中节点性能的限制,文中提出一种方案来调度访问率超过一定阈值的副本数据,通过调度方案保证高访问量数据的使用效率,但是评价模型中的参数需要经过不断地选取调整才能促使模型达到预期的效果。

参 考 文 献

[1] CHEN K,ZHENG W M. Clouding Computing, System Instances and Current Research[J]. Journal of Software, 2009, 20(5): 1337-1348. (in Chinese)

陈康,郑纬民. 云计算:系统实例与研究现状[J]. 软件学报, 2009, 20(5): 1337-1348.

- [2] Apache Hadoop [OL]. [2013-07-10]. <http://hadoop.apache.org>.
- [3] MILOJICIC D,WOLSKI R. Eucalyptus, delivering a private cloud [J]. Computer, 2011, 44(4): 102-104.
- [4] 蔡斌,陈湘萍. Hadoop 技术内幕:深入解析 Hadoop Common 和 HDFS 架构设计与实现原理[M]. 北京:机械工业出版社, 2013.
- [5] REN C, YANG D J. The two-stage Dynamic Optimized Scheduling Mechanism Based on Cloud Storage[J]. Computer & Digital Engineering, 2014, 42(9): 1553-1557, 1716. (in Chinese)
- 任川,杨冬菊. 基于云存储的二阶段动态优化调度机制[J]. 计算机与数字工程, 2014, 42(9): 1553-1557, 1716.
- [6] TAO Y C, SHI L. Research on Dynamic Management of Data Replicas of Cloud Computing in Heterogeneous Environments [J]. Journal of Chinese Computer Systems, 2013, 34(2): 97-102. (in Chinese)
- 陶永才,石磊. 异构资源环境下的 MapReduce 性能优化[J]. 小型微型计算机系统, 2013, 34(2): 97-102.
- [7] KARGER D, LEHMAN E, LEIGHTON T, et al. Consistent hashing and random trees: distributed caching protocols for relieving hot spots on the world wide web[C]//ACM Symposium on Theory of Computing. CA, USA, 1997: 654-663.
- [8] XIE J, YIN S, RUAN X J, et al. Improving mapreduce performance through data placement in heterogeneous Hadoop clusters[C]//IPDPS Workshops. Atlanta: IEEE Computer Society Press, 2010: 1-9.
- [9] ZAMAN S, GROSU D. A distributed algorithm for the replica placement problem[C]//Proc. of IEEE Transaction on Parallel and Distributed System, 2011: 1455-1468.
- [10] LUO P, GONG X. Research and Improvement of Data Placement Strategy for HDFS [J]. Computer Engineering and Design, 2014, 35(4): 1127-1131. (in Chinese)
- 罗鹏,龚勋. HDFS 数据存储策略的研究与改进[J]. 计算机工程与设计, 2014, 35(4): 1127-1131.

(上接第 65 页)

[4] GUO P J, KIM J, RUBIN R. How video production affects student engagement; an empirical study of MOOC videos[C]// Proceedings of the First ACM Conference on Learning @ Scale Conference. 2014: 41-50.

[5] CHORIANOPOULOS K, GIANNAKOS M N, CHRISOCHOIDES N. Open system for video learning analytics[C]// Proceedings of the First ACM Conference on Learning @ Scale Conference. 2014: 153-154.

[6] GAŠEVIĆ D, MIRRIHI N, DAWSON S. Analytics of the effects of video use and instruction to support reflective learning [C]// Proceedings of the Fourth International Conference on Learning Analytics and Knowledge. 2014: 123-132.

[7] LI Y, LU J J, ZHANG Y F, LI R, et al. A Novel Video Annotation Framework Based on Video Object[C]// Proceedings of the 2009 International Joint Conference on Artificial Intelligence.

2009: 572-575.

- [8] YADAV K, SHRIVASTAVA K, DESHNUKH O. Towards Supporting Non-linear Navigation in Educational Videos[C]// Proceedings of the 16th International Conference on Multimodal Interaction. 2014: 82-83.
- [9] BURGE J. Insights into Teaching and Learning: Reflections on MOOC Experience[C]// Proceedings of the 46th ACM Technical Symposium on Computer Science Education. 2015: 600-603.
- [10] FIDALGO-BLANCO Á, SEIN-ECHALUCE M L, GARCÍA-PENALVO F J, et al. Improving the MOOC learning outcomes throughout informal learning activities[C]// Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality. 2014: 611-617.
- [11] VideoNot. es[EB/OL]. <http://www.videonot.es>.
- [12] 网易云课堂[EB/OL]. <http://study.163.com>.
- [13] 朴灵. 深入浅出 Node.js[M]. 北京:人民邮电出版社, 2013.