

实验 8

1 工作量统计

1.1 需求分析

需求分析阶段，我们的大致分工是，查看源码，与网络日志博客，分析出 scrapy 的相关需求，由专门负责文档编写的同学完成文档的整体编写，保证了文档的统一性与一致性，后续需要修改的地方由各同学在需求分析规格书的基础上继续修改。

（1）需求分析阶段的工时统计

需求分析阶段的统计分工如表 1.1

表 1.1 需求分析阶段的工时统计

组员	任务分工
胡勇	初步需求讨论、汇报ppt准备
郭炜锋	需求细化
王益飞	微信会议记录
武丁泽宇	软件需求规格说明书初稿

（2）需求规格说明书编写分工

全文由武丁泽宇同学整理发布，全文字数 8340，关于需求规格说明书的工作分工如表 1.2

表 1.2 需求规格说明书任务划分

姓名	完成图表个数	工作重点
胡勇	用例图 RUCM图2个	全文校对、排版
郭炜锋	RUCM图2个	文档整理
王益飞	RUCM图2个	组织统筹
武丁泽宇	全文编辑书写	全文

（3）需求规格说明书版本追踪

本需求规格说明书共经历 13 个版本，需求规格说明书的最终版本为 v1.95，关于需求规格说明书的过程追踪如表 1.3 所示

表 1.3 需求规格说明书追踪

版本号	日期	所修改章节	修改说明	修改人	审核人
1	2017.3.21		完成第一版	武丁泽宇	胡勇
1.1	2017.3.29	4	修改第4章内容	武丁泽宇	胡勇
1.11	2017.3.29	4、5	修改4章内容,删除5章	武丁泽宇	胡勇
1.2	2017.4.4	1、2、3、4	按照修改建议进行修订	武丁泽宇	胡勇
1.3	2017.4.5	3	完善业务需求和用例图	武丁泽宇	胡勇
1.4	2017.4.9	全文	排版修改	胡勇	胡勇
1.5	2017.4.16	1.3、全文字体、全文格式	将1.3定义使用表格表述,将正文的格式统一	郭炜锋	胡勇
1.6	2017.4.23	全文修改	根据老师批注进行修改	武丁泽宇	胡勇
1.7	2017.4.26	全文修改	根据各组评审意见进行修改和调整	武丁泽宇	胡勇
1.75	2017.5.3	批注章节修改增加第5章	按照批注进行修改完善	武丁泽宇	胡勇
1.8	2017.5.4	全文	格式化文本格式	胡勇	胡勇
1.9	2017.5.11	全文	修改RUCM和用例图	武丁泽宇	胡勇
1.95	2017.5.29	第5章	增加改进需求	武丁泽宇	郭炜锋

1.2 需求评审

需求评审又分为需求评审与需求复评审,其中还掺杂部分需求修改与实验 6、7、8 的第一次工作内容,统计分析下

(1) 分工统计

这一阶段的大致分工与实际工时统计如表 2.1 所示

表 2.1 统计分工

组员	任务分工
胡勇	RUCM补充完善、汇报ppt编写、个人评审需求
武丁泽宇	补充业务需求、修改需求规格说明书、个人需求评审
郭炜锋	微信会议记录、实验8、个人需求评审
王益飞	完善后期计划、分配任务、个人需求评审

(2) 对其他组的评审统计

这里主要整理我组对其他组的评审统计,如表 2.2 所示,关于项目的计划的评审不在此统计范围内,一共整合两次评审的结果。

表 2.2 评审次数统计

评审人	评审对象	意见个数	意见总个数
胡勇	Blade	6	13
	Dex2jar	7	
武丁泽宇	Blade	4	8
	Dex2jar	4	
王益飞	Blade	6	9
	Dex2jar	3	
郭炜峰	Blade	5	11
	Dex2jar	6	

1.3 改进与展示

(1) 改进阶段的分工

按每个人的任务，统计各组员在这一阶段的工作分工如表 3.1

表 3.1 改进阶段的分工说明

组员	任务分工
胡勇	提交前端可视化实现、提交改进方案、汇报PPT
武丁泽宇	整理文档，跟踪报告
郭炜峰	后端实现，越过反爬机制实现，提交改进方案，课堂问题记录
王益飞	爬虫与分布式爬虫实现，布隆过滤器实现，提交改进方案

(2) 代码输出

这一阶段的文档产出与代码产出的统计，如表 3.2 所示：

表 3.2 文档与代码输出统计

编号	编号	子系统名称	构件名称	总代码行数		涉及章节		描述方式	简述/备注	负责人
1	1.1	过滤器	布隆过滤器	165	74行	(A)1、2	(A) 1.2	文字说明	涉及到filters.py	王益飞
							(A) 2.2.2			
	1.2		分布式过滤器		91行		(A) 1.2		涉及到dupefilters.py	
							(A) 2.2.1			
2		调度器		365行		(A) 2.1.3		图简表	涉及到mongodb_agent.py	
									dupefilters.py	
									scheduler.py	
3	3.1	网站画像	基础 爬虫部分	924行	14行	(A) 2.1.4		图简表	代码行数包含3.1，3.2,3.3	
									该部分代码和调度器用到的mongodb代理是同一个程序文件。	
	3.2		HTTP后端部分	370行		(B)2.2	(B)2.1 (B)2.2	文字说明，接口描述		郭炜峰
	3.3		前端展示部分	540行		(C) 1改进说明	(C) 1改进说明	文字说明		胡勇
						(C) 技术说明	(C) 技术说明			
4		反爬扩展	反扒扩展	120行		(C)展示说明	(C)展示说明	文字说明		郭炜峰
						(B) 1.1	(B) 1.1			
						(B) 1.2	(B) 1.2			

(3) 文档输出

各同学的输出的改进说明文档工作量统计如表 3.3 所示

表 3.3 文档输出统计

组员	输出文档	字数统计
胡勇	搭建运行环境说明、爬虫改进部分3	1591
郭炜峰	需求细化	2032
王益飞	微信会议记录	1514
武丁泽宇	整理输出文档	

1.4 软件测试需求

(1) 本阶段分工如表 4.1 所示

表 4.1 本阶段任务分工

组员	任务分工
	提交个人测试需求报告
胡勇	制作、汇报ppt
	完成改进部分3的测试工作
武丁泽宇	整理输出文档
	课堂问题记录
郭炜峰	完成改进部分2的测试工作
	提交个人测试需求报告
王益飞	统筹会议，分配任务
	完成改进部分1的测试工作
	提交个人测试需求报告

(2) 测试需求分析

各同学提出了个人的测试需求用例与输出文档，统计如表 4.2 所示

表 4.2 测试需求统计

组员	输出文档	测试用例个数
胡勇	Scrapy改进部分3测试需求分析	5
	改进部分3_测试用例和测试项对照表	
郭炜峰	Scrapy改进部分2测试需求分析	2
	改进部分2_测试用例和测试项对照表	
王益飞	Scrapy改进部分1测试需求分析	6
	改进部分1_测试用例和测试项对照表	

(3) 测试需求规格说明书

测试需求规格说明书由武丁泽宇同学编写，共 13 个测试用例，4 个版本，共计 7563 个字，版本变更过程如表 4.3 所示

表 4.3 测试需求规格说明书变更追踪

日期	版本	修订说明	修订人	审阅人
5.13	1	首次编写	武丁泽宇	胡勇
5.23	2	加入测试用例和测试项对照关系表	武丁泽宇	胡勇
5.29	3	增加测试用例，修改对照表	武丁泽宇	胡勇
6.14	4	增加对应需求的说明	武丁泽宇	胡勇

1.5 测试评审与测试结果

(1) 本阶段分工如表 5.1 所示

表 5.1 本阶段任务分工

组员	任务分工
胡勇	完成B组组间互测
	制作、汇报ppt
	关于AB组的测试需求评审
武丁泽宇	关于AB组的测试需求评审
	整理输出文档
郭炜锋	关于AB组的测试需求评审
	课堂问题记录
	完成B组组间互测
王益飞	统筹会议，分配任务
	关于AB组的测试需求评审
	完成B组组间互测

(2) 关于 AB 组的评审报告

这里主要整理我组对其他组的测试需求评审统计，如表 5.2 所示

表 5.2 测试需求互评统计

评审人	评审对象	意见个数	意见总个数
胡勇	Blade	3	6
	Dex2jar	3	
武丁泽宇	Blade	5	7
	Dex2jar	2	
王益飞	Blade	6	8
	Dex2jar	2	
郭炜峰	Blade	3	5
	Dex2jar	2	

(3) 组间互测

我组主要完成对于 B 组的部分测试，输出文档与 B 组测试报告汇总，如表 5.3 所示

表 5.3 关于 B 组的测试工作

组员	输出文档	测试用例个数
胡勇	组间互测_B组测试报告_1_胡勇	8
郭炜峰	组间互测_B组测试报告_2_郭炜峰	10
王益飞	组间互测_B组测试报告_3_王益飞	8
武丁泽宇	测试报告汇总	共26个测试用例，1800字

以上就是我组对于主要工作的分工统计，其他工作内容由各同学协同完成。不在此赘述。

2 计划与统计

2.1 计划分工

计划分工如表 6 所示

表 6 小组工作分配表

姓名	性别	擅长	分工	职位
王益飞	男	Python，分布式系统，统筹能力	部分代码工作，每周组织会议分配任务，工作日志mpp维护	组长
郭炜锋	男	Django网站后台，数据库SQL语句、文案记录	部分代码工作，每节课的课堂日志记录	组员
武丁泽宇	男	文档能力、文字编写能力	各实验文档编写，跟踪总结，修改整合	组员
胡勇	男	PPT制作、软件工程基础、前端数据可视化	部分代码工作，每周PPT制作汇报，用例图实现	组员

2.2 总工作量统计

项目成员，王益飞、郭炜锋、胡勇、武丁泽宇，共四人。从已经花费的时间来看，郭炜锋花费的时间最长，大于两百工时。其他三人的实际花费时间几乎相同。在两百工时以下。从剩余工时来看，四人的剩余工时几乎相同。

观察四人的剩余工时的变化可以发现，由于项目每周都进行具体和细致的任务分配，所以实际每周都会增加一部分公式。从项目开始到 4 月 9 日，剩余工时的变化基本呈下降趋势。在 4 月 9 日有一个跃迁，说明 4 月 9 日增加了一些新的工作任务。4 月 23 号有开始下降，直至 5 月 21 日，又出现了一个跃迁。总的工作量统计如图 1 所示

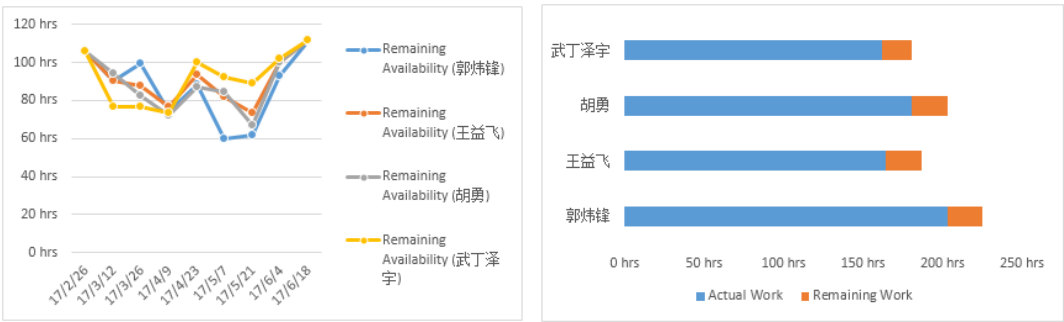


图 1 工作时间

3 分析

3.1 总结

从上文所总结的统计报告来看，各同学各尽所能，从各自擅长的方向参与软件工程的整个过程。从时间统计数据 and 实际工作分工可以看出，各同学任务量大致相当，结果满足基本要求。

3.2 经验

（1）组员间的默契程度能提高工作效率，在工作开始阶段，各同学相互之间不大熟悉，任务分配不具体，导致项目初期，进展缓慢偏离预期，在同学间相互磨合熟悉后，能够快速定位个人在项目组中的位置，找到适合自己的工作，交流习惯以后，相互形成默契，做工作方面各尽所能，大大提高了工作效率与项目进度

（2）出现问题应当及时总结，本组初期，核心需求模糊，工作量偏小，在与老师交流与线下会议后，能够及时调整个人工作进度，提高整个组的工作量，拉到与其他组别大致相当的位置。项目的进展不可能一帆风顺，只要同理合作，及时沟通，认真总结反思，就能够应对。

（3）代码开发只是软件工程的一部分，其他部分同样重要。从工作量统计来看，项目改机即开发计划只占到了一部分，其他的诸如需求分析、软件测试、文档编写、进度控制等，在软件工程过程中的工作不容小觑，同样重要，甚至由于其他方面的不完善会影响实际项目代码质量与进度完成情况

（4）尺有所长，寸有所短。各组员间能够相互学习，谦虚请教，积极发挥个人的能力，共同学习进步，是本次项目的一大收获。