计算机网络实验报告

--BGP篇

学校：哈尔滨工业大学

班级：11403202

学号：1140320206

姓名：霍峻杰

实验3.1

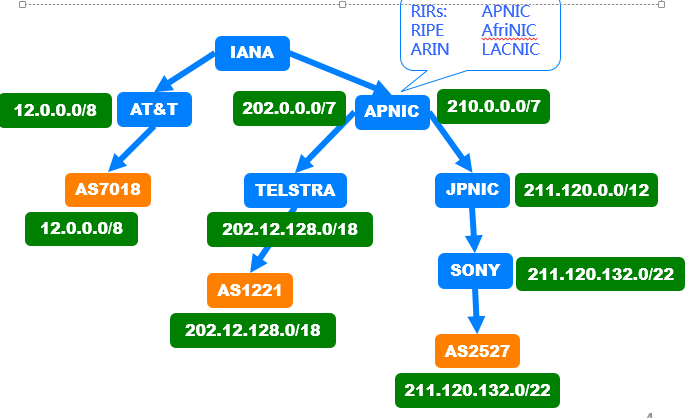
一、实验题目:

* 分析某一天(2015年你生日)的一张路由表，回答以下问题（不限于）：
  + Internet多少个AS，多少IP前缀？
  + 每个AS所声明前缀的数量，所声明前缀中IP地址数量分布？（绘图）
  + 每个AS的节点度（有几个邻居）分布？（绘图）
  + AS节点度与声明前缀数量有关系吗？（绘图）
  + 其他感兴趣的问题，例如判断两个相邻AS间是P2C还是P2P关系？
* 开发语言：Python或Perl等脚本语言
* 绘图工具gnuplot，Python绘图工具包、或R语言绘图工具包
* 实验报告：数据获取和分析过程，你的发现和认识
* 计分：报告7分+实验课代码检查及提问3分

二、实验原理：

BGP全称Border Gateway Protocol，边界网关协议，也就是常说的路由协议 自治域(Autonomous System, AS)：在一个或多个网络运营商的控制下的一组相连的IP路由前缀集合，该集合对互联网呈现出一个共同明确的路由策略。

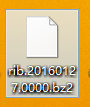
IP前缀与自治域号(ASN)分配



本实验通过对路由表的统计整理进行绘图，验证一些结论，实验结果如下。

三、实验结果:

1. 下载好我的生日（160127）当天的原始路由数据：



2. 通过老师的bgpdump，将路由数据整理成一条一条的记录，写入data.txt

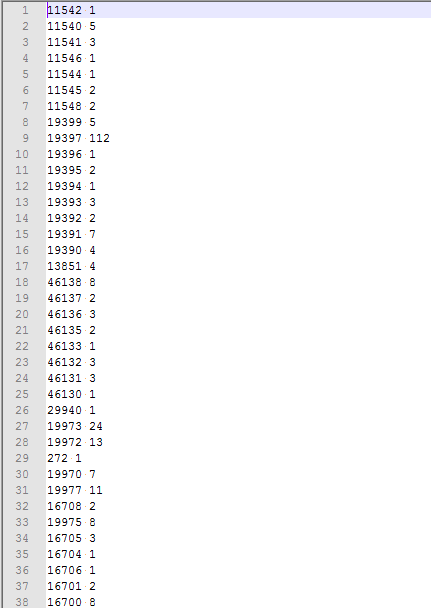
命令为bgpdump –M rib.20160127.0000.bz2 –O data.txt

解析速度还可以，文件数据比较多，当我一打开我的data.txt，我的虚 拟机就死机了，还好数据写入没问题，于是就拿出python大法开始数 据的处理。

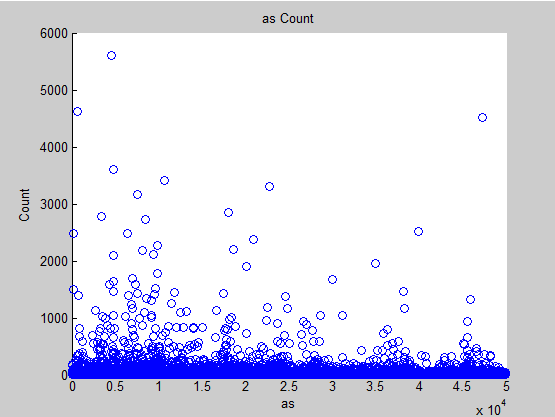
3. 这里在处理数据的时候由于会遇到很多的重复元素，所以我们就想到了 把他们都塞进一个集合，再求集合的长度，就是想要的不重复的as，ip 前缀个数，等等。计算结果见下图：as：53314，ip前缀：614556



4. 每个as声明的ip前缀数量，我把统计完的数据写入了一个apcnt.txt 中：每行表示，一个as及其对应的ip前缀个数



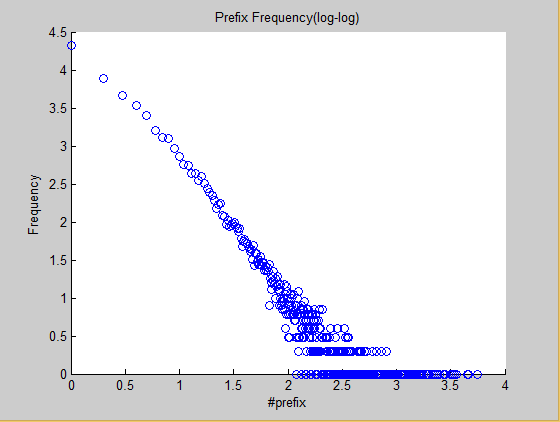
每个as所声明的前缀数量绘图：



分析绘图结果：

我们可以看到as的ip前缀数量大多都分布在300以下，少数as的ip前缀数量会达到很高，达到5600左右。

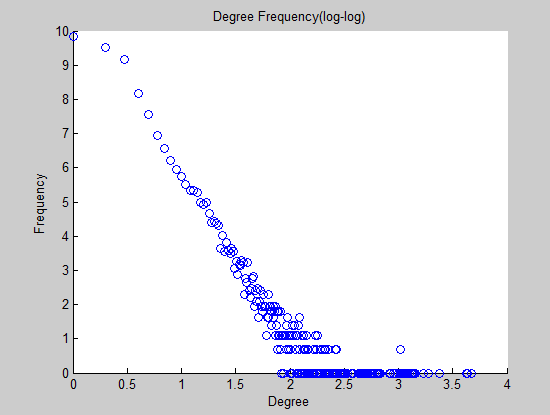
Ip地址数量分布就是统计ip前缀相同数量出现的频率，我将其统计在prefixcnt.txt中，用matlab对数据进行绘图：



分析绘图结果：

Ip前缀数量集中在较小的数量。两者的log-log成负相关。

5. as节点度（有几个邻居）分布如下图（开始时找错了，把整个aspath的as节点当成互为邻居了，慢而且结果不对）

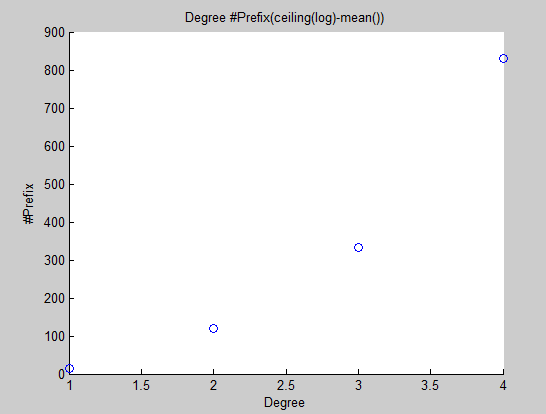


分析绘图结果：

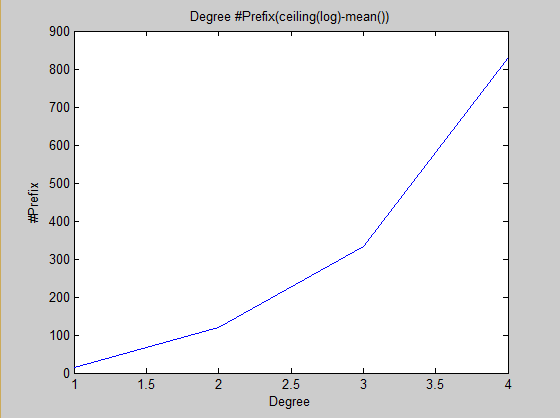
As节点度频率高的节点度越低。两者log-log成负相关（和上面的绘图结果很像）。

6. AS节点度与声明前缀数量的关系，这里按照老师讲的分成几个桶去统计

散点图：



趋势图：

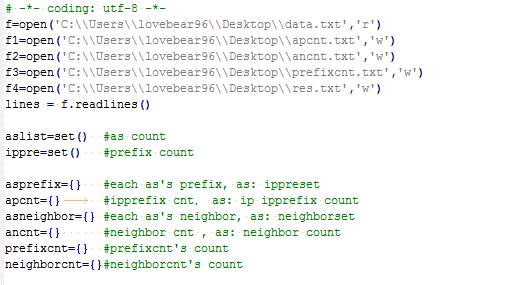


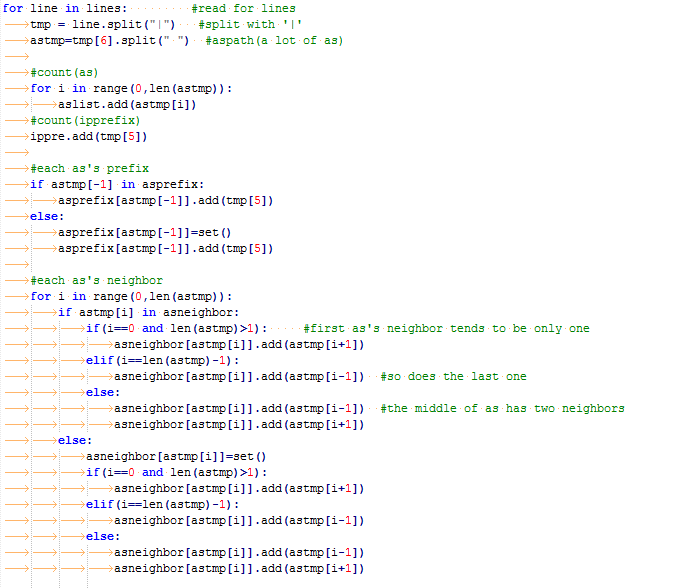
分析绘图结果：

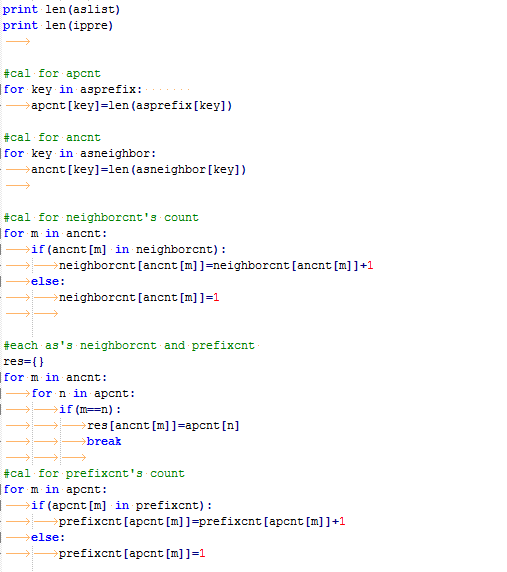
as节点度越高，生命的ip前缀越多。

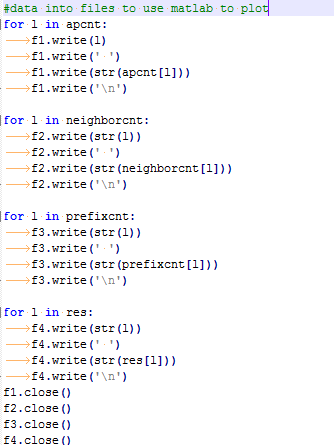
四、实验代码详解

实验代码解释以注释形式给出









Matlab绘图代码就不给出了，就是读取刚才写入的数据，绘图。

实验总结：

这个实验思路上不是很难，但是对大数据的处理上要求比较高，代码的数据结构和算法的要求会比较高，因为需要在最少的时间内计算出所有的结果，我的代码还是运行的挺快的，整体代码运行时间不超过10min。

通过对真实的路由表进行整理分析绘图，验证结果，来更加深刻的理解bgp，as的ip前缀，节点度的关系等。

同时，还复习了用matlab绘图，matlab在画图方面还是非常的方便的。