

## Exposé zur Bachelorarbeit

# Analysis of BitTorrent Trackers and Peers

Stefan Schindler\*

14. Januar 2015

## 1 Motivation

Bram Cohen veröffentlichte 2008 das BitTorrent-Protokoll, mit dem Dateien in einem beliebig großen Computernetzwerk effizient verteilt werden können [1]. Dieses Protokoll macht heute 15 % des europäischen Internetverkehrs aus; im asiatisch-pazifischen Raum liegt der Anteil sogar bei 27 %. Übrige Filesharing-Technologien belegen ca. 2 % [6].

Da BitTorrent größtenteils zur illegalen Verbreitung von urheberrechtlich geschützten Werken verwendet wird [7], ist es von großem Interesse die Aktivitäten über das BitTorrent-Protokoll verstehen und auswerten zu können. Oft wird die Zahl an eindeutigen Peers als Maßeinheit für die Verbreitung eines Torrents herangezogen [3], wohingegen in dieser Studie tatsächlich beobachtete Downloads gewertet werden.

## 2 Problemstellung

Ziel ist es, den Umfang von unrechtmäßigen Downloads sowie deren geographische Verteilung und Geschwindigkeit verlässlich zu erfassen. Größtes Hindernis dabei ist jedoch die Tatsache, dass es kein zentrales Verwaltungssystem gibt. Stattdessen existiert eine Vielzahl an öffentlichen und privaten Trackern, wobei letztere von untergeordneter Relevanz sind [2]. Aufgrund dieser Unübersichtlichkeit muss für die Auswertung eine repräsentative Auswahl von Torrents gefunden werden.

Die rechtlichen Aspekte der Datenerfassung werden kurz betrachtet, insbesondere die gesetzliche Notwendigkeit, Torrents mit urheberrechtlich geschützten Werken ohne inhaltlichen Datenaustausch zu beobachten.

---

\*E-Mail: stefan@kaloix.de, Matrikelnummer: 21676746

## 3 Ziel

### 3.1 Software

Es wird ein BitTorrent-Client entwickelt, der zu einem gegebenen Torrent Statistiken über alle Teilnehmer anzeigt. Dieser kann auf bestehender freier Software aufbauen und ist durch einen modularen Aufbau mit weiteren Analysemodulen erweiterbar. Als Eingabe wird eine Torrent-Datei entgegengenommen und vom darin eingebetteten Tracker-Server eine Liste teilnehmender Peers abgerufen. Außerdem werden Magnet-Links [4] zur Identifikation von Torrents akzeptiert und das DHT-Protokoll [5] zur Suche weiterer Peers verwendet.

Zu den Peers werden folgende Daten ermittelt: Der geographische Standort, die bereits heruntergeladene Datenmenge sowie die Downloadrate. Der Standort wird mit Hilfe einer lokalen Bibliothek anhand der IP-Adresse direkt festgestellt, und die IP-Adresse gelöscht. Wenn ein Peer seine Sitzung beendet, wird der Download registriert. Um unterbrochene Sitzungen zu erkennen, werden Peers anhand ihres mittels Reverse DNS bestimmten ISPs, des verwendeten Clients sowie dem Muster heruntergeladener Segmente wieder erkannt. Für einen gegebenen Messzeitraum kann ein Bericht erstellt werden.

Die Software ist in der Lage, alle Statistiken anzuzeigen, ohne tatsächlich Inhalte des Torrents herunterzuladen oder an andere Teilnehmer wieder zu verteilen. Dadurch können auch Torrents mit urheberrechtlich geschützten Inhalten in einem legalen Rahmen analysiert werden.

### 3.2 Auswertung

Es werden die verbreitetsten Tracker ausgewählt und von diesen eine möglichst umfangreiche Auswahl der beliebtesten Torrents untersucht, um ein repräsentatives Ergebnis zu erhalten. Im Untersuchungszeitraum wird für alle Torrents die Downloadzahl festgestellt.

Das Ergebnis wird nach Kategorien (z. B. Musik, Filme, Software etc.), Legalität und geographischen Gebieten aufgeteilt.

## 4 Meilensteine

1. Recherche (1 Monat): Suche nach passenden Programmbibliotheken für BitTorrent und IP-Ortung, Spezifikation von Datenmodellen und Ausgabeformaten
2. Software-Entwicklung (1,5 Monate): Schreiben des Programms in Python
3. Datenerfassung und Auswertung (1 Monat): Auswahl von Torrents verschiedener Tracker, Sammlung von Statistiken, Evaluation der Ergebnisse
4. Verfassen der schriftlichen Arbeit (1,5 Monate): Entwerfen und Formulieren der Bachelorarbeit

## Literatur

- [1] Bram Cohen. *The BitTorrent Protocol Specification*. BEP 3. 2008.
- [2] *Comparison of BitTorrent sites*. Version 23:00, 14. November 2014. Wikipedia, the free encyclopedia. 2014.
- [3] Anders Drachen, Kevin Bauer und Robert WD Veitch. “Distribution of digital games via BitTorrent”. In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. ACM. 2011, S. 233–240.
- [4] Greg Hazel und Arvid Norberg. *Extension for Peers to Send Metadata Files*. BEP 9. 2008.
- [5] Andrew Loewenstern und Arvid Norberg. *DHT Protocol*. BEP 5. 2008.
- [6] Sandvine. *Global Internet Phenomena Report 1H 2014*. Study. Sandvine, 2014.
- [7] Paul A Watters, Robert Layton und Richard Dazeley. “How much material on BitTorrent is infringing content? A case study”. In: *Information Security Technical Report* 16.2 (2011), S. 79–87.