

# TJHSST Senior Research Project

## End-to-End Publication Using the Bittorrent P2P Filesharing Protocol

### 2006-2007

Andrew Wang

November 3, 2006

#### **Abstract**

Bittorrent is a promising peer-to-peer network that always allows for fast download speeds despite the number of peers downloading the file. Currently, there exist tools to make .torrent files, tools to "track" the peers downloading the file, and tools to host .torrent files. This project aims to unify this process by making an end-to-end software suite that simplifies the process of publishing a file on the Bittorrent network for download. The key to this will be automating and streamlining the process from the perspective of the user. It will involve a complete implementation of the Bittorrent protocol, including encoding torrent files, peer-to-tracker and peer-to-peer communication, and a greater understanding of the benefits and detriments of the Bittorrent protocol.

**Keywords:** Bittorrent, Peer-to-peer, Linux, Publishing, Tracker, Peer

## **1 Introduction**

### **1.1 Scope of Project**

The scope of the project is broad, because it aims to be a complete solution to publishing files using Bittorrent. It will have to handle all aspects of the

Bittorrent protocol, from processing the file to make the .torrent metadata file to hosting and tracking the .torrent file and possibly a download client.

This project started out as writing a better download client, but there is already a plethora of download clients available and finding an easy to use and mature Bittorrent development library proved difficult. It would also be difficult to surpass the quality and features of other download clients developing by myself, and users would be unlikely to switch their choice in download clients unless there was a very good reason.

## **1.2 Expected Results**

The research of this project involves the Bittorrent protocol. Given the specification of the protocol given on the Bittorrent website, I aim to implement the server side aspect. This means encoding and dencoding of .torrent metadata files, hosting of .torrent metadata files, and subsequent tracking and peer-to-tracker communication of download peers by the tracker. It will involve the subject areas of large scale networks, encoding and decoding algorithms, and peer-to-peer communication.

## **1.3 Type of Research**

This project will be use-inspired basic research, because the underlying goal is to gain an understanding of the benefits and limitations of the Bittorrent protocol. There are great practical implications for the end product of this research, but ultimately the project was started to gain an introduction to networking and peer-to-peer technology.

# **2 Background**

Bittorrent is an up and coming filesharing protocol that has emerged in the wake of illegal services such as Kazaa, Napster, or Bearshare that have since been shutdown or forced to end their copyright violations. Bittorrent is a much more legally feisable filesharing protocol than previous attempts, because there is no copyrighted content to be stored on centralized servers that can be subpoenaed or seized, and it has become extremely popular for independent movie makers and other people that need to distribute their legal content without buying an expensive server. A movie distribution method

using Bittorrent is also being developed by major movie companies, as they too see the benefits of peer-to-peer technology.

The whitepaper written by the creator of Bittorrent available on the official Bittorrent website is the most useful reference for this project. Additionally, there is a page on [wiki.theory.org](http://wiki.theory.org) that takes and expands upon the official protocol specification that is useful for more detailed help. These two documents give a total description of all aspects of Bittorrent, and will be the only necessary references throughout the project's extent.

## 3 Procedure and Methodology

### 3.1 Planning

The languages used in this project will be Python and probably PHP for the web interface of the software suite. Performance is not an issue because the processor requirements are low. A webserver of some sort will be needed to host the .torrent files, but this can be done with a third-party solution, or a basic server can be written if needed. The stages of this project can be split up into a number of clearly defined steps:

1. Study and implementation of encoding .torrent metadata files. These files are "bencoded," which is a translated form of dictionaries, lists, strings, and integers. This will also coincide with studying of the various kinds of metadata stored in .torrent files as well as an interface for creating these .torrent files.
2. Study and implementation of a Bittorrent tracker. This is a much more involved step, and further research will be required to split this up into smaller tasks. Currently, the tracker has to process the .torrent metadata file, and then receive messages from peers and process and reply to these messages.
3. The rest of the project is making a web interface that can take a file, make a .torrent metadata file for it, add it to the tracker, and then be the initial uploader for the file. This should be a seamless process.

## 3.2 Testing and Analysis

Testing of encoding .torrent files is done using examples on the Bittorrent website and others. The program will transparently handle errors because it will simple treat the invalid input as a string. This will result in an incorrect .torrent file though, so I will build in checking when I make the frontend for making .torrents. Performance is also not an issue for the bencoding program because it takes minimal time even with the use of Python. I also will validate the torrent files I create against the torrent files created by third party applications and try publishing it with a third party tracker.

```
[03:42:11] awang::hermit $ ./torrentfile.py
```

```
torrentfile.py 0.2
Enter a string: hello world!
Bencoded string: 12:hello world!
Enter an integer: 12345
Bencoded integer: i12345e
Enter a list: apple,orange,pear,grape
Bencoded list: l5:apple6:orange4:pear5:grapee
```

```
Example bencoded dictionary
Dictionary: {'myname': ['andrew', 'wang'],
'dozen': 12, 'apple': 'red', 'banana': 'yellow'}
Bencoded: d6:myname16:andrew4:wange5:dozeni12e5:
apple3:red6:banana6:yellowe
```

## 3.3 Goals and Requirements

The goals for this project are as follows:

1. Easy to use, automated front end for the user
2. Bencoded .torrent file creation and parsing.
3. Correct implementation of tracker software and tracker-peer communication

## 4 Expected Results

I expect a complete, easy to use frontend that will handle and automate as much of the process of publishing a file through Bittorrent as possible. These results will be represented with screenshots and flowcharts describing the process. The website should be easy enough to use and well designed so that the proper steps to take are obvious.

This could be a very useful way to easily distribute files within a bandwidth limited environment. It could be useful to any project that needs to distribute large files or other people who want to use the Bittorrent protocol, because it will be a complete implementation.