



北京邮电大学

BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

计算机应用编程

实验

熊永平@网络技术研究院

ypxiong@bupt.edu.cn

周五10:10-12:00@3-130

2014.9.12



关于我



智感星际
SmartSensingStars



MobiCom



Nokia Research Center

计算机网络与信息安全

工信部电子信息产业发展研究院



Wuxi SensingNet Industrialization Research Institute



哈尔滨工业大学

ISCAS



中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES

软件学报 Journal of Software

论文排行 投稿指南 问题解答 下载区 收费标准 在线投稿

全文下载排行(总排行, 年度排行, 各期排行)
摘要点击排行(总排行, 年度排行, 各期排行)

全文下载Top 15

- 任圭原, 曹滨宁, 林闯, 无线传感器网络, 2003, 14(7):1282-1291 [47673]
- 陈康, 郭利民, 云计算:系统实例与研究现状, 2009, 20(5):1337-1348 [20756]
- 冯晋国, 张毅, 张研, 徐雷, 云计算安全研究, 2011, 22(1):71-83 [15646]
- 孙吉贵, 刘杰, 赵海宝, 聚类算法研究, 2008, 19(1):48-61 [14859]
- 杨博, 刘大有, Liu Jiming, 金弟, 马清亮, 复杂网络聚类方法, 2009, 20(1):54-66 [12436]
- 陈前峰, 何国祥, 曹建农, 内容传递网络处理能力受限代理缓存算法, 2007, 18(1):146-156 [10480]
- 张亨, 杨家洪, 李健王, Internet 流量模型分析与评论, 2011, 22(1):115-131 [9296]
- 李丽君, 刘鹏飞, 杨相元, 曹晓彬, 曹晓彬, 车用自组网信息广播, 2010, 21(7):1620-1634 [8421]
- 苏贵生, 卢慧群, 陈丽琼, 刘冬梅, 基于Petri网的业务组合故障诊断与处理, 2010, 21(2):231-247 [8360]
- 李明树, 杨秋松, 霍健, 软件过程建模方法研究, 2009, 20(3):524-545 [7641]
- 王福豹, 史力, 任圭原, 无线传感器网络中的自身定位系统和算法, 2005, 16(5):857-868 [7522]
- 吴剑峰, 刘伟强, 王天然, 赵吉亮, 一种带噪声的密集三角网络细分曲面拟合算法, 2007, 18(2):442-452 [7452]
- 甄朝辉, 陈海, 陈贵清, 一种统计多路经的移动自组网路由策略, 2004, 15(4):594-603 [7335]
- 王冀浩, 孙伟东, 周松, 裴晓强, 李小勇, 云计算环境下的分布存储关键技术, 2012, 23(4):962-986 [7224]
- 公茂恩, 焦李成, 杨晓峰, 马文萍, 进化多目标优化算法研究, 2009, 20(2):271-289 [7205]

摘要点击排行Top 15

- 任圭原, 曹滨宁, 林闯, 无线传感器网络, 2003, 14(7):1282-1291 [9937]
- 陈康, 郭利民, 云计算:系统实例与研究现状, 2009, 20(5):1337-1348 [9272]
- 冯晋国, 张毅, 张研, 徐雷, 云计算安全研究, 2011, 22(1):71-83 [8134]
- 公茂恩, 焦李成, 杨晓峰, 马文萍, 进化多目标优化算法研究, 2009, 20(2):271-289 [7479]
- 孙吉贵, 刘杰, 赵海宝, 聚类算法研究, 2008, 19(1):48-61 [7182]
- 孙宝林, 林, 蒋, 冰, 叶雪军, Ad Hoc网络多路经需求路由及路径清洗算法, 2008, 19(2):112-120 [6680]
- 吕帅, 刘磊, 石萍, 李军, 基于自动推理技术的智能规划方法, 2009, 20(5):1226-1240 [6437]
- 李明树, 杨秋松, 霍健, 软件过程建模方法研究, 2009, 20(3):524-545 [5932]
- 杨博, 刘大有, Liu Jiming, 金弟, 马清亮, 复杂网络聚类方法, 2009, 20(1):54-66 [5780]
- 郭少华, 李建中, 高宏, 张研, 从不稳定图中挖掘频繁子图模式, 2009, 20(11):2965-2976 [5641]
- 杨博, 刘大有, 林闯, 无线传感器网络, 2003, 14(7):1282-1291 [5444]
- 熊永平, 孙利民, 牛建伟, 刘燕, 机会网络, 2009, 20(1):124-137 [5095]
- 孙吉贵, 刘杰, 赵海宝, 聚类算法研究, 2008, 19(1):48-61 [4938]
- 孔晓, 赵春雷, 可视化语言技术在软件开发中的应用, 2008, 19(8):1902-1919 [4901]
- 王福豹, 史力, 任圭原, 无线传感器网络中的自身定位系统和算法, 2005, 16(5):857-868 [4760]

物联网核心技术丛书

物联网
技术应用

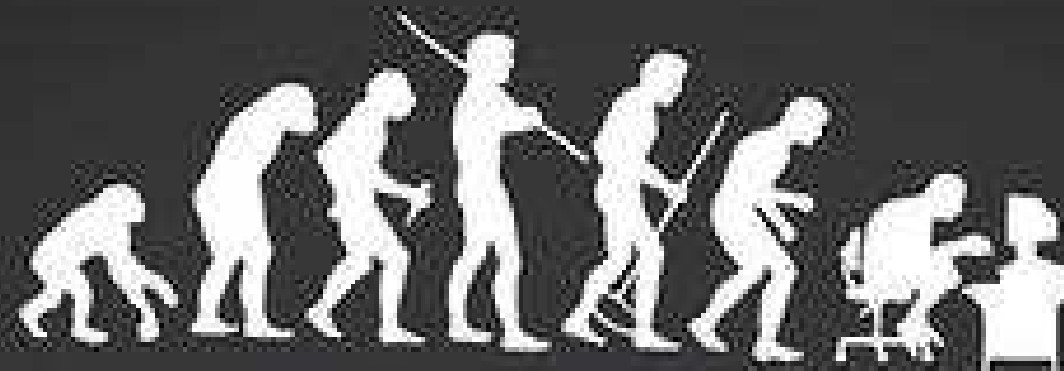
主编 熊永平
副主编 马建 熊永平

① 第一本全方位阐述物联网技术在行业级和消费级应用的书
② 全书由来自20位来自工业界和学术界的专家共同完成
③ 以各领域应用需求为主线, 介绍物联网前沿技术和应用案例
④ 为重点领域应用示范推进和物联网技术产业化提供参考
⑤ 本书编写得到了“感知中国”物联网联盟的大力推动和支持

机械工业出版社
CHINA MACHINE PRESS

本页面中的内容受版权保护

为啥要上编程课？



It is our festival

1024





码农？
中关村夫？

换一个角度



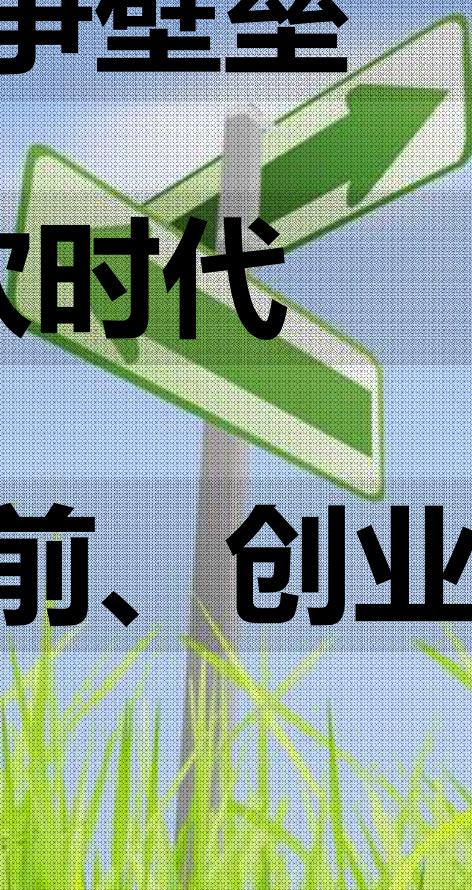


为啥要学编程？(2)

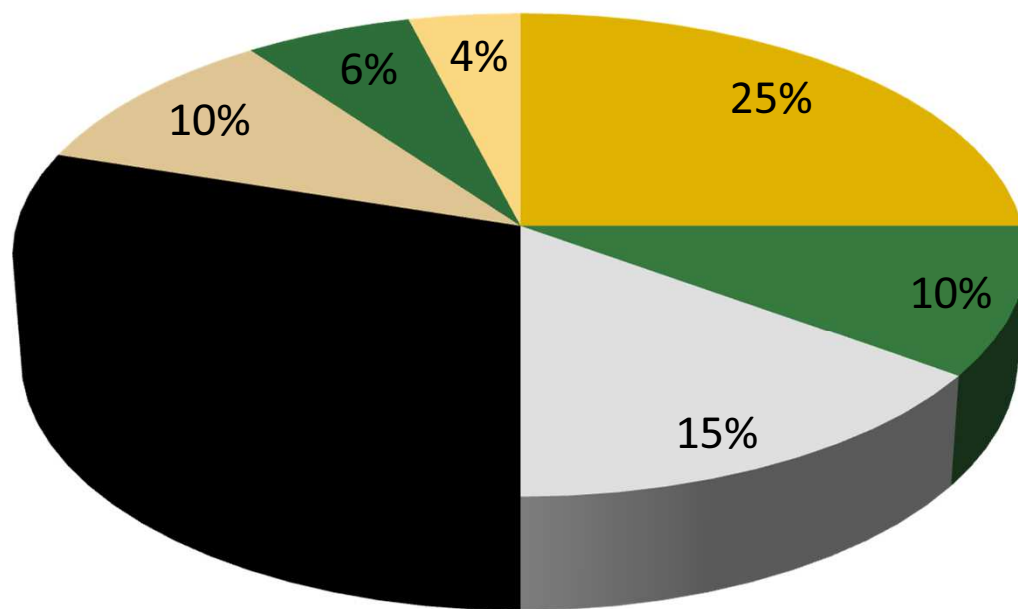
供求关系和竞争壁垒

鲍尔默的微软时代

技术转市场、售前、创业



据不完全统计.....



BUPT信息通信方向硕士毕业去向

■ 互联网公司(16w+)

■ 软件公司(13w+)

■ 运营商(14w+)

■ 通信设备(13w+)

■ 金融IT(10w-30w)

■ 其他IT(unkown)

■ 转行(读博、投行、咨询、公务员...)

看来编程还挺重要的，
你这课能教会我吗？



A photograph of a sandy beach with several footprints leading from the foreground towards the horizon. The sand is a warm, golden-brown color. The footprints are arranged in a line, suggesting a path. The background shows a clear, light-colored sky.

课程**期望**

探索职业定位

提高编程技能

巩固大学基础

Practice makes perfect!

“这门课已经很好了，个人认为研究生阶段做值得上的一门课....”

课程很锻炼人，是研一期间最货真价实的课程”

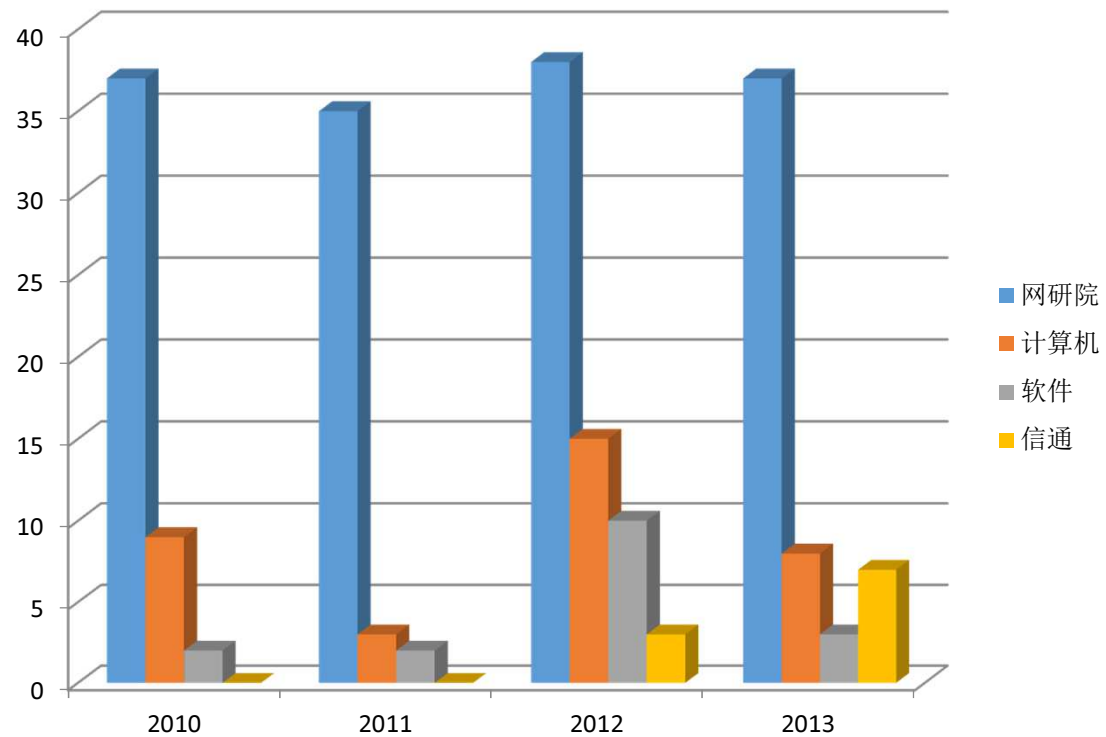
我最大的感觉就是计算机网络编程真的让我学到了很多知识，它是我这学期收获最大的两门课之一

三个作业涉及函数编程，对象编程，过程编程，强化了过去遗留的知识模糊点，掌握了新的知识。对于一些同学来说三个作业的工作量可能有点大，但有压力才有动力，我还是很适应这个工作量。

不夸张的说，这门课在大家心目中是这学期感受最深的一门课，虽然课时不多，但是基本平时的自习时间最多的还是在进行这门课的编程。这门课对大家的编程水平进行了严格的考验。

“没有想到比讨论课更好的形式”

往年选课学生专业分析



这门课**不**学习这些

- 前端开发(PHP/ASP/AJAX/ROR/JSP....)
- 企业信息化系统开发 (Java EE/.Net/Oracle/DB2..)
- 物联网现在很火，基于ARM/51/MSP430/的智能传感器设备.....
- 移动互联网很火，Android/IOS/WP的移动app开发

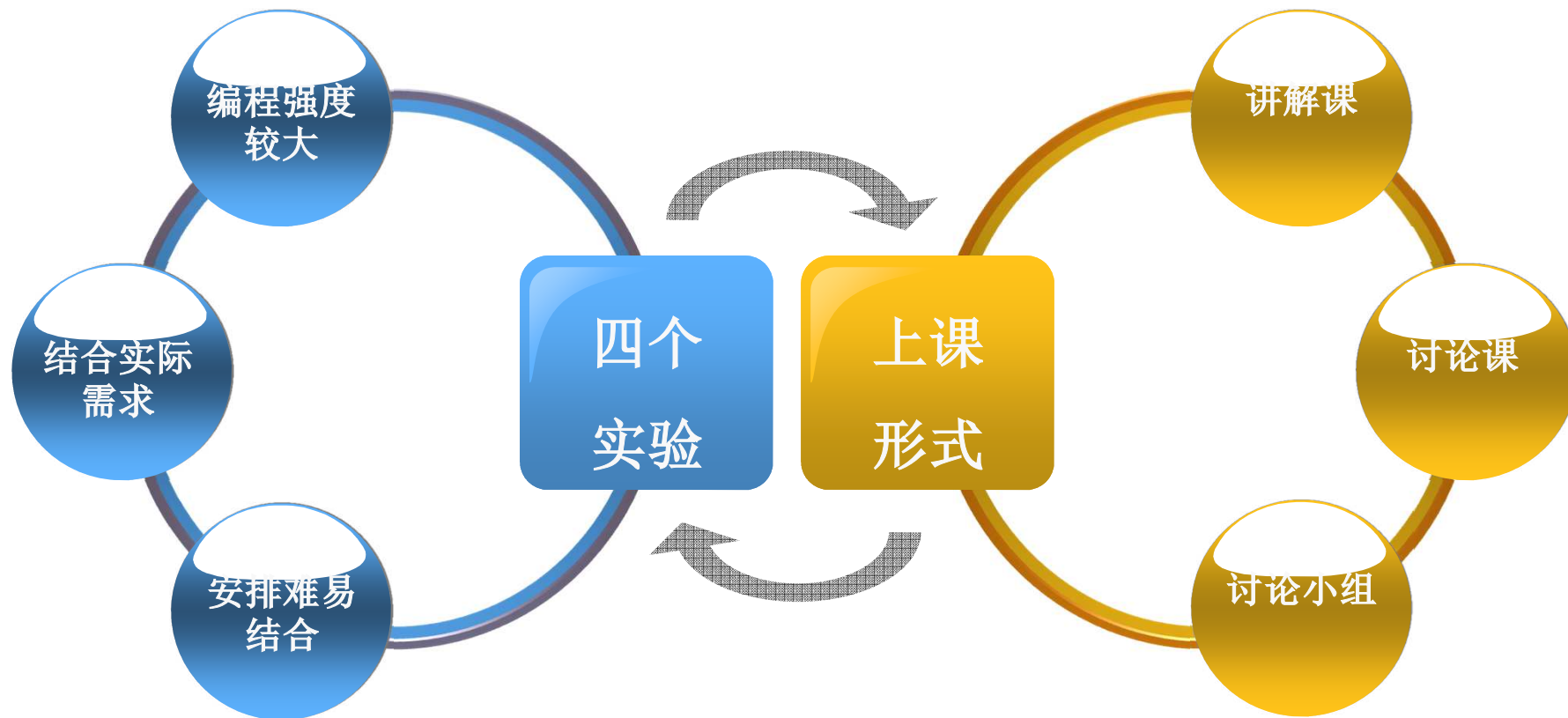


这门课**训练**

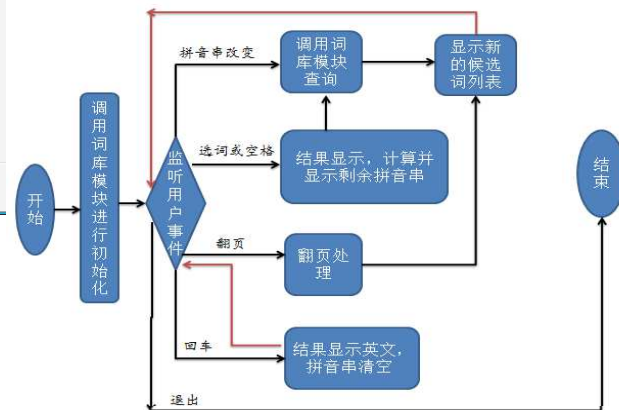
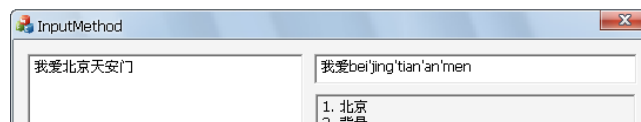
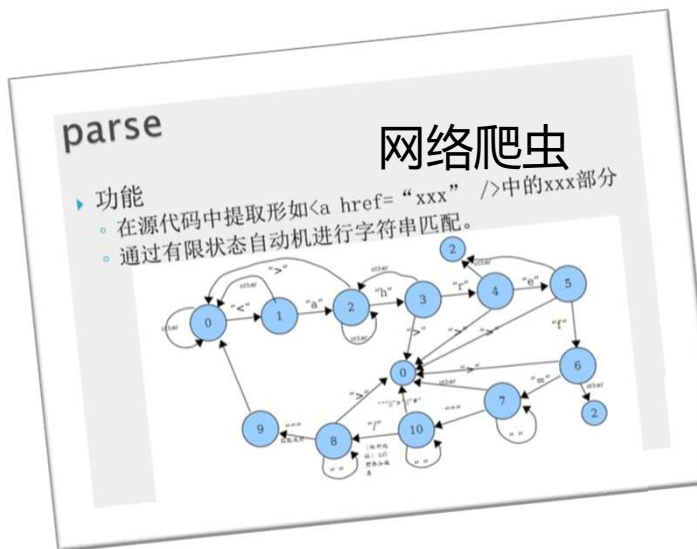


- 计算机系统结构
- 计算机网络原理理解与应用
- 大规模数据处理
- 高级数据结构
- 算法与性能分析

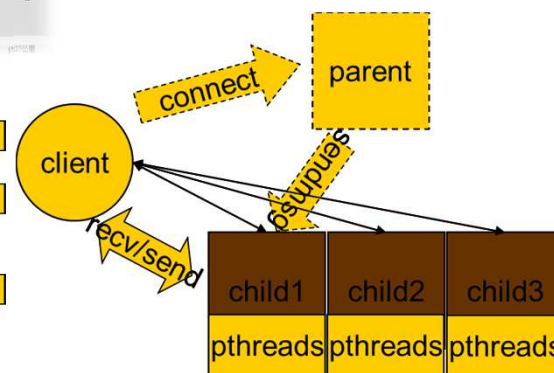
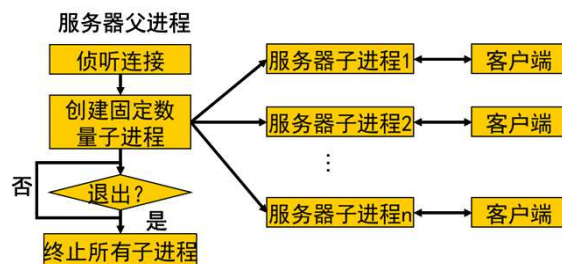
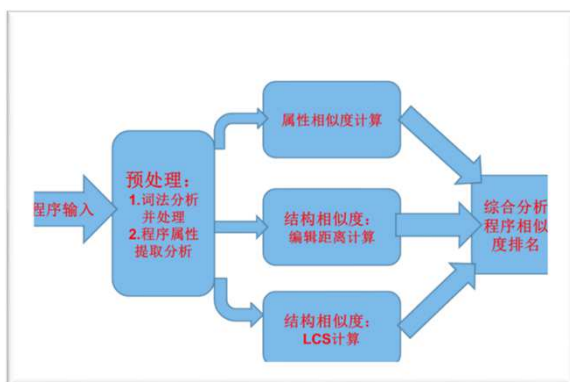
课程方案



以往的实验



121.194.0.224	Apache
220.181.82.213	Apache/2.2.3 (Red Hat)
113.31.30.114	Microsoft-IIS/6.0
118.192.129.234	Microsoft-IIS/6.0
118.192.65.104	Apache/2.2.15 (Red Hat)
123.183.208.215	Microsoft-IIS/6.0
119.97.198.165	Microsoft-IIS/6.0
210.47.0.207	Apache/2.2.8 (Ubuntu) mod_ssl/2.2.8 OpenSSL/0.9.7a PHP/5.2.4
60.173.12.37	Microsoft-IIS/6.0
125.46.13.227	Microsoft-IIS/6.0
221.234.38.221	Apache-Coyote/1.1
218.7.37.194	Apache-Coyote/1.1
61.168.251.140	Microsoft-IIS/6.0
60.173.195.221	Apache/2.0.63 (Win32) PHP/5.2.14
60.173.9.211	Microsoft-IIS/6.0
222.186.34.163	Microsoft-IIS/6.0
58.215.133.51	Microsoft-IIS/6.0
222.186.31.226	Microsoft-IIS/6.0
221.8.18.15	nginx
202.101.233.67	Apache-Coyote/1.1
222.191.251.22	Microsoft-IIS/6.0



课程表

课程表																											
(2014—2015学年 秋季学期)																											
学 期	秋季学期													寒 假													
年 份	二 〇 一 四 年													二 〇 一 五 年													
月 份	八 月	九 月				十 月				十 一 月				十 二 月				一 月				二 月					
周 次	〇	一	二	三	四	五	六	七	八	九	十	十一	十二	十三	十四	十五	十六	十七	十八	十九	二十	廿一	廿二	廿三	廿四	廿五	廿六
星期一	25	1	实验一	15	实验二	29	实验三	13	20	27	3	10	17	24	1	8	15	22	29	5	12	19	26	2	9	16	23
星期二	26	实验一	实验二	16	23	30	实验三	14	21	28	4	11	18	25	2	9	16	23	30	6	13	20	27	3	10	17	24
星期三	27	实验一	10	17	24	实验二	8	15	22	29	5	12	19	26	3	10	17	24	31	7	14	21	28	4	11	18	25
星期四	28	4	11	18	25		9	16	23	30	6	13	20	27	4	11	18	25	实验三	8	15	22	29	5	12	实验一	26
星期五	29	实验一	12	19	26		10	17	24	31	7	14	21	28	5	12	19	26	实验二	9	16	23	30	6	13	20	27
星期六	30	6	13	20	27		1	8	15	22	29	5	12	19	26	3	10	17	24	31	7	14	21	28	25	1	8
星期日	31	7	14	21	28	实验三	2	9	16	23	30	6	13	20	27	4	11	18	25	1	8	15	22	29	5	12	
注：①校历若有临时调整，以党政办的通知为准。②本科生开学典礼、研究生新生报到日期为2014年9月4日。																											

注：①校历若有临时调整，以党政办的通知为准。②本科生开学典礼、研究生新生报到日期为2014年9月4日。

课程介绍

实验一讨论课

实验二讨论课

实验三讨论课

实验四讨论课
论结束

四个实验

多线程、网络编程、索引树.....

1. 海量字符串

处理大规模字符串

2. 并行爬虫

分布式并发爬取
站点所有页面

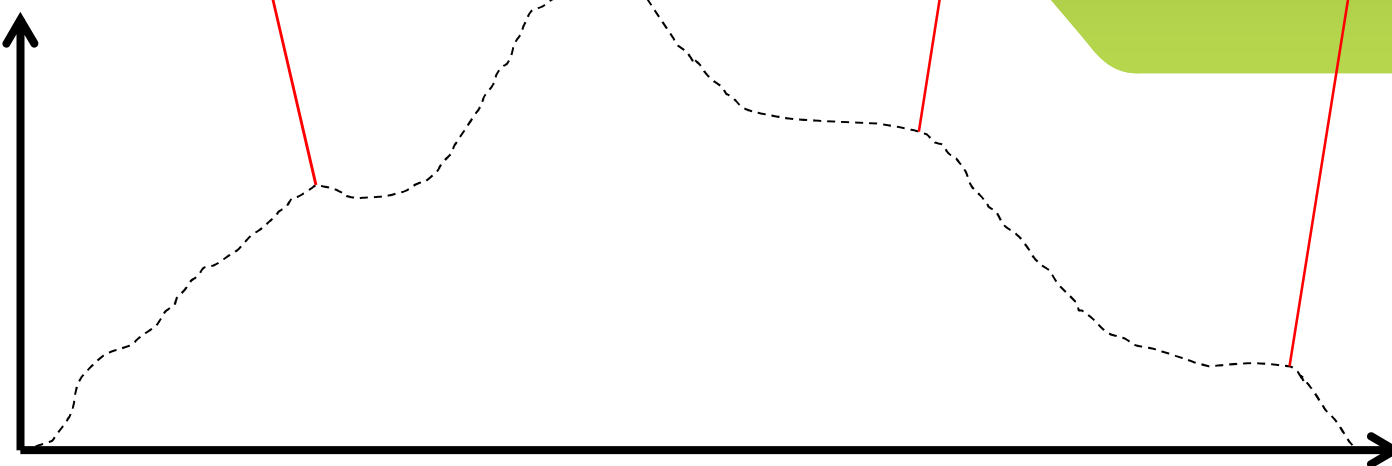
3. 页面分析器

大规模链接图算
法, 社会网络与
矩阵运算

4. 文本处理

基于NLP的文档
相似性分析

难度
系数



➤ 实验任务1

- ❑ 检索一个特定的字符串是否出现在海量字符串集合中
- ❑ 2000万个目标字符串
- ❑ 应用于拼写检查字典、公安疑犯检索/GFW、网络爬虫URL去重、垃圾邮件过滤、IDS特征串分析、封包路由转发.....

➤ 训练技能

- ❑ C编程
- ❑ 大规模字符串存储和检索
- ❑ 字符串匹配算法BM、Trie树、Bloomfilter
- ❑ Hash算法
- ❑ 大量动态内存管理
- ❑ 用数学分析算法性能



➤ 实验任务2

- ❑ 设计一个并发爬虫程序，实现
- ❑ 爬取目标网站页面
- ❑ 去掉无效页面和链接
- ❑ 存储本地

➤ 训练技能

- ❑ C编程
- ❑ 多线程与线程池
- ❑ HTTP协议
- ❑ IO复用高级网络编程
- ❑ 分布式系统设计



➤ 实验任务3

- ❑ 任务2爬取的页面提取链接关系
- ❑ 利用图模型和矩阵对链接网络进行分析
- ❑ 寻找最重要的网页

➤ 训练技能

- ❑ C++实现
- ❑ 大规模复杂网络分析算法
- ❑ 图分析
- ❑ 稀疏矩阵
- ❑ 高级代数原理与应用



➤ 实验任务4

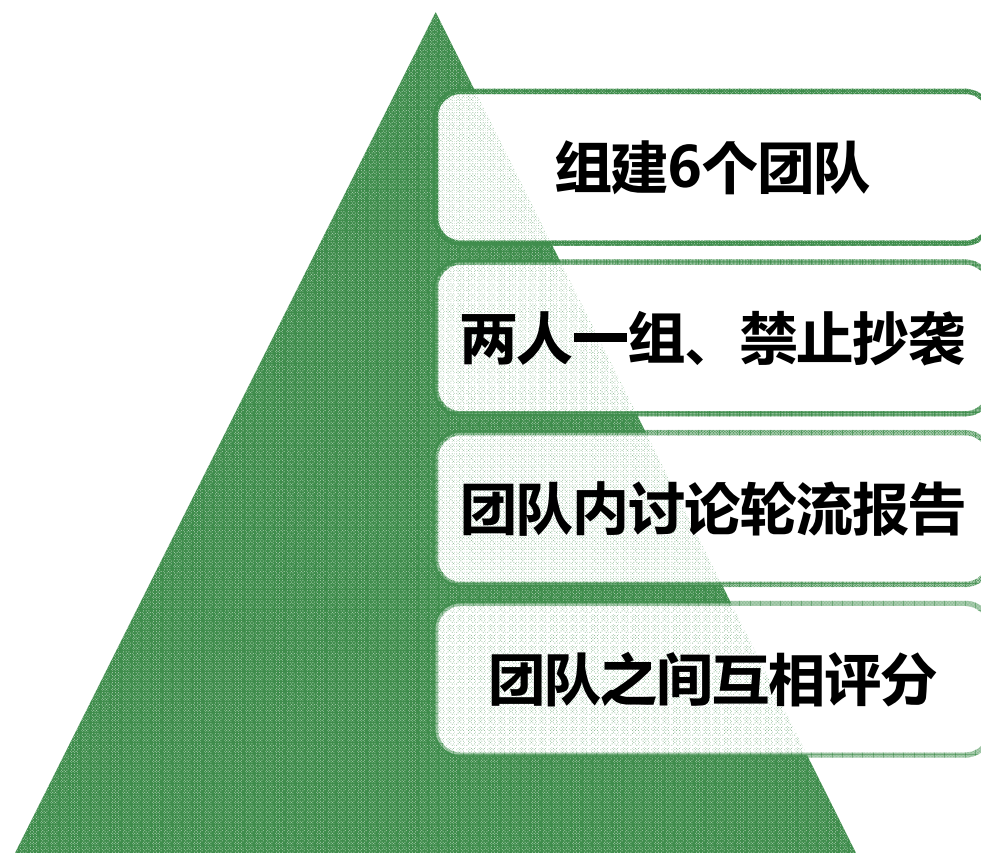
- ❑ 给定巨大的中文文本语料库
- ❑ 给定若干篇文档
- ❑ 对文档进行相似度分析

➤ 训练技能

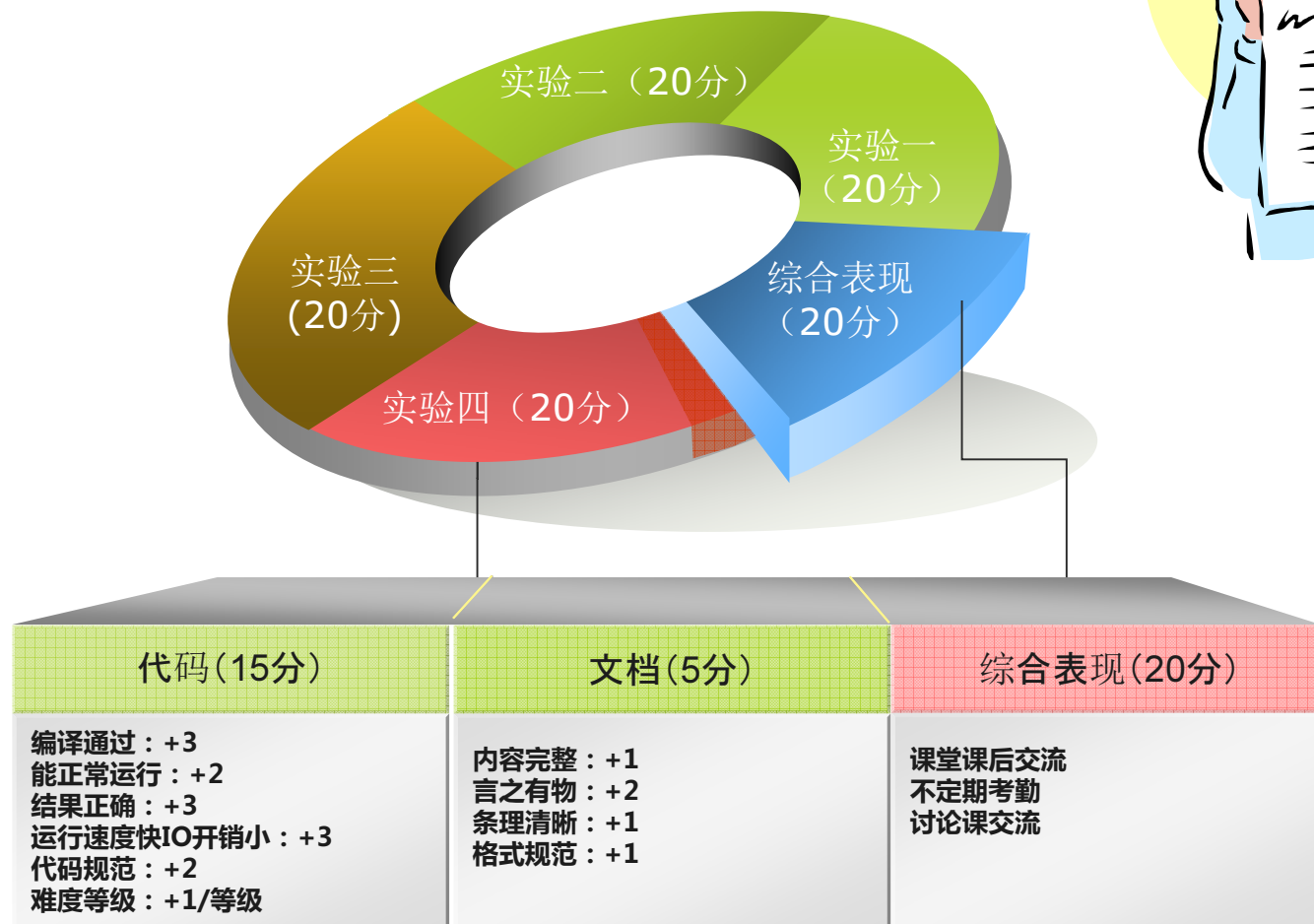
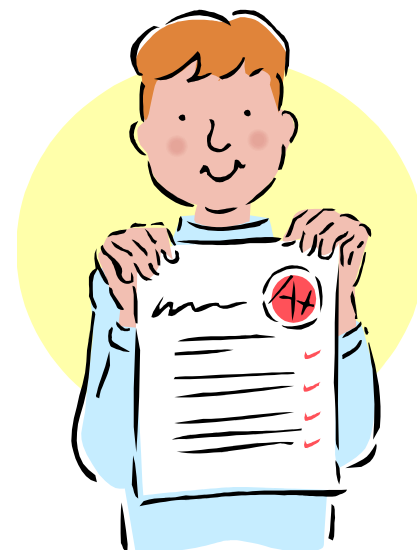
- ❑ Java实现
- ❑ 中文编码理解
- ❑ 中文分词
- ❑ 基于概率模型的特征词抽取
- ❑ 空间向量模型
- ❑ 数据挖掘算法
- ❑ 初步的大数据分析



关于讨论课和程序提交



考核方式



编程环境工具

➤ 操作系统

- ❑ Linux操作系统

➤ 编辑器

- ❑ Source Insight或合适的

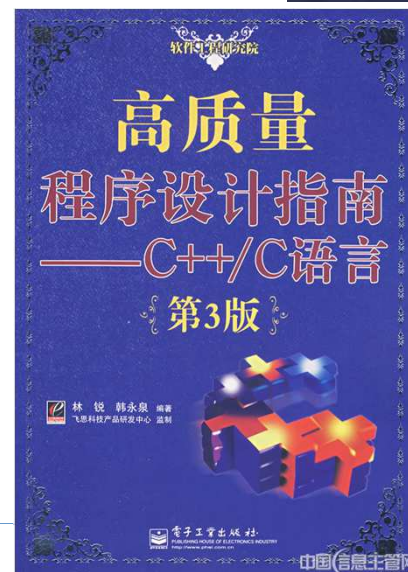
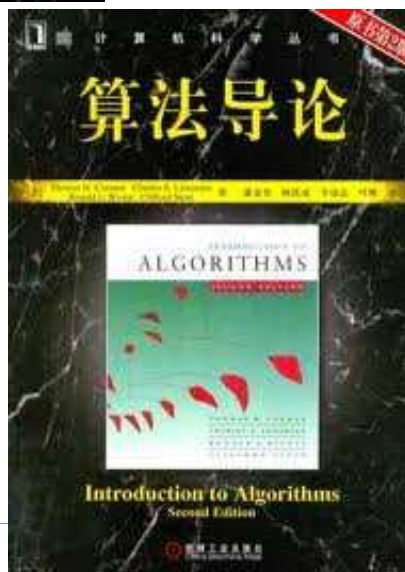
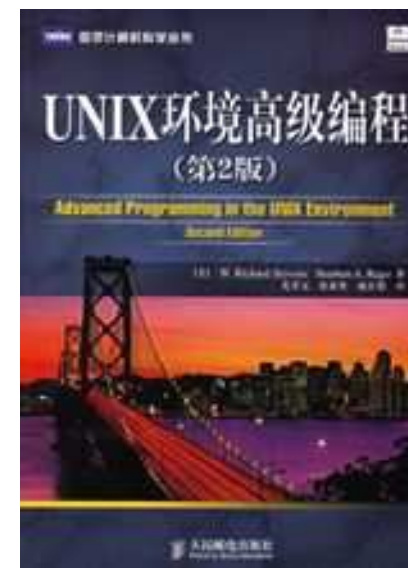
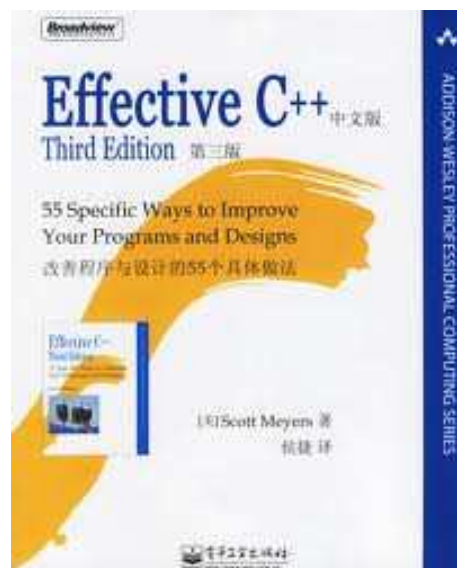
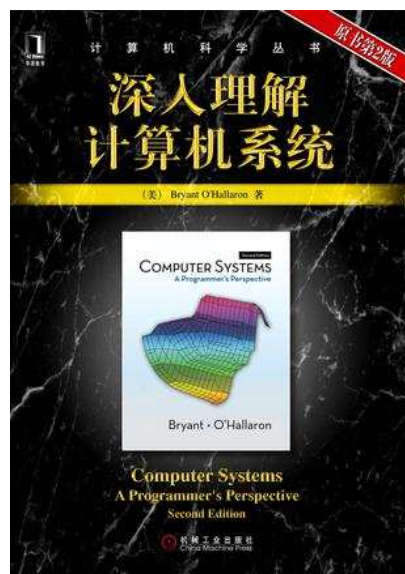
➤ 编译和调试

- ❑ GCC/G++
- ❑ gdb
- ❑ Makefile

➤ 使用的库

- ❑ STL

参考书籍



课程特色



大量编程实践

团队化互动学习

IT相关职业分享



THE END