

1 引言	2
1.1 编写目的	2
1.2 背景	2
1.3 定义	2
1.4 参考资料	3
2 任务概述	3
2.1 目标	3
2.2 用户的特点	3
2.3 假定和约束	4
3 需求规定	4
3.1 对功能的规定	4
3.2 对性能的规定	8
3.2.1 精度	8
3.2.2 时间特性要求	9
3.2.3 灵活性	9
3.3 输入输出要求	9
3.4 数据管理能力要求	9
3.5 故障处理要求	9
3.6 其他专门要求	9
4 运行环境规定	10
4.1 设备	10
4.2 支持软件	10
4.3 接口	10
4.4 控制	10

软件需求规格说明书

1 引言

1.1 编写目的

本需求说明书目的在于：将用户提供的需求描述系统化、精确化、全面化。从而实现：

1. 便于用户、分析人员和设计人员进行理解和交流。
2. 支持目标软件系统的确认。
3. 控制系统进化过程。

预期读者：软件设计者和测试者。

1.2 项目背景

随着科学技术的高速发展，我们已步入数字化、网络化的时代。网上购物愈益流行，然而，受到产品质量检验和实际体验感的限制，线下商店当然是不可替代的，很多顾客仍然希望进入实体店亲身试用挑选。

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。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

还可以通过本系统随时查看店铺内的客流量情况，并根据客流高峰时段，对店内工作人员进行合理分配，提高人力资源利用率，并在一定程度上降低经营成本。

1.3 定义

WIFI 探针：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。

大数据：指无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合，是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。

1.4 参考资料

- [1] 《Spark 快速大数据分析（第一版本）》
- [2] 《Hadoop 基础教程》
- [3] 《精通 Hadoop(第一版本)》
- [4] 《响应式 Web 设计：HTML5 与 CSS 实战》
- [5] 《JavaScript 权威指南（第 6 版）》
- [6] 《疯狂 Ajax 讲义（第 3 版）》
- [7] 维克托·迈尔·舍恩伯格，肯尼思·库克耶，周涛. 《大数据时代》[J]. 教育科学论坛, 2013, 8(7):27-31.
- [8] 殷人昆, 郑人杰, 马素霞, 白晓颖. 实用软件工程(第三版)[J]. 计算机教育, 2010(24):95.
- [9] 《猫酷室内行为采集系统》 <http://www.mallcoo.cn/action.html>

2 任务概述

2.1 目标

本系统旨在通过 WIFI 探针收集顾客 MAC 及与探针的距离、出现时间地点等信息，来分析门店的客流情况、精准监控客流质量、实时展示客流转化的情况，从而帮助检测营销效果、发现潜在机会和改进措施，为便捷、高效精细化运营提供全方位的数据参考。

基本的技术目标包括：

- 1) 用 WIFI 探针收集顾客信息，实现实时的客流量监测并实现环比与历史比较；
- 2) 根据历史客流量，预测未来时刻的店内客流量，以便商家进行人员调度；
- 3) 获取顾客实时入店量并予以实时展示并实现环比与历史比较，从而了解进入店铺或区域的客流及趋势；
- 4) 分析比较得出实时入店率并予以实时展示并实现环比与历史比较，从而获取进入店铺或区域的客流占全部客流的比例及趋势；
- 5) 快速分析得出顾客来访周期从而实现对进入店铺或区域的顾客按照距离上次来店不同间隔实现动态归类；
- 6) 顾客活跃度：按顾客距离上次来访间隔，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）；
- 7) 快速分析进入店铺的顾客在店内的停留时长并实现动态归类予以实时展示；
- 8) 根据驻店时长来判断进入店铺后很快离店的顾客及占比（占总体客流）即

- 跳出率，并实现实时展示与小时、日、周、月多维度环比以及历史比较；
- 9) 根据顾客停留时长判定计算进入店铺深度访问的顾客及占比(占总体客流)即顾客深访率并予以实时展示与小时、日、周、月多维度环比以及历史比较
 - 10) 统计商家推送的营销方案使用率，辅助商家进行调整。
 - 11) 实现使用短信控制模块控制 WIFI 探针的开关，与通过数据获取情况简单甄别判断探针的状态予以实时监控。

2.2 运行环境

2.2.1 硬件环境：

CPU：IntelCoreI51.8GHz 以上

内存：2G 以上

探针：双核探测距离半径>100 米频率 2.4GHz-2.5GHz

2.2.2 软件环境：

服务器：Tomcat7.0

操作系统：windows7 及以上

数据库：NOSQL

PC 端：IE6.0 及以上版本；IE 内核的其它浏览器；Chrome21.0 等

手机端：自带浏览器即可

2.3 假定和约束

1. 此系统有且仅有分店、管理员两类用户使用；
2. 预测客流量数据的计算模型，是根据历史客流量数据得来的，不能保证完全正确，仅供参考；
3. WIFI 探针通过客户手持设备的信号强度估测其所在区域，并不能实现精准定位且存在顾客手持多台设备或者设备未开启 WIFI 的情况从而数据获取存在一定误差；
4. 目前的版本支持 IE6.0 及以上版本的浏览器，Chrome21 等，对于较低版本的浏览器可能会出现页面错乱等现象；
5. 条件限制不能提供 1000 台设备同时运行的测试环境从而可能发生未预知的错

误。

3.数据描述

3.1 静态数据

名称	描述	用途	备注
历史数据	用做模型建模的数据输入	建立入店量、活跃度、驻店时长等数据预测模型和用作环比与历史对比	\

3.2 动态数据

名称	类型	描述	备注
原始数据	INPUT	采集数据	\
结构化数据	INPUT	对采集数据进行结构化后储存转至界面显示	\
模型数据	INPUT	模型数据文件	\
采集的数据	INPUT	用于模型学习趋势预测 例如顾客活跃度、客流量等	\
采集的数据	OUTPUT	原始数据文件	\
客流量的分析、预测结果	OUTPUT	数据可视化	\
入店量的分析、预测结果	OUTPUT	数据可视化	\
入店率的分析、预测结果	OUTPUT	数据可视化	\
新老顾客比例的分析、预测	OUTPUT	数据可视化	
顾客活跃度的分析、预测结果	OUTPUT	数据可视化	\
深访率的分析、预测结果	OUTPUT	数据可视化	\
跳出率的分析、预测结果	OUTPUT	数据可视化	\

3.3 数据库介绍

类型	名称	备注
DB	实时数据	存储所有探针采集到的实时数据
DB	客流量及历史	存储当天每个时段的客流量数据
DB	当前店内人数	存储每个时段店内人数，
DB	入店量入店率及历史	存储当天以及过去日、周、月的的入店量入店率数据

DB	顾客驻留时间及历史	存储根据驻店时长长短分类的各类人数划分阈值为 [30s, 30s-1min, 1min-5min, 5min 以上]
DB	新老顾客人数比例及历史	存储根据顾客上次来访时间判别的新老顾客（间隔一个月为阈值）的人数及比例
DB	顾客访问次数及时间历史	存储顾客在规定时间内访问次数的访问次数
DB	深访率及历史	存储根据驻店时长判定的深访顾客人数
DB	跳出率及历史	存储根据驻店时长划分，标准为低于30s 算作跳出顾客来判定的人数
DB	活跃度及历史	存储根据顾客访问次数得出的活跃度及历史数据

4. 需求规定

4.1 对功能的规定

基于 WIFI 探针的商业大数据分析系统有以下几项功能需求：

- **WIFI 探针设备：**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、时间信息（特别针对 ANDROID6.0 和 IOS10 版本后的移动终端设备进行测试需能采集到 MAC 地址）。
- **数据传送：**服务端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。接收数据不能有数据丢失或者数据失真，探针每 3 秒发送一次数据，数据采集并发量即单台服务器的数据吞吐量，不得低于 1000 台设备；
- **数据存储：**实时数据量获取后及需及时结合历史数据分析，需将已采集生成的数据快速并定期处理后，存储并生成相应的文件，方便以查阅及处理。
- **数据处理：**利用大数据技术 3s 之内完成各门店客流量统计，顾客来访次数和来访时长，与顾客活跃度。
- **数据可视化：**探针采集的原始数据要经过比较处理后以图表或者其他直观的方式实时展现出。

- 活跃度趋势预测：根据当前采集到的数据，分析并预测顾客的来访活跃度。
- 营销方案热度判断：系统会定期向管理层提交营销方案的热度判断报告。
- 数据更新：系统定期更新历史数据及其模型。

对功能的需求可用如下的数据流图描述：

4.2 功能描述

4.2.1 数据流的定义

数据流名称：原始数据
描述：探针采集到的数据
组成：mac 地址+时间戳+信号强度+手机距离嗅探器的距离，经纬度，地址
来源：探针 终点：采集数据

数据流名称：结构化数据
描述：系统分类统计的数据
组成：结构化 mac 地址+结构化时间戳+结构化信号强度
来源：采集数据 终点：linux 计算平台以及数据库

数据字典：

4.2 对性能的规定

4.2.1 精度

数值要求精确到整型个位即可，要求学号完全准确，姓名可以进行部分匹配。

4.2.2 时间特性要求

- a. 响应时间：小于 0.5 秒；

- b. 更新处理时间：小于 1.5 秒；
- c. 数据的转换和传送时间：小于 2 秒；

4.2.3 灵活性

- a. 操作方式上的变化：该系统适用于任一 Windows xp 以上 操作系统，同时也能在手机端通过网页访问。
- b. 运行环境的变化：要求数据计算平台使用 linux 平台，要求离线计算或者实时计算平台必须是分布式环境，此外该系统可以适用于 Windows xp 以上的所有操作系统。
- c. 同其他软件的接口的变化：可以满足 B/S、C/S 两种类型。操作尽量简单，好用、易用。
- d. 系统升级历史数据的变化：升级后的系统会自动保留用户数据。
- e. 精度和有效时限的变化：可以根据实际情况自行设置。

4.3 数据管理能力要求

- 1. 需要对实时数据进行处理后分析结果存储。
- 2. 需要将阶段性的数据整合处理后以指标的形式存储例如日、月客流总量总量等。
- 3. 采用 NOSQL 的数据库,需要对键值关系有统一的规范，便于查询。

4.4 故障处理要求

发生错误时，保证数据完整，对于数据库发生故障时要能够进行故障恢复，以保证数据的一致性同时也要定期进行数据备份。

名称	处理方式	备注
硬件传输数据故障	由硬件负责人进行故障排查和修复	\
模型欠学习或过拟合	重新训练模型	\
推送或者接入 WIFI 故障	重启	\
服务器故障	重启	\

数据库无响应	重启	\
断电	挂起任务	\
网络故障	硬件负责人进行故障排查	\

4.6 其他专门要求

1. 可维护性：维护人员会在定期进行维护和检验，利用可靠的密码技术，掌握特定的记录 或历史数据集，便于维护。
2. 安全保密性：所有用户的信息经过加密后存储在数据库，如需查询有关信息，必须通过 严格的身份验证，以防数据泄露。
3. 可使用性：该系统界面友好，功能详细简单，方便使用。

5 运行环境规定

5.1 设备

列出运行该软件所需要的硬设备。说明其中的新型设备及其专门功能，包括：

- a. 处理器型号及内存容量：P1 及以上，内存大于 32M；
- b. 外存容量、联机或脱机：10G，联机
- c. 输入及输出设备的型号和数量，联机或脱机：联机或脱机均可；

- d. 数据通信设备的型号和数量：服务器 Sever;
- e. 功能键及其他专用硬件：快捷键

5.2 支持软件

支持 Windiw9x/NT/Me/2000/xp， VFP \Delphi\ c\ c++编译程序等。

5.3 接口

5.3.1 硬件接口

名称	描述	备注
探针	通过 WIFI 模块收集数据并进行网络数据传输	\

5.3.2 软件接口

用户接口为用户界面，通过系统提示用户可输入数据以得到自己所需要的服务；外部接口为鼠标、键盘、触摸屏等交互工具，另需配置一台中心服务器，本系统与其他系统通过调用相同的数据库中的数据而保持一致性；内部接口为学生的学号、课程号（名）的一致性。

5.4 控制

本系统通过网络提供服务，用户通过浏览器访问服务器，向服务器发出服务请求。因此，需要使用 TCP/IP 网络协议，作为标准的通信控制接口。

