

软件项目计划书

项目名称：基于 wifi 探针的商业大数据分析技术系统

项目起止时间：2017.3-2017.6

文档编写：黄振洋、张航、许成铭

目录

1 引言	1
1.1 背景	1
1.3 定义	1
1.4 参考资料	2
1.5 标准、条约和约定	2
2 项目概述	2
2.1 项目目标	2
2.2 产品目标与范围	2
2.3 假设与约束	3
2.4 项目工作范围	3
2.5 应交付成果	3
2.5.1 需完成的软件	4
2.5.2 需提交用户的文档	4
2.5.3 须提交内部的文档	4
2.5.4 应当提供的服务	4
2.6 项目开发环境	4
2.7 项目验收方式与依据	4
3 项目团队组织	5
3.1 组织结构	5
3.2 人员分工	5
3.3 协作与沟通	5
3.3.1 项目团队内部协作	5
3.3.2 项目接口人员	5
3.3.3 项目团队外部沟通与协作模式	6
4 实施计划	6
4.1 风险评估及对策	6
4.2 工作流程	6
4.3 总体进度计划	8
4.4 项目控制计划	8
4.4.1 质量保证计划	8
4.4.2 进度控制计划	9
4.4.3 预算监控计划	9
4.4.4 配置管理计划	9
5 支持条件	9
5.1 内部支持	9
5.2 客户支持	9
5.3 外包（可选）	10
6 预算	10
6.1 人员成本	10
6.2 设备成本	10
6.3 其它经费预算	10
7 关键问题	10
8 专题计划要点	10

基于 wifi 探针的商业大数据分析技术系统

软件项目计划书

根据《GB8567—88 计算机软件产品开发文件编制指南》中项目开发计划的要求，结合实际情况调整后的《项目计划书》内容如下：

1 引言

1.1 背景

(1) 项目的名称

基于 wifi 探针的商业大数据分析技术开发

(2) 项目的委托单位

无

(3) 项目的用户（单位）：

线下实体营销商店、大型商场等。

(4) 项目的任务提出者

黄振洋

(5) 项目的主要承担部门

设计承担部门：张航

编码承担部门：黄振洋

测试承担部门：许成铭

(6) 项目建设背景

利用探针数据的客流分析打破模式束缚，不仅仅只是提供可信的客流数据分析，同时还利用延伸的标杆管理才能，深刻洞悉并提供有助于推动实际客流量和消费者习惯行为的一系列因素。这种专业才能呈现了经济分析，社交和环境等一些超出你控制范围的因素，却对商业绩效产生主要的冲击力。

首先开发探针设备能够采集客户唯一的定位标识，比如 MAC 地址，通过数据分析技术，采用离线计算和实时计算结合的方式，为商业环境提供科学的、全面的数据决策依据。不仅对营销能力的评估，也可以对管理上进行优化。

(7) 软件系统与其他系统的关系

本系统需求如下：

数据需求：探针连接路由器的账号密码，云端服务器的 ip，域名

环境：需要有联网的路由器

依赖系统：云端 linux 系统

系统结果反映为电脑桌面应用程序。

(8) 软件系统与机构的关系

委托单位：无

使用单位：线下实体营销店铺

不需要外包

组织机构合作：无

1.3 定义

WiFi 探针技术是指基于 WiFi 探测技术来识别 AP(无线访问接入点)附近已开启 WiFi 的智能手机或者 WiFi 终端(笔记本, 平板电脑等), 无需用户接入 WiFi, WiFi 探针就能够识别用户的信息。当我们走进探针信号覆盖区域内且我们的 wifi 设备打开, 我们的设备就能被探针探测出来, 无论是 IOS 或者安卓系统都能轻易检测到, 并且获取设备的 MAC 地址。

有以下几个特点: 1. 用户无需参与, 无需连接到网络 2. Android, IOS 全兼容 3. 自动探测区域内手机 MAC 地址 4. 手机, 平板均能探测

1.4 参考资料

参考资料: 第六届中国软件杯 A 组 2 题赛题要求。

文档格式要求按照我国 GB/T8567-1988 国家标准和 IEEE/ANSI830-1993 标准规范要求进行。包括以下文件

软件工程项目开发文档范例

软件工程国家标准文档

需求说明书

软件需求说明书编写规范

1.5 标准、条约和约定

本项目遵从以下标准:

GB/T 13702-1992 计算机软件分类与代码

GB/T 20918-2007 信息技术

GB/T 19003-2008 软件工程

GB/T 5538-1995 软件工程标准分类法

GB/T 9386-2008 计算机富安居测试文档编制

GB/T 9385-2008 计算机软件需求规格说明

GB/T 5532-2008 计算机软件测试规范

GB/T 18221-2000 信息技术程序设计语言

GB/T 11457-2006 信息技术 软件工程

GB/T 8567-2006 计算机软件文档编制规范

2 项目概述

2.1 项目目标

总体目标: 通过 WIFI 探针收集 MAC 地址、出现时间、出现地点、与探针距离, 探针设备会定时(3s)发送数据到服务端, 通过一段时间数据平台产生大量的用户数据, 然后使用大数据分析技术, 采用离线计算和实时计算技术相结合的方式, 能够对商业环境中门店的门前人

流量、进店客流量、进入跳出量、新老顾客数量及新增量、在店平均时长进行分析，客流情况更新迅速，便于实时查看线下人群密集程度，有利于总部统一管理并及时调整营销活动方案。新老顾客比例，来访周期，顾客活跃度清晰展示，便于及时掌握入场顾客的构成，为销售策略调整提供参考。通过驻店时长可以快速掌握店招、产品或营销活动的吸引力。

2.2 产品目标与范围

功能目标：主要是实现三个方面的功能：一是通过探针设备采集可监测范围内的手机 MAC 地址、地理位置、与探针距离、时间等信息；二是探针采集的数据可以定时发送到服务端保存；三是利用大数据技术对数据进行人流量等指标的分析。

2.3 假设与约束

项目必须在三月内完成，预计 12 周，人员构成为三人，预算在 500 以内，主要用于购买 wifi 探针的设备，以及云端服务器的租赁。由于开发时间较为短暂，所以需要在日常不断推进，其中重点在于实现数据的图表化显示，满足基本的功能性需求：数据分析平台健壮，数据展现界面友好，探针设备配置简单、运行稳定。

如若开发过程顺利，项目将增加分布式系统架构实现的额外功能。

2.4 项目工作范围

工作范围依照功能目标展开叙述：

WIFI 探针：

探针设备可以在其他渠道（淘宝）购买，也可以自行开发硬件和固件，探针设备可以进行服务端的相关配置（服务端 IP、端口、路径、发送数据时间间隔），能够采集 MAC 地址、地理信息、与探针大概距离、采集时间等信息（不允许采集个人隐私信息，比如手机号等信息），特别针对 ANDROID6.0 和 IOS10 版本后的移动终端设备进行测试能否采集到 MAC 地址。

数据采集

服务端主要接收探针定时发送的数据，将数据保存到数据分析平台待用，文件系统可以使用 HDFS 或者其他适合的分布式文件系统。接收数据不能有数据丢失或者数据失真。

数据分析

基本能够分析以下 9 大指标。

- 客流量：店铺或区域整体客流及趋势
 - 进店量：进入店铺或区域的客流及趋势
 - 进店率：进入店铺或区域的客流占全部客流的比例及趋势
 - 来访周期：进入店铺或区域的顾客距离上次来店的间隔
 - 新老顾客：一定时间段内首次/两次以上进入店铺的顾客
 - 顾客活跃度：按顾客距离上次来访问隔，划分为不同活跃度（高活跃度、中活跃度、低活跃度、沉睡活跃度）
 - 驻店时长：进入店铺的顾客在店内的停留时长
 - 跳出率：进入店铺后很快离店的顾客及占比（占总体客流）
 - 深访率：进入店铺深度访问的顾客及占总体客流（可以根据定位轨迹或者停留时长判定）
- 以上数据指标支持环比和历史对比，并且可以从小时、日、周、月多维度分析。数据分析中的关于范围的界定支持自定义阈值。

2.5 应交付成果

2.5.1 需完成的软件

程序名称：基于 wifi 探针的商业大数据分析技术

编程语言：java, javascript

软件将以 Web 交付给用户。

源程序文件将存储在 github 开源网站上。

2.5.2 需提交用户的文档

用户帮助手册：指导用户安装配置使用该应用。 存储形式：txt

需求规格说明书：在项目开始的应用需求规格。 存储形式：txt

2.5.3 须提交内部的文档

需求分析报告（软件规格说明书）：对所开发软件的功能、性能、用户界面及运行环境等做出详细的说明。

可行性分析报告：说明该软件开发项目的实现在技术上、经济上和社会因素上的可行性，评述为了合理地达到开发目标可供选择的各种可能实施方案，说明并论证所选定实施方案的理由。

软件开发计划书：为软件项目实施方案制订出具体计划，应该包括各部分工作的负责人员、开发的进度、开发经费的预算、所需的硬件及软件资源等。

概要设计文档：该说明书是概要实际阶段的工作成果，它应说明功能分配、模块划分、程序的总体结构、输入输出以及接口设计、运行设计、数据结构设计和出错处理设计等，为详细设计提供基础。

详细说明文档：着重描述每一模块是怎样实现的，包括实现算法、逻辑流程等。

源代码：软件开发过程中的全部代码及注释。

测试计划报告：为如何做好集成测试和验收测试，需为如何组织测试制订实施计划。计划应包括测试的内容、进度、条件、人员、测试用例的选取原则、测试结果允许的偏差范围等。

测试分析报告：测试工作完成以后，应提交测试计划执行情况的说明，对测试结果加以分析，并提出测试的结论意见。

开发进度月报：该月报系软件人员按月向管理部门提交的项目进展情况 报告，报告应包括进度计划与实际执行情况的比较、阶段成果、遇到的问题和解决的办法以及下个月的打算等。

项目开发总结报告：软件项目开发完成以后，应与项目实施计划对照，总结实际执行的情况，如进度、成果、资源利用、成本和投入的人力，此外，还需对开发工作做出评价，总结出经验和教训。

2.5.4 应当提供的服务

本项目为用户提供服务如下：

安装教程服务：提供程序安装使用教程。

维护服务：软件发生不正常错误，或服务器和数据库发生故障时，将提供维护服务。

2.6 项目开发环境

所需硬件设备：wifi 探针，电脑，网络路由器

开发操作系统：win10, linux

开发工具：Eclipse , MySQL, MyEclipse, Dreamweaver, Sublime Text 3

所需第三方库：Echarts，java 图形界面库,Mahout
网络环境：需联网。

2.7 项目验收方式与依据

内部验收方式：

代码验收：最后在提交项目之前进行小组内评审，代码编写符合 HB6465 标准，与文档说明保持一致，代码书写风格统一，采用标准规范，没有下列错误：由于软件缺陷造成丢失数据，不符合设计要求，响应时间太长无法接受等问题。

文档验收：最后在提交项目之前进行小组内评审，文档格式符合 HB6465 标准，功能符合与客户的合同要求，清晰易读，没有语病与歧义。

服务验收：暂无。

说明项目内部验收和用户验收的方式，如验收包括交付前验收、交付后验收、试运行（初步）验收、最终验收、第三方验收、专家参与验收等等。项目验收依据主要有标书、合同、相关标准、项目文档（最主要是需求规格说明书）。

3 项目团队组织

3.1 组织结构

项目负责人:黄振洋

项目主研: 张航 许成铭

其中：黄振洋负责整个项目进度的掌控，张航负责项目设计，许成铭负责编码和测试。

3.2 人员分工

工作内容	负责人	参与人
需求分析报告	张航	黄振洋
可行性分析报告	黄振洋	许成铭
项目计划书	许成铭	黄振洋
概要设计文档	许成铭	张航
详细说明文档	张航	黄振洋
源代码	黄振洋	张航
测试计划报告	许成铭	黄振洋
测试分析报告	黄振洋	许成铭
开发进度月报	黄振洋	张航
项目总结报告	张航	许成铭
软件安装、维护	许成铭	张航
用户帮助文档	张航	许成铭
需求规格说明书	黄振洋	许成铭

3.3 协作与沟通

3.3.1 项目团队内部协作

协作模式：每周两次会议，设计任务，寻找问题，共同解决，每次会议由小组成员轮流做会议记录，上传到小组开发 qq 群。小组成员在各自任务中如果遇到难题，则由组长组织共

同讨论解决方案。

3.3.2 项目接口人员

张航

3.3.3 项目团队外部沟通与协作模式

外部沟通：与指导老师进行沟通、与大型商场相关人员进行需求沟通。

协作模式：在软件开发实践课程中与老师交流，小组成员记录沟通结果。

4 实施计划

4.1 风险评估及对策

任何项目在实现过程中都存在不确定性，所以任何项目都是存在风险的，我们提前预估了本项目在技术，市场，财务，管理上可能遇到阻碍项目进度发展的潜在风险，并针对各项风险给与了相应的解决对策。

技术风险

1. 本项目涉及硬件 wifi 探针，由于市场上的 wifi 探针品种繁多，良莠不齐，低质量的 wifi 探针所获得的数据量和准确程度可能与正常探针所获取的数据差别较大，将对我们的项目后续的分析处理产生误差和错误导向。
2. 大量数据的分析处理和并行计算对于技术方面存在挑战，对于处理实时的数据所涉及到的技术较为复杂，可能对项目开发有所影响
3. 由于本项目采用集群的形式进行分布式计算，这就需要我们拥有两台以上的主机进行测试并相互关联。
4. 对元数据进行处理并以可视化的形式展现给用户，其美观程度对项目的展示显得比较重要，这也是我们后期以美化界面为主要任务的原因。

应对措施

1. 我们通过在淘宝上对多家店铺进行筛选，综合销量和评分等诸多因素予以考虑，并通过与店家进行详细的沟通，选取了多家 wifi 探针进行测试，以保证硬件所获取的数据的相对真实与正确。
2. 由于 wifi 探针获取数据刷新时间短，在短时间内将获得大量数据，我们将对数据进行预处理，使得相邻时间相同数据并不予以存储，从而节约的存储空间，并对数据进行分类和并采用关系数据库进行存储，保证了数据易查易用。对于大量数据的并行处理，我们将采用分布式集群的方式，并将多线程与多进程相结合，充分利用 CPU 进行高效运算。
3. 我们将采用阿里云作为测试主机
4. 我们采用 java swing 进行 GUI 开发，由于 java 编程语言各方面较为健全，其相关库也较为完善，通过第三方 lib 我们将打造美观，易懂的图表来对数据进行可视化显示与分析，以追求更加完美的用户体验。

市场风险

1. 本项目的目标用户——商店业主对本产品的认知理解能力低，接受能力差，对本项目对其的益处存在怀疑，以保守的态度不予接纳。

2. 同类型产品的竞争使得用户忠诚度降低，市场的波动导致用户购买因素以价格主导而并非以技术主导
3. 同类产品的模仿和用户对技术更新速度的要求导致产品开发周期缩短，功能需求点增多。
4. 对小微店铺业主吸引力不大，对各类行业特定需求难以细分并针对实现。相关客户可能对新兴事物的兴趣不大，不愿意学习相关操作。

应对措施

1. 通过多渠道，多角度进行宣传，并进行耐心的讲解以告知相关目标用户客户数据存在的价值以及对店家销量，客流量，营销质量等多方面的益处。并通过试用，优惠等多种方式让其进行试用以得到客户的好感与认可。
2. 针对同类型产品，我们会以更加平民的价格，更加高质量的功能以及更加友好的服务来取胜。
3. 我们会设置通道来接收用户对产品的反馈，并通过统计选取同类型产品需求进行开发，以最大限度的满足用户所期待的功能。同时，我们崇尚技术为主导，因此我们也会不断添加和完善新的功能点来丰富自己的产品，增加竞争力。
4. 我们的可视化图表极大的方便了使用者对数据的分析工作，通过简单的操作即可对近一天，一周，一个月的客流量，入店量，新老用户等客户关注进行查询，使用户轻松方便的得到大数据分析之后的结论以调整营销策略，增加店铺吸引力。

财务风险

本项目涉及的费用支出在于 wifi 探针硬件部分以及对虚拟主机的购买。

应对措施

我们将以质量第一，价格第二的标准对 wifi 探针进行选取，以最大限度的平衡高质量 wifi 探针和减少支出之间的关系。与此同时，对于阿里云主机，我们通过学生认证来购买得到性价比高的云主机。

管理风险

1. 由于本项目涉及的技术点较多，且诸如大数据分析，分布式构建等本项目关键技术对我们来说都是较大的考验。
2. 开发时间的限制意味着我们将在较为忙碌的条件下进行开发测试，同时对于每个队员的高效工作也会产生一些影响。如何保持在规定时间内持续高效的开发是项目成功的关键。
3. 在协作过程中可能由于信息交换不及时，信息不对等等原因，使得团队成员对项目理解有偏差，进而影响整体项目开发进程。

应对措施、

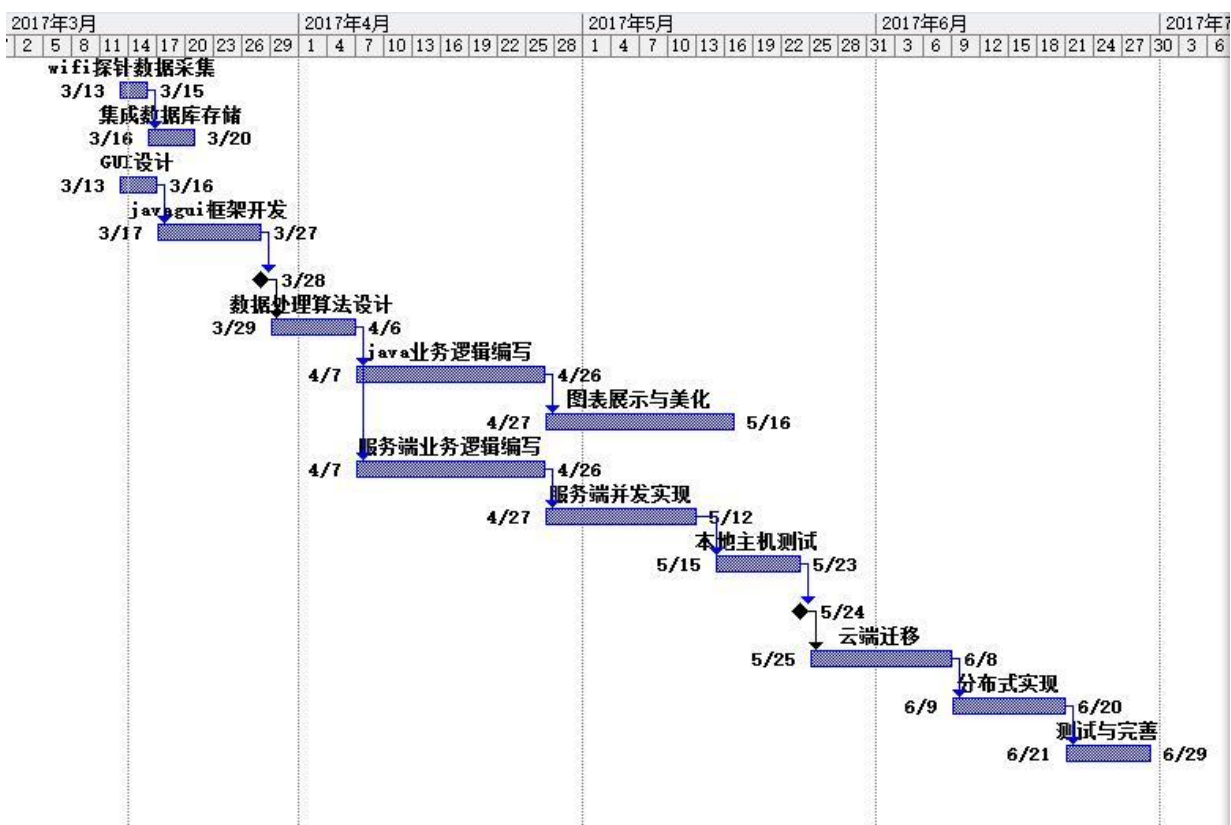
1. 针对于较为复杂的技术，我们将与相关老师进行沟通并通过广泛阅读资料与论文将其攻克。通过科学的分工和进度安排，我们将对每一项技术点进行开发与实现。
2. 通过定期的研讨和共同协作来维持，增进开发热情，保持开发效率，从而保证在规定时间内完成预期的目标

3. 及时，经常的沟通与讨论将会推进开发进度。

4.2 工作流程

我们将采用 XP（极限编程）的方式对整个项目进行开发。由于本团队由三个人组成，团队成员少，相比于传统的螺旋式开发流程和迭代式开发，XP 可以将复杂的开发过程分解为一个个相对比较简单的小周期；通过积极的交流、反馈以及其它一系列的方法，开发人员和客户可以非常清楚开发进度、变化、待解决的问题和潜在的困难等，并根据实际情况及时地调整开发过程。通过简单设计，持续集成，小型发布与验收来实现最终需求。

4.3 总体进度计划



项目开发计划甘特图

4.4 项目控制计划

4.4.1 质量保证计划

执行质量评审活动，对过程质量进行控制。规模较大的项目应当单独编写《软件开发项目质量计划》。根据 GB/T 12504 计算机软件质量保证计划规范，内容包括：

◆ 管理

描述负责软件质量管理的机构、任务及其相关的职责

◆ 文档

列出在该软件的开发、验证与确认以及使用与维护等阶段中需要编制的文档，并描述对文档进行评审与检查的准则

◆ 标准、条例和约定

列出软件开发过程中要用到的标准、条例和约定，并列出监督和保证执行的措施

◆ 评审和检查

规定所要进行的技术和管理两个方面的评审和检查工作，并编制或引用有关的评审和检查规程，以及通过与否的技术准则。至少要进行软件需求评审、概要设计评审、软件验证与确认评审、软件系统功能检查、程序和文档物理检查

◆ 软件配置管理

编制有关配置管理条款，或在“4.4.4 配置管理计划”中说明，或引用按照《GB/T 12505 计算机软件配置管理计划规范》单独制定的文档

◆ 工具、技术和方法

指明用于支持特定软件项目质量管理工作的工具、技术和方法，指出它们的目的和用途

◆ 媒体控制

说明保护计算机程序物理媒体的方法和设施，以免非法存取、意外损坏或自然老化

◆ 对供货单位的控制

供货单位包括项目承办单位、软件销售单位、软件开发单位。规定对这些供货单位进行控制的规程，从而保证项目承办单位从软件销售单位购买的、其他开发单位开发的或从开发单位现存软件库中选用的软件能满足规定的要求。

◆ 记录的收集、维护和保存

指明需要保存的软件质量保证活动的记录，并指出用于汇总、保护和维护这些记录的方法和设施，并指明要保存的期限

4.4.2 进度控制计划

我们将以一周作为审查周期对该周所完成的进度进行审查，并针对该周所遇到的有关问题进行及时的研讨和处理，以保证能够与预期进度相吻合的完成整个项目。

4.4.3 预算监控计划

无。

4.4.4 配置管理计划

无。

5 支持条件

说明为了支持本项目的完成所需要的各种条件和设施。

5.1 内部支持

支持需求	工作阶段	用途
Wifi 探针	项目起始点-项目终止点	提供数据来源，作为数据采集的渠道，为整个项目后续进行打下基础
云主机	项目开发后期-项目终止点	将本机测试后的程序移到远端主机进行测试，进行分布式部署和运行。
编程工具 IDE	项目开发后期-项目终止	编写程序

网络环境	点	
------	---	--

5.2 客户支持

客户需提供场地供本项目进行实地测试和分析，与此同时，客户可以在开发过程中对项目进行检验和提出意见。

5.3 外包（可选）

无

6 预算

6.1 人员成本

无

6.2 设备成本

无

6.3 其它经费预算

无

7 关键问题

关键问题描述	技术难点	风险和影响
顾客手机Mac地址等相关数据采集可行性的问题	Wifi 探针采集顾客手机 Mac 地址等相关信息的前提是顾客手机必须开启 wifi 功能，而很多手机都有节电保护模式，这就增加了采集数据的难度	数据是整个项目的基础，没有相关的数据的话，会使得整个软件系统的数据准确率降低，可用性降低。影响整个软件系统的商业价值。
Wifi 探针能够监听以及采集到范围内的大部分手机的Mac地址等相关信息；并且，采集的频率也快，因此可以产生大量的数据	数据量级大。对如此量级的数据的处理就成了实现整个软件系统的一个关键问题；此外，如何高效调度和使用计算资源以确保对客户的服务品质也是一个问题	数据的处理和分析的正确性直接影响到整个技术、软件系统的准确性、可靠性；数据分析处理的效率则影响了整个软件系统的用户体验

采集到的数据及分析后的结果呈现给客户的形式	采集数据的目的是为了能够处理分析，从而给客户一个参考指标帮助其制定商业计划或决策；而以什么样的形式展示出来能够有较好的效果，需要我们进行深入调查、不断进行测试以及改善	这个问题是本项目中的一个难点，呈现形式将直接影响展示给客户的结果的直观性进而影响本项目的价值
服务端并发问题	在实际的业务场景中，会有很多的顾客，因此同一时间也会有很多数据和并发量，面对如此的吞吐量要求，如何解决是一个很大的挑战	如果不能很好的解决在较大并发用户数下保持较高的吞吐率，便会影响服务端继而影响整个系统的使用

8 专题计划要点

1. 人员培训计划：暂无
2. 测试计划：测试时所有人员都要参与，并随机选择普通用户参与测试，并记录不足之处，以便人员可实行软件修正与优化
3. 质量保证计划：开发过程中互相监督与检查，尽量避免错误；开发小组可以不定期对用户进行软件系统的维护，必要时可对用户进行必要的软件使用培训。
4. 分合同计划：暂无
5. 用户培训计划：基本上不需要就可以使用我们的产品。