

西南交通大学
本科毕业设计（论文）

题目（黑体 2 号字）

年 级： 2016 级

学 号： _____

姓 名： _____

专 业： _____

指导教师： _____

二零二零年五月

西南交通大学

本科毕业设计（论文）学术诚信声明

本人郑重声明：所呈交的毕业设计（论文），是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日期： 年 月 日

西南交通大学

本科毕业设计（论文）版权使用授权书

本毕业设计（论文）作者同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权西南交通大学可以将本毕业设计（论文）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本毕业设计（论文）。

保密☐，在_____年解密后适用本授权书。

本论文属于

不保密☒。

（请在以上方框内打“√”）

作者签名：

指导教师签名：

日期： 年 月 日

日期： 年 月 日

西南交通大学本科毕业设计（论文）

院系 信息科学与技术学院 专 业_____

年级 2016 级 姓 名_____

题目 _____

指导教师

评 语 _____

指导教师 _____ (签章)

评 阅 人

评 语_____

评 阅 人 _____ (签章)

成 绩 _____

答辩委员会主任 _____ (签章)

年 月 日

毕业设计（论文）任务书

班 级 软件 2016-3 班 学生姓名 _____ 学 号 _____

发题日期：2019 年 11 月 22 日

完成日期：2020 年 5 月 22 日

题 目 _____

1、本论文的目的、意义

2、学生应完成的任务

3、本论文与本专业的培养目标达成度如何？（如在知识结构、能力结构、素质结构等方面有哪些有效的训练。）

此为计算机科学与技术专业的最新毕业要求，其他专业应替换为各自专业的毕业要求！

本论文支撑本专业以下毕业要求的达成：（1）能够通过查阅和分析文献，为计算机系统及工程的问题求解寻找方案，并认识到所求解的问题具有多种可能的解决途径（指标点 2.3）；（2）能够针对特定需求确定目标，设计计算机系统框架、组成模块，合理组织/存储数据，基于适当的模型进行系统设计与实现，并体现一定的创新意识（指标点 3.3）；（3）能够在解决方案中从技术、非技术（如经济、社会、健康、安全、法律、文化以及环境等）角度，对设计方案的可行性进行评价和分析（指标点 3.4）；（4）能够采用科学方法对计算机系统及工程问题进行研究，通过实验对比、文献综合、归纳整理得到合理有效结论，并对其进行规范表述（指标点 4.3）；（5）能够利用开发环境和工具，对计算机系统及工程问题进行模拟仿真和数据分析（指标点 5.3）；（6）能够识别、分析、评价特定需求的计算机系统在设计 and 实现中对社会、健康、安全、法律以及文化的影响，并明确自己应承担的责任（指标点 6.2）；（7）能够评价计算机系统的设计、开发、运行和维护对环境保护和社会持续发展的影响（指标点 7.2）；（8）能够通过口头、文稿、图表等方式、陈述和表达自己的观点，能够就计算机系统及工程问题与同行和相关人员进行交流（指标点 10.1）；（9）能够根据对工作内容和过程的记录与整理，撰写技术报告和设计文稿、陈述发言或回应质询（指标点 10.2）；（10）了解计算机系统工程管理原理与经济决策方法，理解计算机系统项目的组织模式和实施过程，掌握项目管理原理和内容（指标点 11.1）；（11）正确认识自主学习的必要性和重要性，认识到本专业是一个发展迅速的学科，具有自主学习和终身学习的意识（指标点 12.1）；（12）具备自主学习新技术和新方法的能力，能够通过学习不断提高、适应信息技术和职业的发展（指标点 12.2）。

4、论文各部分内容及时间分配：（共 17 周）

第一部分_____（ 周）

第二部分_____（ 周）

第三部分_____（ 周）

第四部分_____（ 周）

第五部分_____（ 周）

评阅及答辩_____（1 周）

备 注_____

指导教师： _____ 2019 年 11 月 22 日

审 批 人： _____ 2019 年 11 月 22 日

本页所处页数如为**偶数**，则保留；否则，如为**奇数**则删除此页。

摘要

注：

(1) 本页前面页数应为偶数，以保证此页双面打印时为正面！

(2) 至少 500 字。

(3) 推荐的写作段落顺序及内容如下：

①第一自然段（2~4 行）：选题的提出；

②第二自然段（2~4 行）：论文研究内容总概；

③第三自然段（10~12 行）：描述论文所做具体内容以及其中的一些亮点（如，依据什么提出或改进了什么算法或模型，利用何方法或技术解决了什么关键问题，理论分析或仿真实验表明结果如何 等等）。

以下为示例中文摘要的前 2 个自然段：

Web 信息的爆炸性增长使 Internet 成为我们获取信息资源的重要途径，而在全球一体化的今天，人们对翻译质量和翻译速度的要求也日趋严格。网络上拥有大量的双语对照信息，而传统的搜索引擎无法对其进行充分的利用。因此，本文从搜索技术和翻译技术上进行研究，提出了一个基于双语翻译的搜索引擎系统。

本文利用 Heritrix 和 Lucene 工具，在计算机辅助翻译的基础上利用搜索引擎技术，实现检索具有双语对照信息的网页。本文主要研究如何从海量信息库中爬取网页资源，设计出识别双语网页和提取双语语料的方法，并构造出合适的索引器和检索器，通过用户接口将网页信息输出给用户。

.....

关键词：词 1；词 2；词 3；词 4；词 5

注：① 关键词与摘要正文之间空一行；

② 关键词应按其涵义外延从大到小顺序排，列 3-5 个。

Abstract

注：特别注意此部分不能用中文标点符号，而应用英文标点+空格（一个）来分隔。
以下为示例英文摘要的前 2 个自然段：

The explosive growth of Web information makes Internet be an important way to obtain information resources for us. Because of today's global integration, the quality and speed of translation work are becoming stricter and stricter. There is a large amount of bilingual information in the network, but traditional search engines cannot make full use of them. Therefore, this thesis studies search and translation technologies and proposes a search engine system based on bilingual translation.

This thesis realized searching Web pages that have bilingual information, and combined computer-aided translation technology with search engine technology with Heritrix and Lucene tools. This thesis mainly researches how to crawl Web resources from massive information, and designs a method for recognizing bilingual Web pages and extracting bilingual corpora. The search engine implements in this thesis will output Web information to users through the user interface with the help of indexers and searchers.

.....

Keywords: kw1; kw2; kw3; kw4; kw5

注：① **Keywords** 与英文摘要正文之间空一行；
② 与中文关键词对应排列。

目 录

| | |
|-----------------------|---|
| 第 1 章 绪 论 | 1 |
| 1.1 背景与意义 | 1 |
| 1.2 国内外发展（应用）现状 | 1 |
| 1.3 论文所做工作及思路 | 2 |
| 1.4 论文章节安排 | 2 |
| 第 2 章 标题 | 4 |
| 2.1 一级节标题 | 4 |
| 2.1.1 二级节标题 | 4 |
| 2.1.2 二级节标题 | 5 |
| 2.1.3 二级节标题 | 5 |
| 2.2 一级节标题 | 5 |
| 2.2.1 二级节标题 | 5 |
| 2.2.2 二级节标题 | 5 |
| 2.3 本章小结 | 5 |
| 第 3 章 标题 | 6 |
| 3.1 一级节标题 | 6 |
| 3.1.1 二级节标题 | 6 |
| 3.1.2 二级节标题 | 6 |
| 3.2 一级节标题 | 6 |
| 3.2.1 二级节标题 | 6 |
| 3.2.2 二级节标题 | 6 |
| 3.3 本章小结 | 6 |
| 第 4 章 标题 | 7 |
| 4.1 一级节标题 | 7 |
| 4.1.1 二级节标题 | 7 |
| 4.1.2 二级节标题 | 7 |
| 4.2 一级节标题 | 7 |
| 4.2.1 二级节标题 | 7 |
| 4.2.2 二级节标题 | 7 |
| 4.3 本章小结 | 7 |

| | |
|-------------------|----|
| 第 5 章 标题 | 8 |
| 5.1 一级节标题 | 8 |
| 5.1.1 二级节标题 | 8 |
| 5.1.2 二级节标题 | 8 |
| 5.2 一级节标题 | 8 |
| 5.2.1 二级节标题 | 8 |
| 5.2.2 二级节标题 | 8 |
| 5.3 本章小结 | 8 |
| 结 论 | 9 |
| 致 谢 | 10 |
| 参考文献 | 11 |
| 附录 1 标题 | 13 |
| 附录 2 标题 | 14 |

第 1 章 绪 论

注：(1) 本页前面页数应为偶数，以保证此页双面打印时为正面！
(2) 正文中文用宋体，英文与数字用 Times New Roman，均为小四号。
(3) 论文篇幅要求从第 1 章至参考文献原则上 40 页。

1.1 背景与意义

注：所有章节标题后不能带标点符号。

1.2 国内外发展（应用）现状

注：本节应标注文献引用！尽量**避免**直接引用方式，且尽量**用自己理解后的话来叙述**！引用的**上标**用小四号字，**Times New Roman** 字体^[1]，多篇文献的标注如^[2, 5-8]。当提及的参考文献为文中直接说明时，其序号应该用小四号字与正文排齐，如“由文献[8, 10-14]可知”。文中其它章节如有引用亦应标注！注意：全文**不能在标题上标注**！

1.3 论文所做工作及思路

注：本节应从以下方面叙述：本论文将要做什么内容，准备分几步来做，各步分别准备怎么做，其中可能会涉及什么关键问题或技术，准备如何解决等等。

1.4 论文章节安排

论文共分 xx 章。

注：以下分章简述各章的内容。

第 1 章……

第 2 章……

……

注：

(1) 文中的图、表不要被分页断开！表如被分到两页，则应对分在另一页的表的右上方写上“续表”。图与图名不能分到两页上。图应尽量紧凑些。表的左右两个边框应隐去。图表应分别编号（不能混编），且均以“图 章号-序号、表 章号-序号”形式按章编号（例如，图 1-1，表示第 1 章的第 1 张图；图 2-1 表示第 2 章的第 1 张图；表 1-1 表示第 1 章的第 1 张表；表 2-1 表示第 2 章的第 1 张表）。引用别人的图或表（亦应在文中标注引用），尽量自己重新绘制，不要直接复制。图表中文字（包括图名、表名）字号建议用 5 号字。

以下的表 1-1 和图 1-1 为表和图的示例。

表 1-1 Corpus 表

| 列名 | 数据类型 | 说明 |
|-----|--------------|-----------------|
| Cid | Int | 自增型单页语料编号，初值为 1 |
| Url | Varchar(200) | 网页地址 |
| Ccn | Text | 单页互译文本的中文句子 |
| Cen | Text | 单页互译文本中的英文句子 |

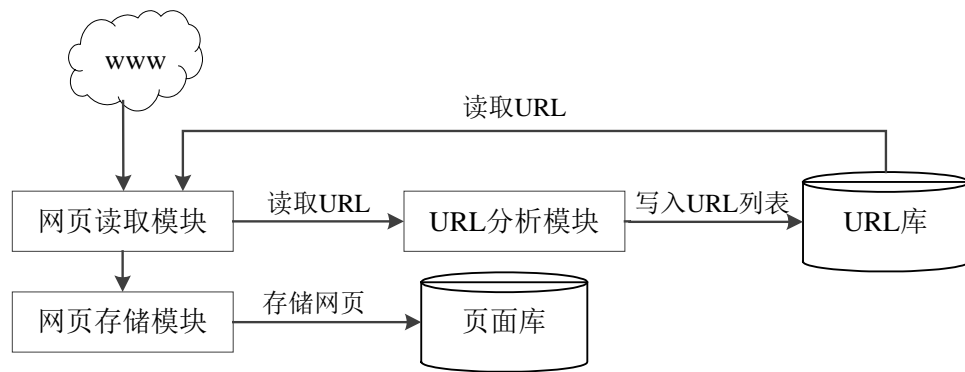


图1-1 网络爬虫工作流程

- (2) 按照 GB3100~3102 及 GB7159-87 的规定，物理量符号、物理常量、变量符号（如： $a(t)$ ， $(i-1)T_h \leq t < iT_h, m, n$ ）用斜体，计量单位（如：km、KB）等符号均用正体。
- (3) 公式居中写，大小与正文相同，公式末不加标点，序号按章节编排，如有“假定，解”字样，文字空两格写，若有对公式变量的说明，以分号结束。公式中用斜线表示“除”的关系时应采用括号，以免含糊不清，如 $1/(b\cos x)$ 。通常“乘”的关系在前，如 $a\cos x/b$ 而不写成 $(a/b)\cos x$ 。具体格式如下：

$$r(t) = R\cos(2\pi f_c t + \theta) \quad (1-1)$$

式中， R 表示幅度； θ 表示相位； f_c 表示载波频率。

- (4) 正文中不能出现大段的空白部分（每章结束处除外）！

第 2 章 标题

2.1 一级节标题

2.1.1 二级节标题

2.1.1.1 三级节标题

注：节最多列到三级，但非必须，需要时才列；如内容过多，还可按条款（见下，最多 3 级）进行组织。如内容不太多，也可以不设三级节，直接用条款来组织。

1. 第 1 级条款标题（退 2 格，不能太长，不能以标点结束，也不能加引用标注）

（1）第 2 级条款内容（首行退 2 格，后面的文字不要悬挂缩进直接顶格换行，以标点结束）

（2）……

（3）……

……

① 第 3 级条款内容（首行退 2 格，后面的文字不要悬挂缩进直接顶格换行，以标点结束）

② ……

③ ……

……

2. 第 1 级条款标题

2.1.1.2 三级节标题

2.1.1.3 三级节标题

2.1.2 二级节标题

注：各级节至少要有 2 个同级的节。

2.1.3 二级节标题

2.2 一级节标题

2.2.1 二级节标题

2.2.2 二级节标题

2.3 本章小结

第 3 章 标题

3.1 一级节标题

3.1.1 二级节标题

3.1.2 二级节标题

3.2 一级节标题

3.2.1 二级节标题

3.2.2 二级节标题

3.3 本章小结

第 4 章 标题

4.1 一级节标题

4.1.1 二级节标题

4.1.2 二级节标题

4.2 一级节标题

4.2.1 二级节标题

4.2.2 二级节标题

4.3 本章小结

第 5 章 标题

5.1 一级节标题

5.1.1 二级节标题

5.1.2 二级节标题

5.2 一级节标题

5.2.1 二级节标题

5.2.2 二级节标题

5.3 本章小结

结 论

注：应客观地总结性说明本论文已经做了哪些方面的工作，各方面又是采用什么方法/手段/技术做了哪些主要内容，取得什么结论/效果；
对之后的展望，应说明在本论文研究工作基础上，今后可进一步研究或完善的问题，或本论文所做工作还需进一步完善的地方，列 2~3 条即可。

致 谢

注：

- （1）查重会覆盖整篇论文（包括致谢），因此所有部分都不得抄袭、复制！
- （2）除对导师感谢外，对于校外毕业设计，也需要感谢校外导师及其所在单位！

参考文献

要求：(1) 篇数 ≥ 20 。

(2) 格式要规范（见下，其中蓝色为示例）。

(3) 英文文献中人名以“姓+名的缩写”方式（参见下述示例）。

(4) 文献应在文中（特别是在 1.2 节）引用！

(5) 序号应按文中标注出现的先后顺序排。

(6) 遵照以下各种类型文献格式标准列出参考文献，多个作者间用逗号（，）分隔。注意各文献各项信息之间的分隔符均应使用英文半角，并加一空格。

- [1] 作者名. 期刊论文题目[J]. 期刊名称, 年, 卷(期): 起止页码.
- [2] 刘东飞, 陈智贤, 邓俊涛, 余俊. 一种辅助翻译搜索引擎的实现[J]. 计算机工程, 2007, 33(1): 216-218.
- [3] Zhang S., Liu L., Diao L.H. Problems on Large-scale Speech Corpus and the Applications in TTS [J]. Chinese Journal of Computers, 2010, (4): 687-696.
- [4] 作者名. 书名[M]. 第 n 版. 出版社名称, 出版年份. （注：如为第 1 版则省略不写版次信息）
- [5] 陶宏才. 数据库原理及设计[M]. 第 3 版. 清华大学出版社, 2014.
- [6] Stallings W. Data and Computer Communications [M]. 10th ed. Pearson Education, Inc., 2014.
- [7] 作者名. 书名[M]. 译者. 第 n 版. 出版社名称, 出版年份. （注：此为译著格式）
- [8] Abraham S., Henry F.K., Sudarshan S. Database System Concepts. 杨冬青 等译. 6th ed. McGraw-Hill Companies, 2012.
- [9] 作者名. 学位论文题目[D]. 博/硕士论文. 学校, 年份.
- [10] 杜亚军. 搜索引擎智能行为的研究及实现[D]. 博士论文. 西南交通大学, 2006.
- [11] 曹桂锋. 搜索引擎中网页分类和网页净化的研究与实现[D]. 硕士论文. 武汉理工大学, 2013.
- [12] 作者名. 会议论文题目[C]. 学术会议论文集名称, 出版年份: 起止页码.
- [13] Zhang X., Dong Y. Designing and Implementing of the Webpage Information Extracting Model Based on Tags[C]. International Conference on Intelligence Science and Information Engineering, 2011: 106-111.
- [14] 作者名. 标准名称[S]. 标准机构名称, 年份.
- [15] Beser B., Droms R., Stump G., et al. The User Class Option for DHCP [S]. RFC3004. 2000.

[16] 作者名. 报纸文章名[N]. 报纸名称, 年-月-日.

[17] 王菲. 移动社交 APP 大比拼[N]. 上海金融报, 2014-12-16.

[18] 作者名. 技术报告题目[R]. 机构名称, 年份.

[19] EMC Corporation. Symmetric Remote Data Facility Product Description Guide[R].
EMC Corporation, 2000.

[20] 作者名. 网上文章题目[EB/OL]. 网址, 查阅的年-月-日.

[21] 中国报告大厅. 2015 年全球智能手机用户数量预测数据分析 [EB/OL].
<http://www.chinabgao.com/stat/stats/39758.html>, 2015-3-24.

文献类型和标志代码对照表（参考后须删除此表）

| 文献类型 | 标志代码 |
|-------|------|
| 普通图书 | M |
| 会议录 | C |
| 汇编 | G |
| 报纸 | N |
| 期刊 | J |
| 学位论文 | D |
| 报告 | R |
| 标准 | S |
| 专利 | P |
| 数据库 | DB |
| 计算机程序 | CP |
| 电子公告 | EB |

附录 1 标题

注：对需要收录于毕业设计中且又不适合在正文中书写的附加数据、资料、详细公式推导等有特色的内容，可作为附录排写。

附录 2 标题