

Techniques for Monitoring and Mapping Data in Peer-to-Peer Networks

Raymond C. Borges Hink

Universidad del Turabo

Research Alliance in Math and Science

Computational Sciences and Engineering Division, Oak Ridge National Laboratory

Mentor: Dr. Robert Patton

<http://sites.google.com/a/g.ornl.gov/borges-r/>



Background

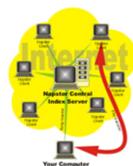
The World Wide Web has made great things possible but it's also enabled crime to reach unprecedented heights. Crime has evolved into cybercrime. Illegal activities can be carried out with a certain anonymity where the crime cannot be easily monitored or traced back to the source. With the creation of peer-to-peer systems like Napster it's become even more difficult to track information throughout the internet.

In 2001 peer-to-peer giant Napster was sued for copyright violations, and this gave rise to decentralized file sharing networks being developed. The focus of this project was to develop a platform to monitor p2p decentralized networks such as BitTorrent as well as trace the information geographically and or demographically back to the sources.

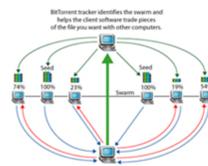
Research Objectives

- Study multiple p2p file sharing networks
 - Determine which are most used and predict which will be used most in the near future
 - Determine for which it would be possible to obtain the IP addresses of the users
- Develop an application to monitor these networks
 - Track IP addresses in p2p networks of seeds and peers
 - Map IP addresses to geographic locations
 - Obtain as much information about IP address as possible, such as the ISP
- Consider future goals of the application
 - Architecture scalability

Napster and BitTorrent Architecture

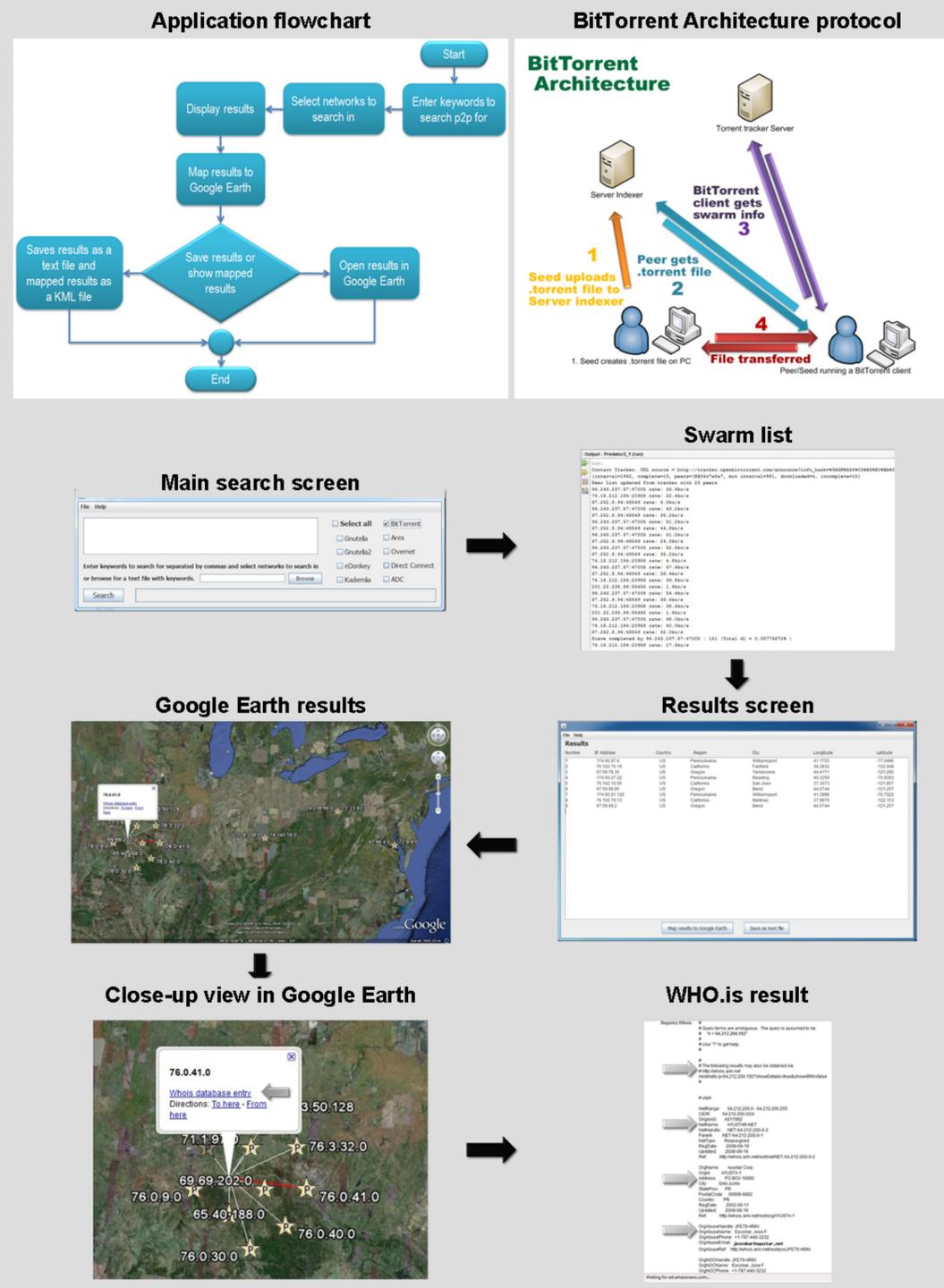


1998-2010 HowStuffWorks, Inc.



1998-2010 HowStuffWorks, Inc.

Methodology



Results

An HTTP get method to scrape the tracker server followed by opening a TCP port connection to the server obtains the peer list for that torrent. By using the peer list and a MySQL IP to country database the IP's can be plotted on Google Earth using latitude and longitude. Finally, each IP is hyperlinked to the Who.is database so a user can click on the pinpoint in Google Earth and obtain the ISP information for that IP address.

The results for BitTorrent differ from the rest of p2p networks in that the peer list is not dependent on hops through the network. False positives can cause problems, tracker peer lists intentionally or unintentionally can be corrupted with bad IPs or inserted to confound government agencies.

Conclusions

BitTorrent is a very robust and evolving overlay network protocol and will most likely continue to grow since more and more organizations are using peer-to-peer networks for its speed and reliability. As peer-to-peer networks become fully decentralized, tracking data will become more complex.

Future Research

Future work will be aimed at determining the feasibility of expanding the platform to more file sharing networks and developing an extension to the application to allow for tracking data in the two new models: (1) Distributed Hash Tables (DHT), and (2) Peer Exchange (PEX) also known as the Gossip protocol.

Selected References

- Baptiste Dubuis "jBitTorrentAPI," http://icwww.epfl.ch/~portabel/studentprojects/files/jBittorrentAPI_repo.rt.pdf, 2007.
- Carmen Carmack "How BitTorrent Works," <http://computer.howstuffworks.com/bittorrent2.htm>, 2005.