

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

计算机通信网络

COMPUTER COMMUNICATION NETWORK

HAC 聊天程序项目说明报告



姓 名：_____ 杨超琪

学 号：_____ 515030910067

年 级：_____ 大三

所在院系：_____ 电子信息与电气工程学院

任课教师：_____ 蒋兴浩

2017 年 12 月 12 日

目录

一. 项目说明	2
二. 运行环境	2
1. 物理主机环境配置	2
2. 开发平台物理配置	2
3. 数据库配置	3
4. 程序文件列表	4
三. 主要功能的实现	5
1. 多用户的一对一通信	5
2. 多用户的一对一文件发送	5
3. 群文本通信与群文件通信	6
4. 表情的发送	6
5. 实时更新在线	7
6. 客户端登录与 HAC 动态圈登录	7
7. HAC 动态圈的功能	8
四. 一些技术细节的说明	10
1. 文本发送与连接状态检测原理	10
2. 表情的发送原理	10
3. 文件的发送原理	11
4. HAC 动态圈操作原理	11
五. 遇到的问题及解决方法	11
六. 存在的缺点与对未来的展望	14
1. 存在的缺点	14
2. 对未来的展望	14
七. 体会与建议	14
1. 体会	15
2. 建议	15
参考文献与链接	15

HAC 聊天程序项目说明报告

——计算机通信网络大作业

一、项目说明

此项目名称为：“Have a Chat!”（简称 HAC），是一个集聊天、文件传输、表情发送、多用户一对一聊天、群组聊天、动态分享、聊天记录查询等功能于一体的聊天软件，由 Qt4 前端编程框架、Python 语言及 XAMPP 数据库技术支持。

二、运行环境

1. 物理主机环境配置

操作系统	Windows 10 家庭版
处理器	Intel(R)-i7-CPU @2.60GHz
内存	8.00 GB
系统类型	64 位操作系统

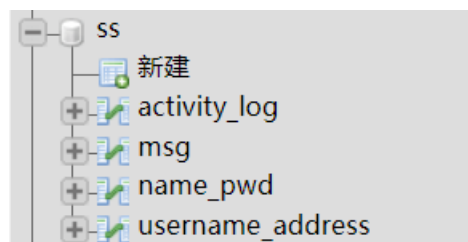
2. 开发平台物理配置

开发语言	Python 2.7.13
编译工具	Sublime Text & 原生编译器
服务器工具	XAMPP

注：依赖于 Python 2.7.13 下的库：os、PyQt4、time、socket、struct、math、webbrowser、PIL、Image、MySQLdb、threading、socketserver、flask、hashlib、random、urllib2、re、bs4、BeautifulSoup；依赖于 XAMPP 下的 phpMyAdmin 数据库。

3. 数据库配置

- ✓ 数据库名称: ss
- ✓ 数据表一: activity_log
- ✓ 数据表二: msg
- ✓ 数据表三: name_pwd
- ✓ 数据表四: username_address



1) activity_log

用于储存用户发送动态的信息，一共七个字段：**name**（字符串类型，用户的注册名）、**nickname**（字符串类型，用户在HAC圈的自设昵称）、**log**（字符串类型，用户动态的内容）、**time**（字符串类型，友好格式的发动态时间）、**location**（字符串类型，发动态的地点）、**mood**（字符串类型，当时的心情）、**time2**（整数类型，用于排列发动态的时间顺序）

name	nickname	log	time	location	mood	time2
ycq2	Mr.ycq	旧瓶装新酒，翻拍的电影都难免和旧作比较。我对旧电影有印象但不深刻，只觉得旧版的波罗更聪明，新版其他还...	Sat Sep 16 12:32:48 2017	山东省:临沂	just so so	1505536368
zhouyiran	Mr.Box	我们年轻时候，能够吃特别甜的点心，比如蜜三刀；能够吃特别腻的肥肉，比如烧白。能够看特别俗辣的电影，比...	Mon May 01 17:39:39 2017	湖南省:株洲	good	1493631579
ycq_text2	Mr.Box	IMdB公布的电影时长：120分钟 院线版本电影时长：100分钟 首先，为什么说两个版本，因为不只...	Thu Jan 12 09:19:04 2017	湖南省:株洲	good	1484183944
ycq2	Mr.ycq	米格的曾曾爷爷为了追求音乐抛弃家人，拆散了整个家庭。而曾曾奶奶通过做鞋修补了这个家庭，并且将做鞋的本...	Sat Nov 12 09:27:41 2016	江西省:景德镇	good	1478914061
ycq2	Mr.ycq	故事，就从某一种盘据全亚洲的女孩类型说起。这种女孩，叫“便利贴女孩”。她，朴素、简单安分、愿望也...	Fri Jul 08 05:41:53 2016	江西省:九江	just so so	1467927713

图 1：随机生成的用户动态

2) msg

用于储存用户之间的聊天记录，一共五个字段：**receiver**（字符串类型，信息接收者）、**sender**（字符串类型，信息发送者）、**message**（字符串类型，信息内容）、**time**（字符串类型，友好格式的发动态时间）、**time2**（整数类型，用于排列发动态的时间顺序）

receiver	sender	message	time	time2
ycq1	yangchaoqi	zE5SIEO9SAWNYDCxU22T	Sun Sep 17 00:47:13 2017	1505580433
zhouyiran	ycq_text2	Mg0zs8SCxrv11rhAhQJ9	Sat Sep 16 13:14:16 2017	1505538856
ycq_text2	yangchaoqi2	PWAaeqzspeWKovdXJ3lr	Wed Sep 06 01:22:41 2017	1504632161
yangchaoqi2	ycq_text1	HpnyJlqupRhJbBuYtrlw	Wed Aug 09 06:17:27 2017	1502230647

图 2：随机生成的聊天信息

3) name_pwd

用于存储用户的个人信息，一共六个字段：**name**（字符串类型，用户的注册名）、**pwd**（字符串类型，用户密码）、**img**

(字符串类型, 用户头像在服务器上的相对路径)、**motto** (字符串类型, 用户的个性签名)、**profession** (字符串类型, 用户的身份)、**nickname** (字符串类型, 用户的昵称)

name	pwd	img	motto	profession	nickname
group	123	static/uploads/group.png	Let's Chat !	Undergraduate in SJTU	Mr.Box
yangchaoqi	15821188556	static/uploads/icon.png	fly higer	Undergraduate in SJTU	Mr.Box
ycq1	aaa	static/uploads/09e8e09be017032	fly	student	ycq
ycq10	asd	static/uploads/new_arrow_d.png	Love a flower bloom to accompany it, love a man t...	Undergraduate in SJTU	Mr.Box

图 3: 已注册的用户信息展示


4) username_address

用于存储当前的用户的连接信息 (服务端对其实时更新), 一共五个字段: **name** (字符串类型, 用户的注册名)、**address** (字符串类型, 用户的 IP 地址)、**pid** (字符串类型, 唯一进程标识)、**status** (整数类型, 1/0 表示 socket 连通/不连通)、**time** (字符串类型, 表示该用户与服务器连通的时间节点)

name	address	pid	status	time
yangchaoqi1	127.0.0.1	62508	0	1511683657.08
ycq3	127.0.0.1	57994	1	1511772143.91
ycq4	127.0.0.1	58022	1	1511772168.08
ycq5	127.0.0.1	58283	0	1511772600.19

图 4: 部分用户与服务器连通信息

4. 程序文件列表

 flowLayout.py	 generateActivityLog.py
 generateMsg.py	 getDouban.py
 getTool.py	 hacDemo1.py
 hacDemo2.py	 msgList.py
 Panellist.py	 server_file.py
 server_moment.py	 server_msg.py

注: hacDemo1.py、hacDemo2.py 为开启两个客户端 Demo 文件; server_file, server_moment.py、server_msg.py 为开启 Demo 前需要双击运行的服务器端文件; flowLayout.py、msgList.py、Panellist.py 为程序的界面配置文件; getTool 为客户端与服务器端与数据库交互的函数文件、generateActivityLog.py、generateMsg.py 为随机生成默认信息的文件、getDouban.py 为爬取豆瓣 Top 影评榜的文件。

三、主要功能的实现

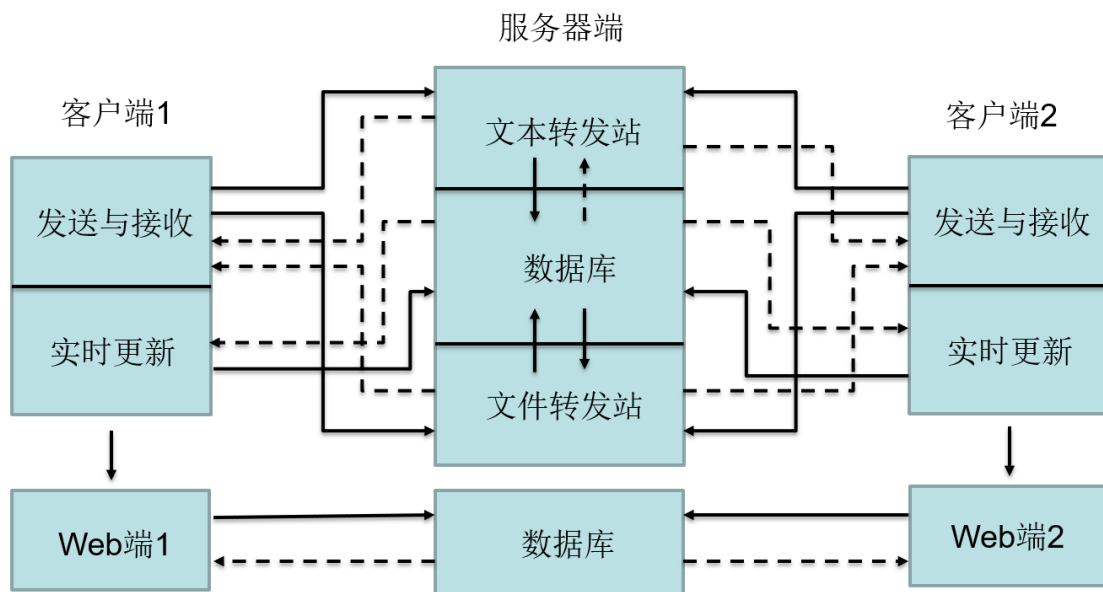


图 5：系统顶层视图

1. 多用户的一对一文本通信

- 1) 从发送端来看：1. 点击列表中的在线用户，获取接受者的 id；2. 点击发送，将文本，接受者 id，自身 id 打包成固定的格式发送给服务器
- 2) 从服务器端来看：1. 接受用户文本消息（不必知道发送的用户）；2. 按照格式解析消息得到发送者 id，接受者 id 与文本消息；3. 根据接受者 id，将消息按照固定格式打包发送给接受者；4. 同时将这条消息与时间存入数据库。
- 3) 从接受者来看：1. 接收到服务器发送的消息；2. 按照固定格式解析消息；3. 如果当前用户为发送者，则将消息显示在对话框。

2. 多用户的一对一文件发送

- 1) 从发送端来看：1. 点击列表中的在线用户，获取接受者的 id；2. 选择文件，获取文件名称与路径；3. 文件名称以文本的形式发送给服务器；4. 文件通过另一个 socket 线程发送给客户端
- 2) 从服务器端来看：1. 接受用户文本信息，以文本发送的形式发送给相应的客户端；2. 接受用户的文件信息，并发送给所有的客户端；
- 3) 从客户端来看：1. 相应的客户端接受到服务器发送的文本

信息，所有客户端都收到服务器发送的文件；2. 没有收到文本的用户调用系统函数删除该文件，收到文本的用户将文本信息显示在对话框，表示已收到。

3. 群文本通信与群文件通信

- 1) 从客户端来看：与上述两个功能实现几乎一样，只需将接受者 id 转换成群组的 id 即可
- 2) 从服务器端来看：与上述两个功能实现几乎一样，只需将接受者 id 设为所有用户

4. 表情的发送

- 1) 从发送端来看：1. 点击列表中的在线用户，获取接受者的 id；2. 选择表情，获取表情的编码（如：0x1f1ec1f1e7）；3. 表情编码以文本的形式发送给服务器；
- 2) 从服务器端来看：1. 接受用户文本消息（不必知道发送的用户）；2. 按照格式解析消息得到发送者 id，接受者 id 与表情编码消息；3. 根据接受者 id，将消息按照固定格式打包发送给接受者；4. 同时将这条消息与时间存入数据库。

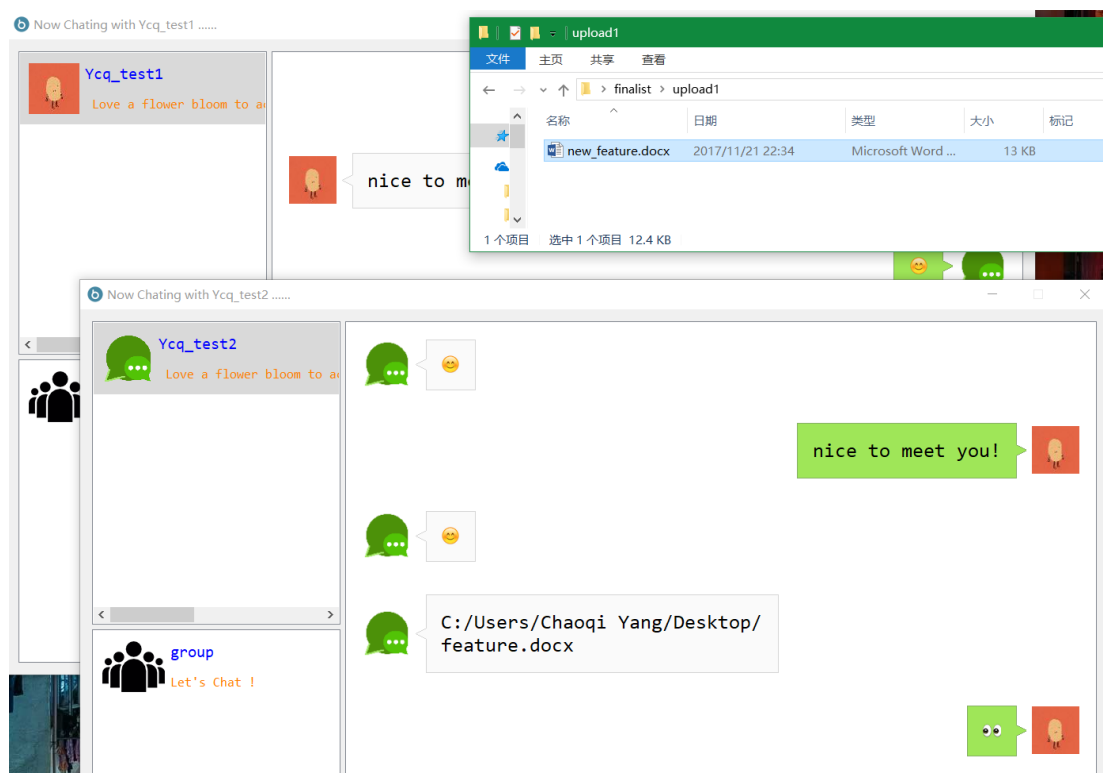


图 6：表情发送、文件传输等功能截图

- 3) 从接受者来看：1. 接收到服务器发送的表情编码；2. 按照固定格式解析表情编码；3. 如果当前用户为发送者，则根据表情编码从表情库中查询得到表情显示在对话框。

5. 实时更新在线联系人

- 1) 从服务器端来看：服务器各线程每隔 1 秒向所有客户端发送消息“#”，并设置 timeout，若未超时，则表示该用户与服务器保持连通；若超时，则表示该用户与服务器断开连接，并把状态信息存入数据库。
- 2) 从客户端来看：每个客户端每隔 1 秒从数据库读取所有用户的状态信息，并查询其他信息（例如注册名、个性签名等），并把在线的用户实时显示在用户列表上。

6. 客户端登录与 HAC 动态圈登录

在客户端需要注册用户名、密码、个性签名、并选择头像。在点击登录时，如果用户名已存在，则会提示重新输入用户名，此时需要关闭程序，重新运行。如果登陆成功则会跳转到聊天界面。

在登录 HAC 动态圈时，同样需要输入用户名与密码，此时服务器会根据数据库的信息，检测用户名与密码是否匹配，登录界面下会显示提示信息。

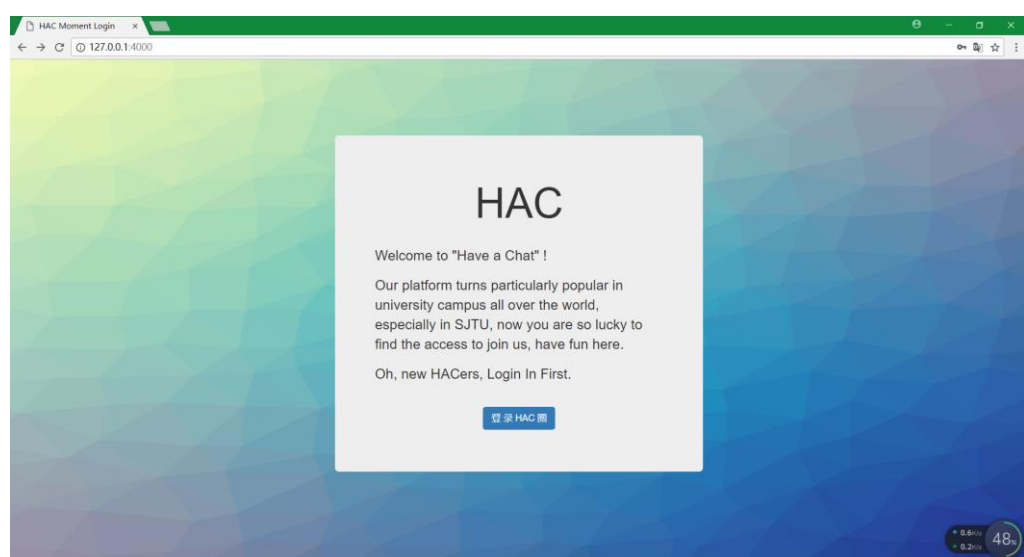
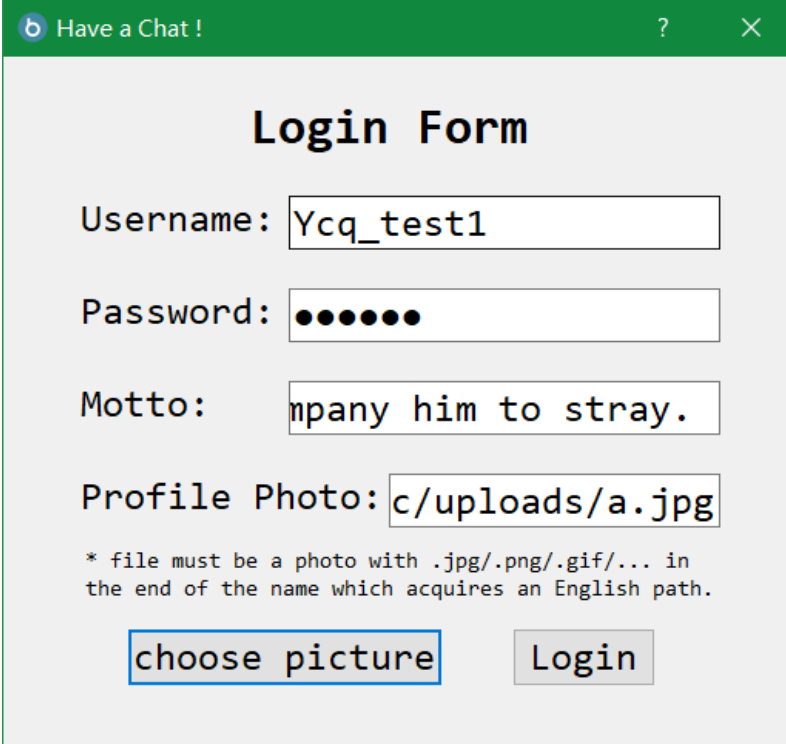


图 7：Web 端动态圈登录界面



Have a Chat !

Login Form

Username: Ycq_test1

Password: ●●●●●●

Motto: npany him to stray.

Profile Photo: c/uploads/a.jpg

* file must be a photo with .jpg/.png/.gif/... in the end of the name which acquires an English path.

choose picture Login

图 8：客户端聊天登录界面

7. HAC 动态圈的功能

HAC 动态圈采用 Python Web 端的方式，所有的交互均与服务器的数据库相关，分为个人信息的修改、个人动态的发送与好友动态的查询、聊天记录的查询。

- 1) 个人信息的修改：分别是昵称、身份信息、头像、个性签名。如果用户的更改信息没有达到要求，网页会提示错误的信息；如果要求均满足，则此时客户端将会在底层请求服务器数据库的信息，及时进行更新并显示。
- 2) 个人动态的发送与好友动态的查询：在这里模仿微信朋友圈的模式，用户可以添加一条新的动态，加入心情与地点。也会在登录的时候立刻看到当前所有用户发送的动态，按照时间顺序逆序排列，因为在登录的时候，底层程序将访问数据库，将信息按照固定的格式显示在页面上。
- 3) 用户与群组聊天记录的查询：在这里也是模仿微信保存历史聊天记录的功能，在页面上可以点击用户头像，将弹出对话框，按照时间顺序自己与此用户或者群组所有的聊天记录。

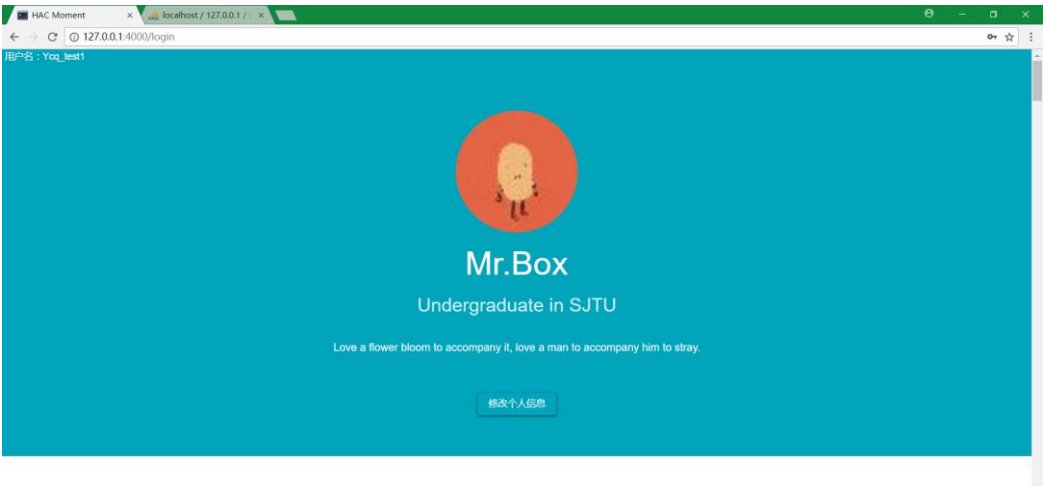


图 9：用户 HAC 动态圈个人信息板



图 10：社区用户动态展示（动态为爬取的豆瓣影评）

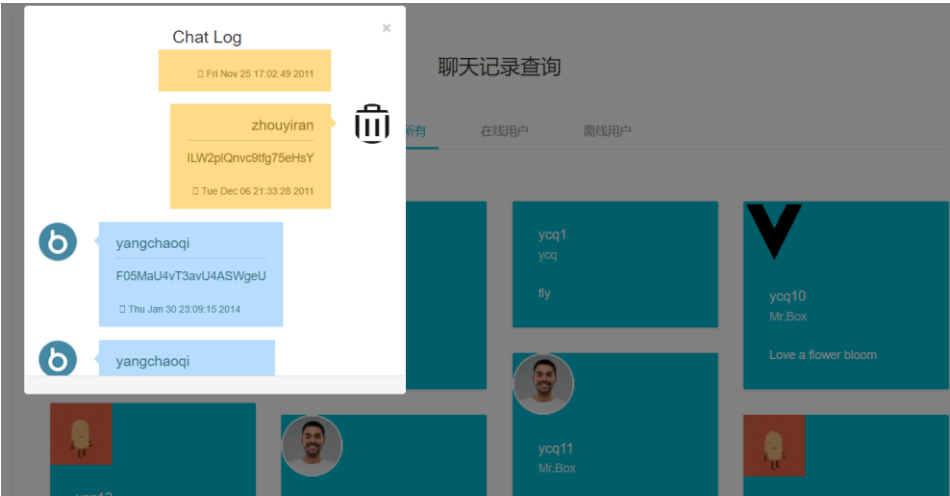


图 11：聊天记录查询（此处聊天记录由代码随机生成）

四、一些技术细节的说明：

1. 文本发送与连接状态检测原理

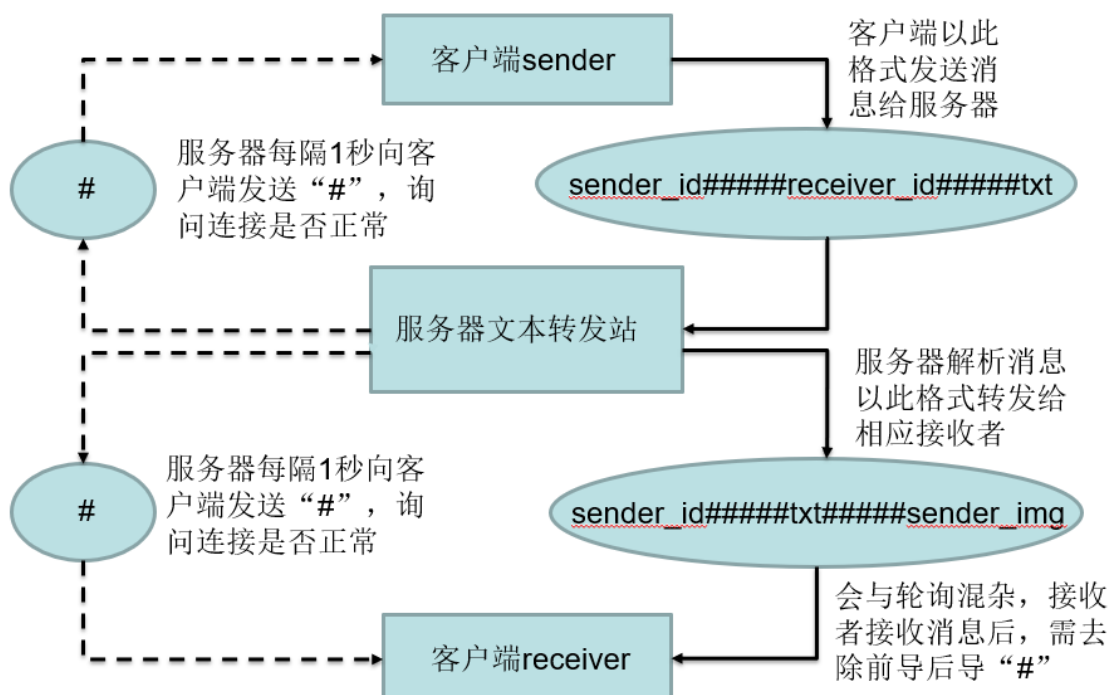


图 12：文本发送与连接状态检测原理图

2. 表情的发送原理

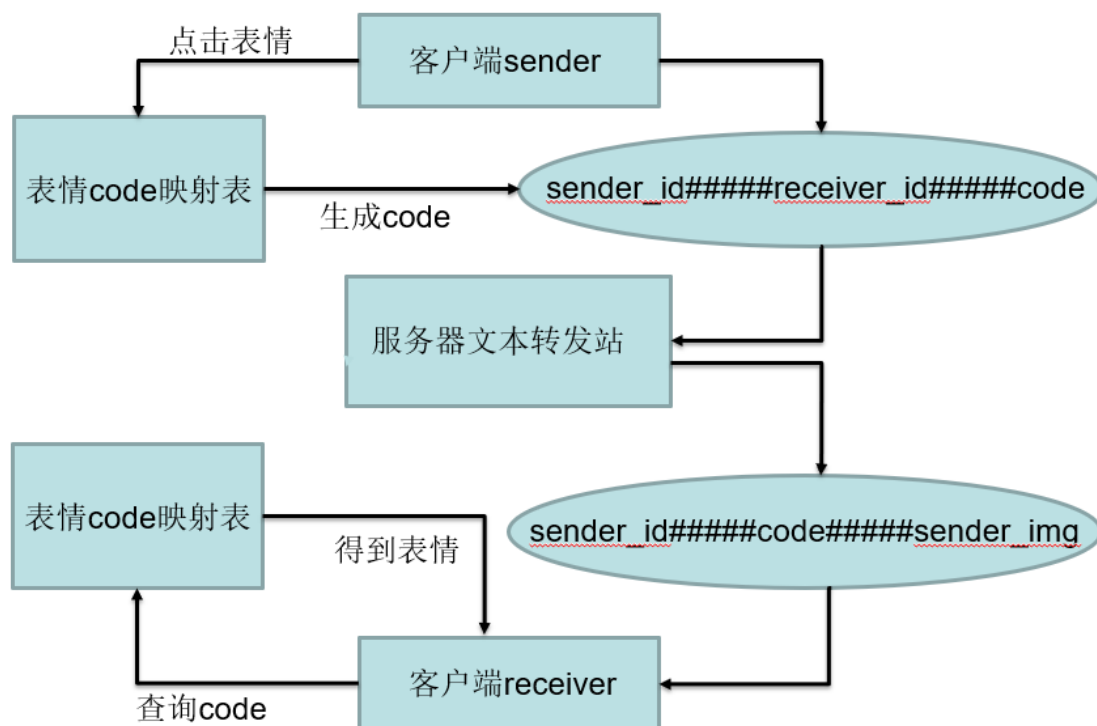


图 13：发送表情原理图

3. 文件的发送原理

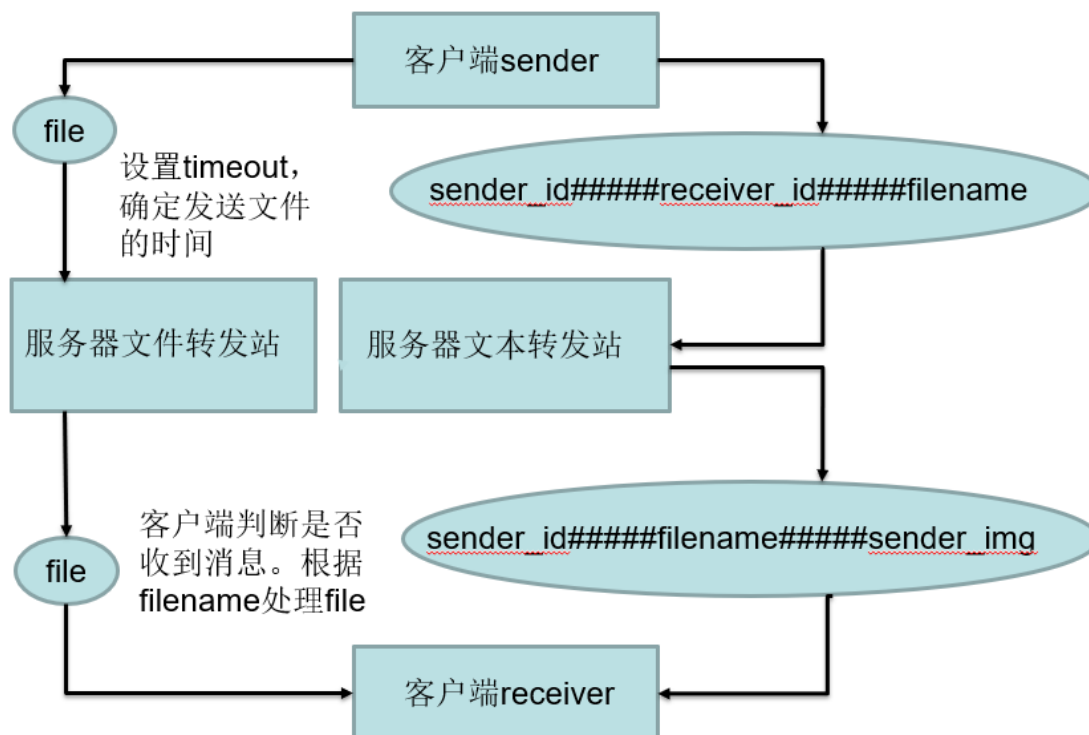


图 14: 文件发送原理图

4. HAC 动态圈的操作原理

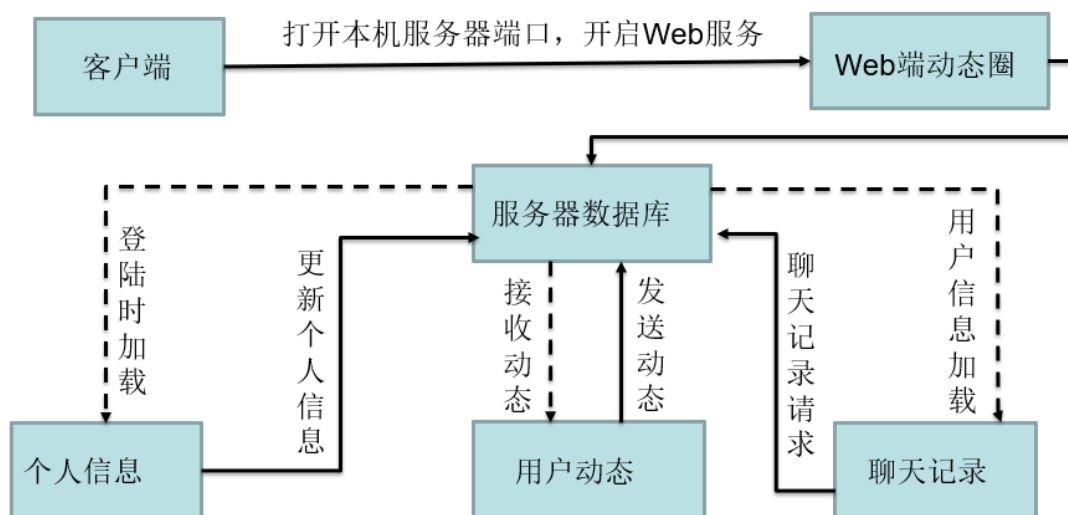


图 15: HAC 动态圈与服务器交互

五、遇到的问题及解决方法

1. 问题：服务器与多个用户的连接问题。

描述：服务器可与一个用户正常连接，多用户时，服务器只连接最后一个用户。考虑到单线程只能一对一通信的问题，利用

多线程,服务器成功连接多用户。调用传统的多线程库 thread, 又出现进程调度问题, 无法判断用户与线程之间的对应关系。
解决: 参考 blog, 调用 socketserver 库, 调用多线程服务器类, 设置列表, 同时与数据库交互, 实时存储用户与线程的对应关系。

2. **问题:** 界面模态与非模态的问题。

描述: 在一个界面中通过调用类产生另一个界面, 出现两种情况: 一种情况是锁定当前应用所有功能, 只能操作新界面, 直到新界面调用 close 函数时, 原程序继续执行; 另一种情况是顺序执行程序, 新出现的界面与原应用操作并行。当在使用类函数时, 一部分构件默认是前一种形式, 另一部分默认是后一种形式, 严重阻碍程序的逻辑执行。查阅官方文档与 blog, 得知前一种模式叫做模态, 后一种叫做非模态, 可以根据需求设置模态与非模态。问题是在设置完模态与非模态模式以后, 没有产生效果。

解决: 不断地尝试与查阅资料, 找到模态与非模态的构件分类, 并根据需求, 将部分界面换成模态类构件, 另一部分为非模态类构件。

3. **问题:** Python 的缩进问题。

描述: 在完成这个项目过程中, Python 的缩进问题成为一个非常大的阻碍。直接的复制粘贴非常容易产生错误信息, 并且 Python 原生的 IDLE 编辑器无法检测产生错误的位置。当使用 Sublime 时, 出现了此类 (如图 16) 现象:

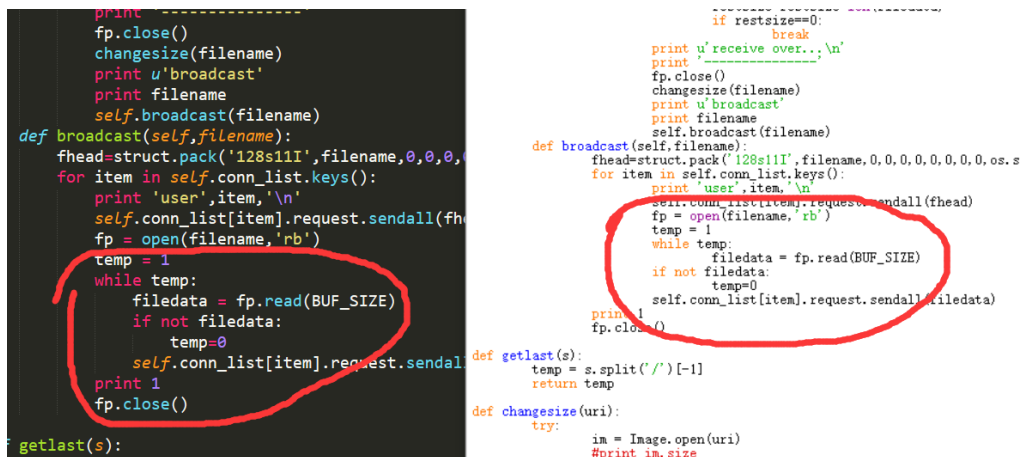


图 16: 同一程序 Sublime 显示与 Python IDLE 显示

并不产生语法错误, 但缩进层次在视觉上无误, 运行时有误, 非常难发现问题。导致进度被拖慢, debug 的难度更进一步。

解决: 注释其他部分, 定点运行程序, 尽量减少不同编辑器间

的复制粘贴，尽量手写。同时在改程序的时候，留意多备份！

4. **问题：**如何正确处理文本、文件与表情的发送。

描述：在开始构思时，想法非常简单，不论用户发送的是文本还是文件亦或是表情，都可以按照文件来处理，用户接受到文件及时显示在对话框。由于考虑到服务器需要自动处理转发问题，在实现的时候，无法将文件与接受者信息同时打包给服务器，因此想到文本与文件分端口发送的方法，同时将表情映射成文本编码。难点在于：由于服务器的自动性，连带客户端的接受也具有自动性，我们必须在两端都开启死循环，但是在客户端死循环时，必须使用多线程，而多线程又涉及到线程通信的问题，因此我在这里巧妙地利用 Qt4 内置的 QTimer 定时器类，与 socket 的 timeout 类似，对服务器端循环发送信号。

解决：文本与文件分端口发送，表情映射成文本编码。

5. **问题：**用户列表实时更新的问题。

描述：首先考虑用户列表实时跟新的问题，这个问题困扰了我很久，如何获取所有用户的状态呢，只能依赖于服务器的数据库了，又由谁来实时向数据库更新用户的连接状态呢，只能依赖全连通的服务器了，因此我设计客户端定时访问数据库得到所有用户的状态信息。接着，如果用户列表需要实时更新，涉及到点击用户的信息，弹出与该用户的对话框，如果实时更新，需要设计一种用户列表与用户 id 匹配的映射关系，如果更新速度太快，界面与程序内部存在延迟，就会出现点击用户，但程序底层获得的是另一个用户的 id 的情况。另外，如果更新速度过快，非常吃电脑的内存与磁盘。

解决：设定每隔 1 秒更新一次用户列表，服务器连续向用户发送请求，如果发送失败，则设置状态为未连接。

6. **问题：**如何正确使用 PyQt4 库。

描述：这个问题是阻碍大作业进程最大的“拦路虎”，因为 Qt4 作为一个著名的 GUI 库，主要是用于 C++ 的编程，因此网上对 Python Qt4 的介绍也比较稀少，官方的 Tutorial 写得非常简略只有寥寥几种构件的介绍，完全不足够用于此项目的书写，如 QTimer 定时器的应用，Qthread 的使用，exec_ 与 close 的区别等等。由于在这个项目中要用到大量的比如按钮触发，信号类间传递，点击构件获得文本等等具体的函数，而 Tutorial 帮助甚微，最后只能运行各种举例，寻找其中的规律。在这个过程中，其实我花了大量的时间在界面的完善上，几乎找遍了网上关于 PyQt4 的介绍，才得以顺利完成。

解决：查阅官方库的文档，参考 blog。

六、存在的缺点与对未来的展望

1. 存在的缺点：

- 1) 在我的项目中，虽然文件传输成功了，但传输文件时并没有显示文件的缩略图，因为不知道该如何使用缩略图，因此只能显示文件的路径名称。
- 2) 由于管理客户端状态时，服务器发送的关键词是“#”，而用于文本传输的分割符为“#####”，因此用户之间无法发送单个“#”，或者超过五个“#####”，否则将会被服务器当成通信关键词或分隔符而丢弃。
- 3) 由于传输表情使用的是表情与编码映射，因此不能发送如：0x1f1ec1f1e7 等的图片编码关键词，否则服务器将会当成表情发送来处理

2. 对未来的展望：

- 1) 因为网上对于用 Python 做语音或视频聊天的内容非常少，而且暂时也考虑不出来如何将语音信息转化为可以利用 socket 传输的数据信息，因此没有添加语音功能。未来希望能成功地利用 Python 的语音模块与 Opencv 模块实现语音聊天与视频聊天功能，或者与 C++或其他语言交叉编程，实现这一功能。
- 2) 其实在项目开始前，我就去腾讯租了一个服务器，原先想把项目放到服务器上，但是近期实在抽不出时间在全新的服务器上配置环境，未来有望将服务端假设到服务器上，让小伙伴们一起来体验与测试。

七、体会与建议

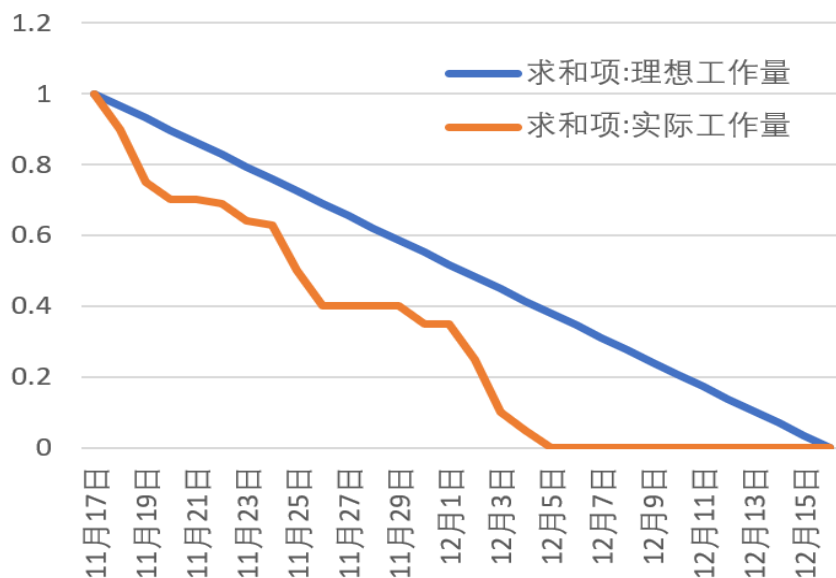


图 17：项目开发燃尽图

开发周期	两星期（100 小时左右）
代码量	2500+行

体会：这个项目花费了我非常长的时间，期间遇到了非常多令人奔溃的瞬间，甚至还想中途放弃去开始尝试另外的项目。但最终我还是饱含兴趣地完成了这个任务，其实因为使用过 QQ 与微信，从读大学前就非常想独立开发一个聊天程序，碍于时间，一直没有找到契机能够实现这个愿望，所幸遇到了这次的大作业，毫不犹豫地选择了聊天程序。

一开始设想直接架设 Web 端，成为像微信网页端一样的网页聊天程序，因为个人对可视化比较感兴趣，对自己要求也非常高，又参加过 Web 开发的许多项目，比较擅长这方面，而且 Web 更灵活，界面更美观。之后查阅资料得知 websocket 与通常所说的 socket 有区别，websocket 是作用于应用层，而普通的 socket 多用于传输层，因此将 Web 端转化成用户动态圈的形式，另外又增加了普通的客户端，用于聊天与文件传输等功能。

由于之前写程序比较喜欢将所有程序集中在一个文件里，为了改掉这个坏习惯，这次大作业采用模块的方式，有自定义类，函数的导入与调用等，多文件相互调用的形式。由于不是很习惯，因此过程中会产生一些逻辑上的混乱，但最终都被我一一克服，圆满完成了任务。

建议：我觉得大作业的要求蛮合理的，而且项目也比较有趣，也有很强的可扩展性，对我们的能力考察比较全面，也很能锻炼我们的逻辑能力与耐心，同时也很好地结合了课本的知识，做到了实践与理论的很好的结合。

参考文献与链接

- 【1】 Python Software Foundation, SocketServer Documentation
<https://docs.python.org/2/library/socketserver.html>
- 【2】 第一个程序 — The PyQt4 tutorial 0.1 documentation
http://www.gaulau.com/books/PyQt4_Tutorial/first_programs.html
- 【3】 Socket 通信的 Python 实现--平和的心--博客园
<https://www.cnblogs.com/ajianbeyourself/p/3771198.html>
- 【4】 Welcome | Flask (A Python Microframework) <http://flask.pocoo.org/>