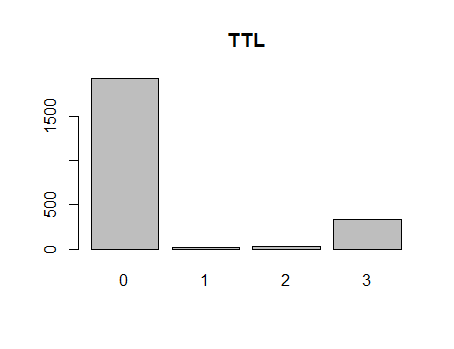
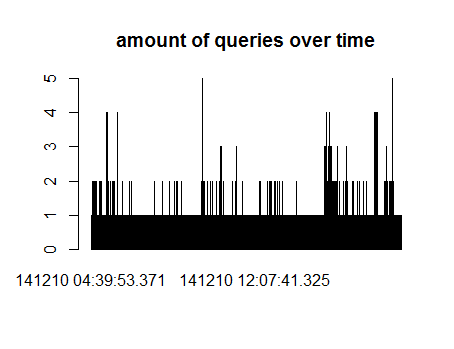
Kiera Crist ([kcrist@u.rochester.edu](mailto:kcrist@u.rochester.edu))

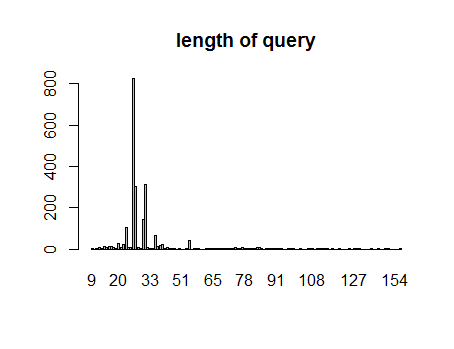
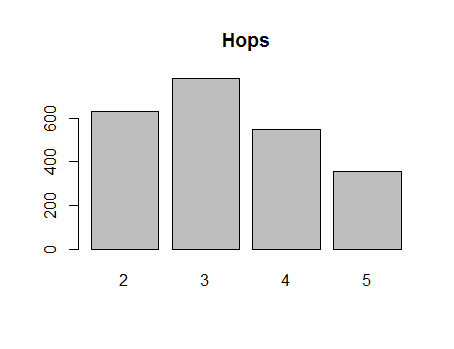
The purpose of this project was to gather queries and analyze them and see any interesting results.

So I used the program phex to connect to the gnutella network and log queries. This was done through the error logging system. This is different from the file i/o I was considering using when I wrote the preliminary report, because I did not take into account the complexity of synchronization needed. The code addition for the query logging was quite small, less than 10 lines of code. The analysis, however, took a bit more time.

I tried to use R to analyze the queries. I did not take into account, however, the difficulty in analyzing strings in R. There are ways, but they are beyond my skills to use. I was able to make graphs to demonstrate the more numerical values, like ttl or hops, but not the search queries. I ended up doing a rough analysis of the search queries by hand. This was done through opening the log file and searching for common words I saw. And also copying and pasting in a separate file to count through line numbers, common categories of searches.

Some interesting graphs





Some interesting results from the search queries

There were 2315 search results total.

Whatisnewxoxo -1739 times. I believe there is a built in button that searches this for you.

There were a lot of ?? ??? looking search queries, which most likely were from characters that weren’t encoded right.

?? -435

Apart from the whatisnew searches, the other ones were mostly for music and videos, or porn. Some searches I could not tell what they were trying to search for; there were unfamiliar abbreviations.

Music and video related searches -47

There were some searches that were not in English, at least 14 that were clearly not English and there may have been more that I missed. The question mark searches were probably also not English, but since I couldn’t tell what they were, I did not count them.

There were a lot of porn related searches, which is not surprising. What was unexpected was how much of it was searching for underaged porn.

Porn related searches-82

Teen-6

Yo – 42 yo as in short for year old, from people searching for porn of kids of a certain age.

Pedo -13

Avi-86

20xx -53