

《计算科学导论》课程总结报告

|  |  |
| --- | --- |
| 姓 名 | 姚河清 |
| 学 号 | 1907010326 |
| 专业班级 | 计科1903 |
| 学 院 | 计算机科学与技术学院 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 课程认识  30% | 问题思考  30% | 格式规范  20% | IT工具  20% | Latex附加  10% | 总分 | 评阅教师 |
|  |  |  |  |  |  |  |

# **1 引言**

随着科技的发展，从一代计算机的发明，计算机的发展从未停住，在发展途中，产生了一些与计算机相关的学科，从20世纪30年代到60年代初，计算机科学与技术的研究与科学发展基本上处在萌芽状态，当时从事计算机科学与技术的科学家主要来自数学和电子科学领域。数学家最初的工作是围绕什么是计算开展理论探索，寻求计算的数学理论模型，弄清计算的极限。由于图灵和冯·诺依曼等人的贡献，使得存储程序式电子数字计算机在40年代诞生，人类使用自动计算装置代替人工计算和手工劳动的梦想成为现实。计算机科学与技术是研究计算机的设计与制造，并利用计算机进行有关的信息表示、收发、存储、处理、控制等的理论方法和技术的学科。计算机科学与技术专业的学习是一项十分艰巨的劳动，不少近年来成长起来的青年科学家和工程师都有同感。

在计算科学导论的课堂上，学到了与计算机科学与技术专业相关的更多知识，学到的不仅仅是与专业相关的基础知识，还学到了不少本来由其他课程的知识，如计算机操作命令与汉字编码等，也增加了一些不应该进入教学计划与课程体系的内容，如某些高级语言和某些数据库系统语言及其应用等。其中一部分是我们进入高年级如果必要，能够很自然、很容易地自学掌握的知识，在这堂课上，还给我们机会，让我们了解更多书本上没有的知识，让我们自主选题演讲，台上一分钟，台下不知道要准备多久，锻炼了我们查找资料，上台演讲的能力。

# **2 对计算科学导论这门课程的认识、体会**

国际著名计算科学家马纳曾在其著作《计算的数学理论》中一开头就写道：“什么是计算？我相信，世界上没有两个计算机科学家会就这一概念给出相同的定义。”计算科学是对描述和变换信息的算法过程，包括其原理、分析、设计、效率自动进行，什么不能（有效地）自动进行。本学科来源于对数理逻辑、计算模型、算法理论、自动计算机器的研究，形成于20世纪30年代后期。计算作为数学的研究对象已有几千年了。计算不等于数学，但数学确实起源于对计算的研究。计算的渊源可以深入扩展到数学和工程。数学为计算提供了理论、方法和技术，而工程为实际计算和应用提供了可以能行自动计算的设备，并为更有效地完成计算和应用任务提供了工程方法和技术。

比如：在其中介绍到了“算法、过程与程序”，“求解一个给定的可计算或可解的问题，不同的人常常会编写出不同的然而都是正确的程序，经过分析，科学家发现这里存在两个层面的问题，一个是与计算方法密切相关的算法问题”，这段话透露出计算科学所给我们介绍了计算机编程的原理分析，以及不同的算法会有相同的结果；另一个是程序设计的技术问题；其中还有计算科学与数学和其他相关学科的关系。

其中写到“计算科学可以在几乎所有的科学领域，甚至我们日常生活的各个方面找到应用，原因是计算（作广义理解）确实是人类最基本的智力活动之一。如果把每一个人看成是一台计算模型或者机器，那么，人的思想是一种计算，相应的行为就是计算的外在表现形式。一旦离开了计算，人类将一事无成。在各个比较大一些的领域范围内，计算科学已或多或少地找到了一些应用，这是计算机具体应用研究与开发同各个学科之间建立的联系。”这显示出计算科学来源于对数理逻辑等的研究，是人们日常必备“用品”，一旦离开计算，人们将一事无成。

**课堂分组演讲总结**

在这门课上，我们进行了分组演讲，我们学生通过查找资料并整合理解，将这些资料整合做为PPT展示给其他同学与老师，通过老师和其他同学的提问让我们更加对这门技术的了解，所谓“台上一分钟，台下十年功”。为了准备这份报告，我和我的伙伴准备了很久。

我和我伙伴的主题是“听歌识曲”。大多数的音乐软件都有这一功能，“听歌识曲”，是很多人会在日常用到的功能，当你在任何时候听到一首正在播放的音乐时，都可以拿出手机，打开听歌识曲功能进行识别，仅仅几秒钟就可以得到识别结果，然后把这首歌放入你的收藏列表。事实上，每天使用各类音乐软件听歌识曲的人数就超过百万级，听歌识曲技术就像是用户的音乐小助手，帮助用户迅速地从"歌海"中找到他们喜欢的那首歌。现在的听歌识曲相对于这门技术刚刚产生的时候已经进步了很多，但提到这项技术，必须要致敬的一家公司便是Shazam公司，以及使这一切成为可能的算法发明人Avery Li-Chun Wang。其实早在2002年时，Shazam的服务就已经存在，在那个非智能手机时代，使用识曲服务的办法是：用户拨打2580号码，然后使用通话传输录音，服务会在30秒后自动挂断，并通过短信的形式发送识别结果。2018年9月，Shazam公司最终被苹果公司以4亿美金收购。

**需要实现的功能**

输入一些音乐，提取这些音乐的声纹信息，存放于数据库中，用于识别音频片段

输入一个音频片段，大约10s到20s左右，该音频片段相对于原曲，允许混杂一些噪音信号，但是要求原曲子的声纹信息已经存在于声纹信息库中

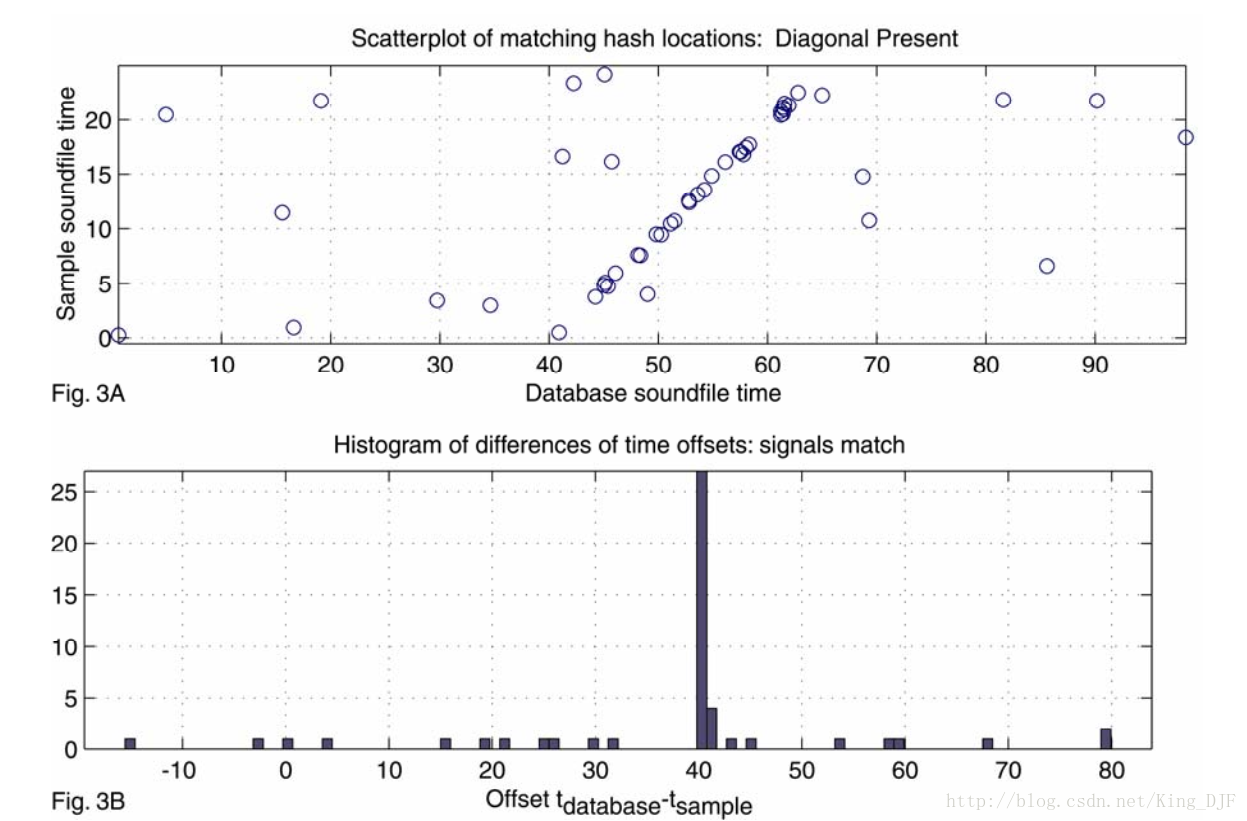
**算法原理**

算法主要分为两个部分，提取特征与检索方式：

提取特征第一步（这里我只说明结果，并不解释为什么用这个特征，具体解释论文中很详细）： 对音频做快速傅里叶变换，得到频谱信息；划分不同的频带，选取每一频带的幅度最大的频率点作为一个特征值的点保留下来。例如：总的频率范围是0-8KHz，那我如果平均划分8个区域：0-1，1-2，2-3…7-8，最后每一帧就会有八个极值点，其余的点就可以舍弃掉，后面我们只需要频谱的极值点。

提取特征第二步： 计算哈希值。我们并不是直接将极值点信息不做处理直接放进数据库中，而是需要将极值点信息与其他信息做一些组合，保留时间，频率等特征。这里引入两个概念，两个概念：锚点(anchor pointanchor pointanchor\ pointanchor point)与目标区域(target zonetarget zone target\ zonetarget zone)。锚点即当前点，锚点的目标区域是指与每一个锚点相对应的一片区域（例如时间线上排在后面的帧）。之后根据锚点与目标区域的相应特征提取出特征值，表征相应的信息，这个特征值的获取方式为：[f1:f2:Δt:t1](Δt=t1−t2)[f1:f2:Δt:t1](Δt=t1−t2) [f1:f2:\Delta t:t1](\Delta t = t\_1 - t\_2)[f1:f2:Δt:t1](Δt=t1​−t2​)，一帧信号可以有多个特征值，因为同一帧内本身就存在不止一个极值点，每个极值点都可以作为锚点，这主要取决于你的组合规则。

之后，将所有提取出的信息存放进数据库中，之后检索时使用检索过程：检索过程比较简单，对待识别片段做同样的特征提取，得到一些特征值，然后用这些特征值去数据库中做匹配，最匹配的歌曲就是识别结果，其实，比较接近的算法就是倒排索引。可以用一个图说明：



如图所示，上面的子图的纵轴是待识别音频片段，横轴为一数据库中完整音频，若该片段来自这个音频，互相匹配的特征值点便能构成一条倾斜的直线，显然线上每一点的横纵坐标之差（可以看作是音频片段的偏移值）相等。由此可以得到一个柱状图，横坐标是一首歌中存在的差值，纵轴表示差值的数量，可以看到差值为40的点最多，也就是说，如果这个音频片段属于这个歌曲，它最有可能是以40为起始点截取下来的。故，把40这个点位置上的数量，作为音频片段与它的匹配值。匹配值最高的就是最后的检索结果。

**以上是比较古老的技术，现在来说，计算机的发展越来越快，对于听歌识曲，也是有了更新的技术，听歌识曲的实现是基于语音识别，对于语音识别，BP神经网络在语音识别上能“有所作为”。**

BP神经网络是一种多层前馈神经网络,该网络的主要特点是信号前向传递,误差反向传播。在前向传递中,输入信号从输入层经隐含层逐层处理,直至输出层。每一层的神经元状态只影响下一层神经元状态。如果输出层得不到期望输出,则转入反向传播,根据预测误差调整网络权值和阈值,从而使BP神经网络预测输出不断逼近期望输出。BP神经网络的拓扑结构BP神经网络可以看成一个非线性函数,网络输入值和预测值分别为该函数的自变量和因变量。当输入节点数为n,输出节点数为m时,BP神经网络就表达了从n个自变量到m个因变量的函数映射关系。BP神经网络预测前首先要训练网络,通过训练使网络具有联想记忆和预测能力。BP神经网络的训练过程包括以下几个步骤。  
步骤1: 网络初始化。根据系统输入输出序列(X,Y)确定网络输入层节点数n、隐含层节  
点数1,输出层节点数m,初始化输入层、隐含层和输出层神经元之间的连接权值ω  
ij,ωjk,初始化隐含层阈值a,输出层阈值b,给定学习速率和神经元激励函数。  
步骤2: 隐含层输出计算。根据输入向量X,输入层和隐含层间连接权值ωij  
以及隐含层阈值a,计算隐含层输出H。  
步骤3: 输出层输出计算。根据隐含层输出H,连接权值ωjk和阈值b,计算BP神经网络预测输出O。  
步骤4: 误差计算。根据网络预测输出O和期望输出Y,计算网络预测误差e。  
步骤5: 权值更新。根据网络预测误差e更新网络连接权值ω  
步骤6: 阈值更新。根据网络预测误差e更新网络节点阈值a,  
步骤7: 判断算法迭代是否结束,若没有结束,返回步骤2。

语音特征信号识别

语音特征信号识别是语音识别研究领域中的一个重要方面,一般采用模式匹配的原理解决。语音识别的运算过程为:首先,待识别语音转化为电信号后输入识别系统,经过预处理后用数学方法提取语音特征信号,提取出的语音特征信号可以看成该段语音的模式。然后将该段语音模型同已知参考模式相比较,获得最佳匹配的参考模式为该段语音的识别结果。语音本案例选取了民歌、古筝、摇滚和流行四类不同音乐,用 BP 神经网络实现对这四类音乐的有效分类。每段音乐都用倒谱系数法提取 500 组 24 维语音特征信号,提取出的语音特征信号。

模型建立

BP 神经网络构建根据系统输入输出数据特点确定BP神经网络的结构,由于语音特征输入信号有24维,待分类的语音信号共有4类,所以BP神经网络的结构为24—25—4 ,即输入层有24个节点,隐含层有25个节点,输出层有 4 个节点。BP 神经网络训练用训练数据训练 BP 神经网络。共有2000组语音特征信号,从中随机选择 1500 组数据作为训练数据训练网络,500组数据作为测试数据测试网络分类能力。BP神经网络分类用训练好的神经网络对测试数据所属语音类别进号。BP神经网络构建根据系统输入输出数据特点确定 BP 神经网络的结构,由于语音特征输入信号有24维,待分类的语音信号共有4类,所以BP神经网络的结构为24 — 25 — 4 ,即输入层有 24 个节点,隐含层有 25 个节点,输出层有 4 个节点。BP 神经网络训练用训练数据训练 BP 神经网络。共有 2000 组语音特征信号,从中机选择 1500 组数据作为训练数据训练网络,500 组数据作为测试数据测试网络分类能力。

# BP神经网络分类用训练好的神经网络对测试数据所属语音类别进 行分类。

# **3 进一步的思考**

结合学习的计算科学知识，对分组演讲涉及的问题作进一步的思考。

在分组演讲中，我们小组查找资料的我查找到的资料文献并不能跟上现在的技术，太落后了，及我在上面所提到的“古老技术”，经过老师的提醒，我们小组在课后再去网上查找了资料，终于找到了新的技术，更加先进的技术，更加完善的技术，及我在上面介绍的BP神经网络进行语音识别的技术，（BP(back propagation)神经网络是1986年由Rumelhart和McClelland为首的科学家提出的概念，是一种按照误差逆向传播算法训练的多层前馈神经网络，是应用最广泛的神经网络。），在未来的路上，我们还更应该深度的学习，应该在网上查找资料时，应该多进行对比。

# **4 总结**

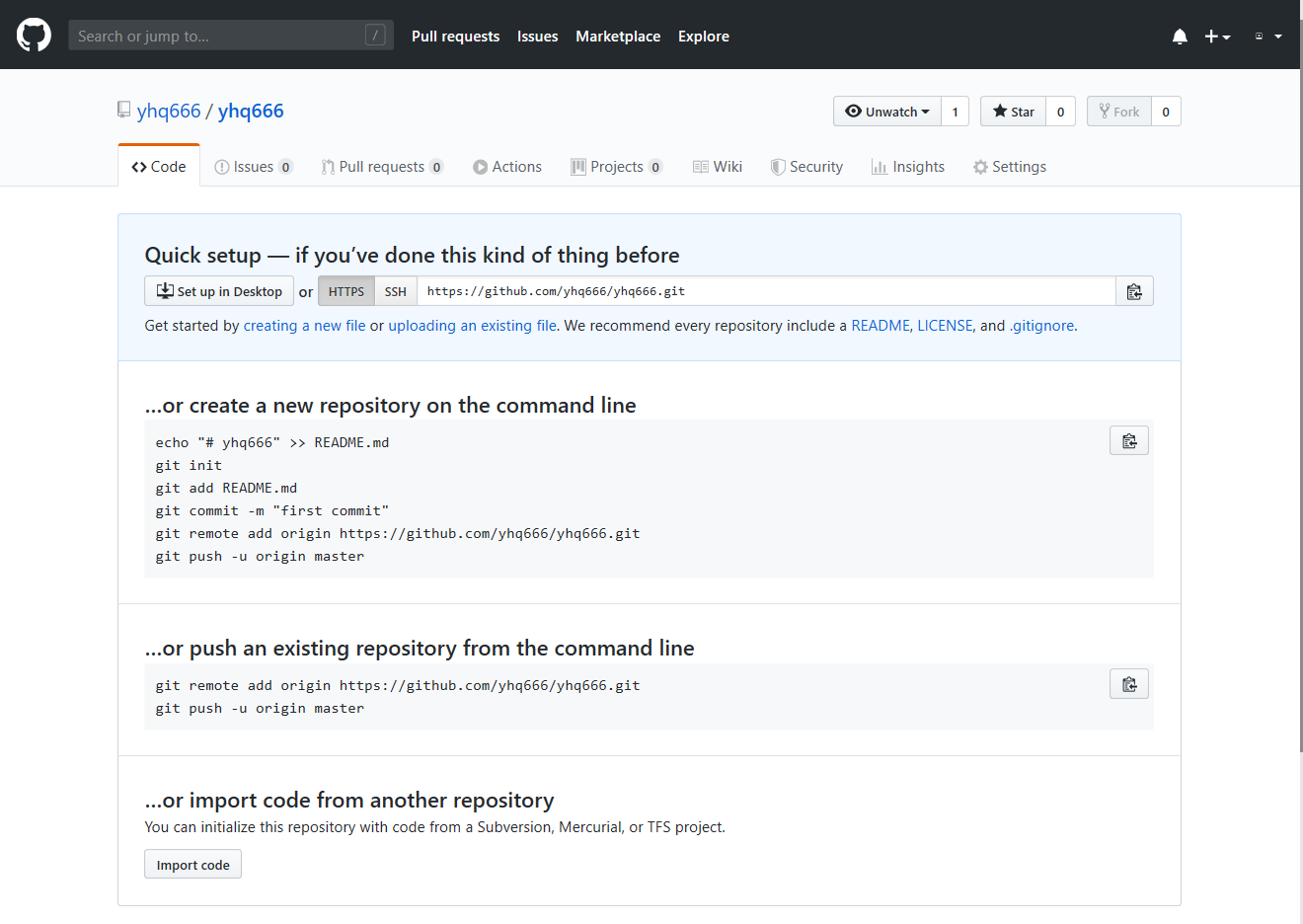
**在王晨，宋亮，李少昆所编写的《工业互联网平台：发展趋势与挑战》中提到：“工业互联网平台是制造业与云计算、大数据、 人工智能等新一代信息技术高度融合的产物，也是 先进制造业与互联网结合的新事物，需要广大科研 人员和产业工作者在实践中一边摸索一边突破。作 为制造业大国，我国拥有全球最多的工业设备，时 刻产生着海量的工业数据，拥有丰富的互联网生态 以及大量的工业与信息化人才。我们应该充分利用 这一条件，创新管理思想，重构产业生态，提升中 国制造在全球产业链分工中的地位。依托工业互联 网平台，推进工业互联网的深度应用与制造业转型 升级，力争在新工业革命时代实现“换道超车”。”**中国已经在第三次工业革命中占居鳌头，作为新中国发展的接班人，我们应当努力学习我们现在所学习的课程与技术，不能够只学课上所学，更应当学课上不能学到的知识，这样，我们这一代人才能让祖国更加的强盛，对于计算机科学与技术这门专业，是当今时代火热的专业，也是区分一个国家在这次工业革命中发展状态的一种体现，计算科学导论让我们更深刻的认识到这门学科所学所用以及如何学习，有助于我们正确理解学科中的思维方法和解决问题的思维。

# **5 附录**

## **Github**

个人网址：https://github.com/yhq666/yhq666

个人网站截图：



## **观察者**

给出对应的截图：



## **学习强国**

对应的截图：



## **哔哩哔哩**

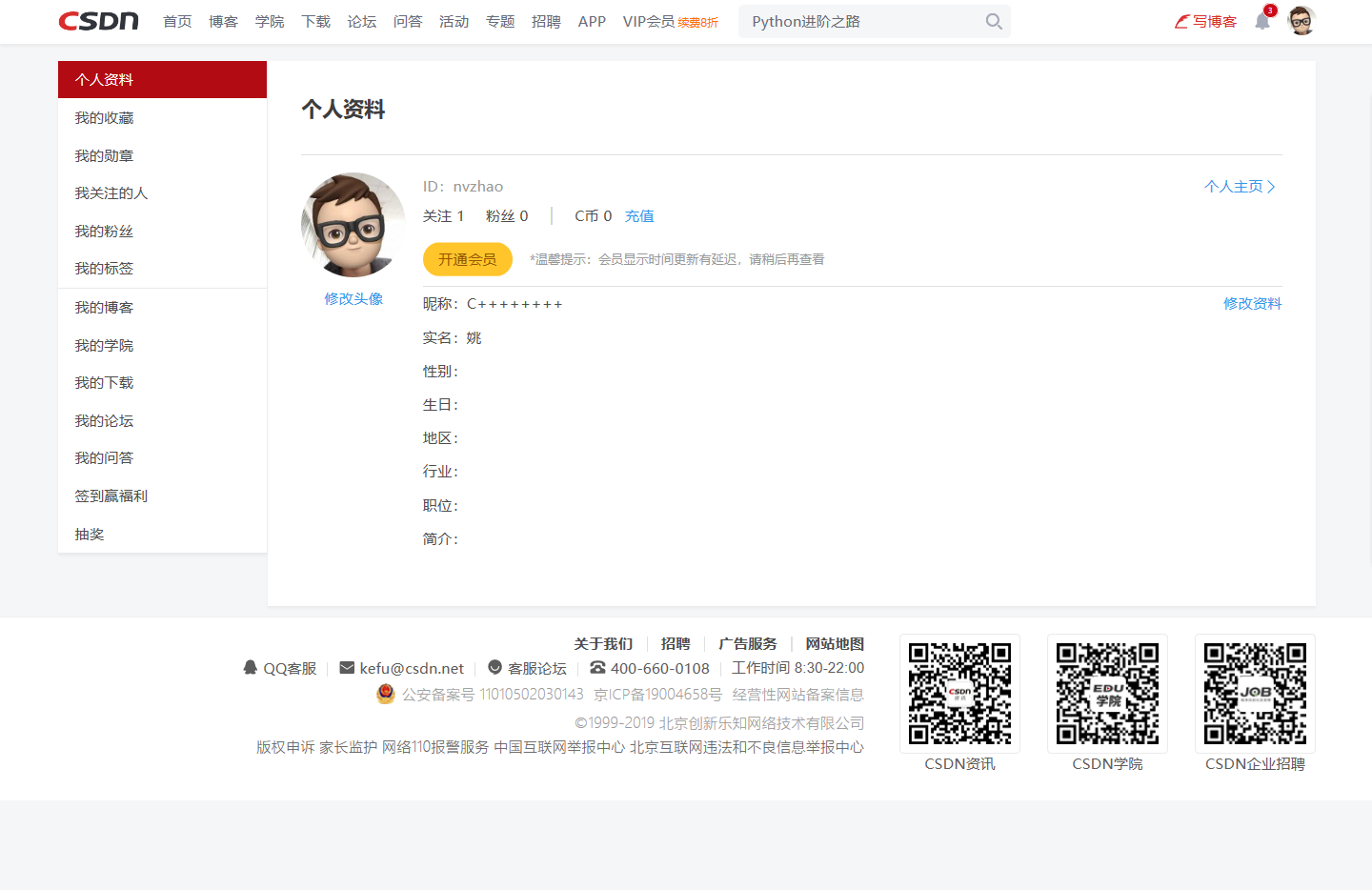
对应的截图：



## **CSDN**

个人网址：https://i.csdn.net/#/uc/profile

个人网站截图：



## **博客园**

个人网址：https://home.cnblogs.com/u/1914068/

个人网站截图：



## **小木虫**

个人网址：muchong.com/bbs/space.php?uid=20368039

个人网站截图：

