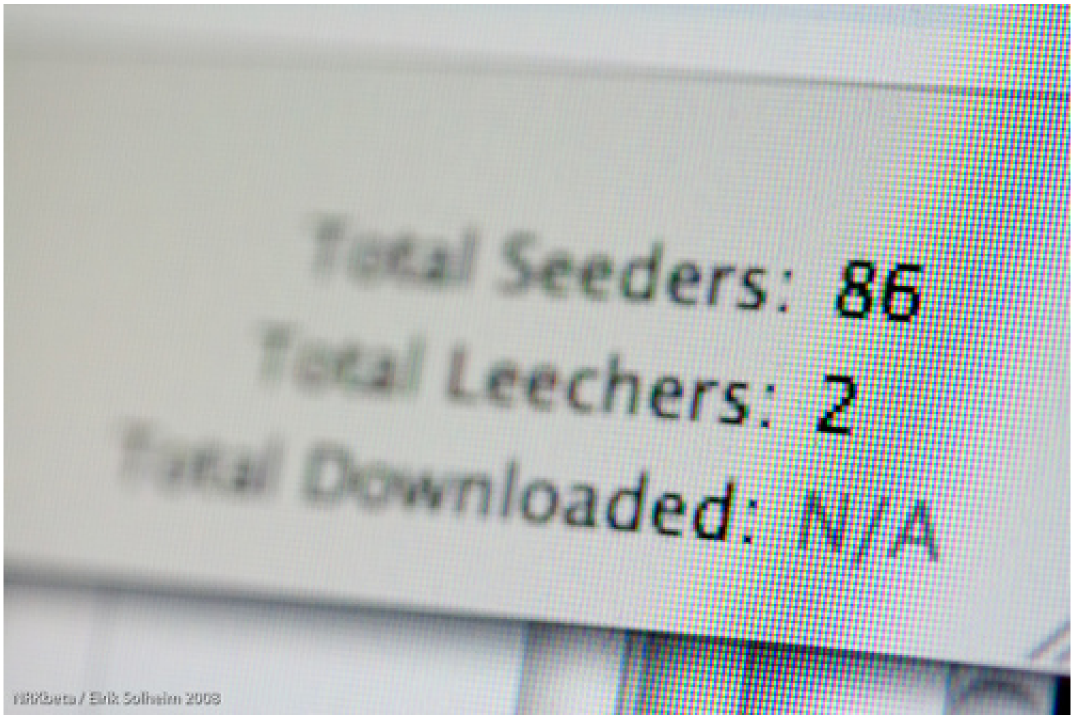
BitTorrent and the digital fingerprints we leave behind

April 10 2015, by Robert Merkel



Finding those responsible for illegal downloading on BitTorrents may prove a challenge. Credit: Flickr/nrkbeta, CC BY-SA

The Dallas Buyers Club LLC v iiNet Limited piracy court case raises many questions about what sort of trail people leave when they use technology to make illegal copies of movies and other copyrighted material.

The Federal Court of Australia has ruled that iiNet and a number of other internet service providers (ISPs) are required to disclose details of 4,726 of their account holders alleged to have been used to illegally download the movie over the internet via BitTorrent.

BitTorrent is a protocol (i.e. a detailed procedure) for transferring files – including, but not limited to, music and video files – between networked computers.

It was invented in the early 2000s by Bram Cohen, a programmer who went on to found a company called BitTorrent Inc that produces official BitTorrent software, which implements the protocol. Many other organisations have written compatible software.

To explain what BitTorrent does and how its users can be traced, it's first worth examining more common examples of file transfer protocols. HyperText Transfer Protocol (HTTP) and its more secure cousin HTTPS are two of many other file transfer protocols.

But there are some key differences between "client-server" protocols such as HTTP and HTTPS, and peer-to-peer protocols like BitTorrent.

The client-server approach

When a browser retrieves a web page or other resource from a web server, the page to retrieve is defined by a Uniform Resource Locator (URL). For example, one of my previous articles at the The Conversation has the following URL:

theconversation.com/how-the-he … nline-security-25536

In this URL, the "https" indicates the protocol and "theconversation.com" is the *host name*. The remaining part of the URL denotes a specific resource (file) on the host server.

When you access a URL, the web browser (i.e. the client) examines the *host name* (theconversation.com) and contacts a *name server* to find out the Internet Protocol (IP) address of the server responsible for hosting "theconversation.com".

It's just like looking up a person by name in a phone directory to get their phone number.

Once the browser knows the IP address of the server, it contacts the server and asks for the content as indicated by the rest of the URL. The server retrieves the content and sends it, in its entirety, to the client's IP address.

It's worth noting that the client also has an IP address, but that only has to remain stable for a relatively short period of time, and doesn't have to appear in any directory.

Most home ISPs only provide IP addresses to their customers on a temporary basis. Furthermore, that "visible" IP address is shared between all the devices connected to a home network. This might include a couple of PCs, a few tablets and smartphones or even internet-connected appliances, possibly owned by different people.

Client-server file transfer protocols work well for many purposes. Unfortunately, media files – particularly high-definition video for movies – can be very large. A high-quality full length movie runs to hundreds of megabytes of data that needs to be transferred to the client. Multiple simultaneous requests for them will overwhelm most standard internet servers.

Companies such as Netflix and YouTube therefore need large "server farms" with extremely fast and expensive network connections to meet peak demand.

Sharing the load

But there is an alternative approach. We don't need to ask the *original* server for the file – any intact copy will do. All we need is a mechanism for finding out which computers have a copy of the file we want and are willing to share it, at this particular moment, and what their IP addresses are so we can contact them and ask for a copy.

And that's precisely what early peer-to-peer file sharing mechanisms such as Napster and Gnutella did. Rather than one server providing the files, Napster and Gnutella had central servers that kept track of the IP addresses of computers (i.e. peers) currently offering particular files on a minute-by-minute basis, and a mechanism for requesting a file from another peer.

BitTorrent has an additional refinement. When your software makes a BitTorrent request, you get a list of the IP addresses of a *swarm* of peers who either have a complete copy of the file ("seeders", in BitTorrent's terminology), or are in the process of retrieving the file (non-seeder peers, or "leechers").

The software then requests "chunks" of the file from both seeders and leechers. Other leechers can request the parts you do have even before you have a complete copy.

Because of this cooperation, a very large number of computers can simultaneously get copies of very large files, without putting undue load on any one computer or network link.

Legal and not-so legal sharing of files

This has a number of very useful non-controversial applications. For instance, Facebook uses the BitTorrent protocol to transfer software updates to the thousands of servers it uses.

But it's undeniable that BitTorrent is also very useful for those who want to share copyrighted material. The only permanent infrastructure required is a server that has links to "torrents" – the originating seed which maintains a list of the computers in a swarm.

Not only is this not particularly costly, it maintains a level of indirection to the possibly copyright-infringing files being shared.

This has not stopped authorities – with the strong encouragement of the movie, television and music industries – using the law to attempt to shut down torrent directories for copyright-infringing material such as the Swedish-based The Pirate Bay.

It's worth noting that BitTorrent Inc itself is not associated with The Pirate Bay or any other copyright-infringing torrent directory. It is not a party in the present lawsuit about the alleged use of BitTorrent technology for copyright infringement.

Despite periodic shutdowns and arrest of The Pirate Bay's creators, it and other torrent directories remain available.

Representatives of copyright holders have resorted to another approach: suing BitTorrent users who have shared copyright-infringing files. To do so they must identify those users, both to contact them and to provide sufficient certainty that they will be held legally liable.

IP addresses revealed

Identifying the IP addresses of the members of a BitTorrent swarm is extremely simple. When a new client connects to the swarm, the IP addresses of the members of the swarm are transferred to the client, and existing clients are updated as new clients enter or leave.

Therefore, if an organisation wishes to identify those participating in trading a particular infringing file, they merely need to write a modified BitTorrent client that connects to the relevant swarm and records the list of participants.

University of Birmingham researchers have reported on the extent of such monitoring, which indicated that at the time of their study in 2012, participants in high-profile torrent swarms would be logged within three hours.

In the current court case, the recording of IP addresses was performed by a product called Maverik Monitor, written by the German firm Maverickeye. The court decision makes amply clear that Maverik uses the general approach outlined above. The judge was satisfied:

[…] that there is a real possibility that the IP addresses identified by Maverik Monitor were being utilised by end-users who were breaching copyright in the film by making it available for sharing on-line using BitTorrent participating in a torrent swarm […]

The judge therefore decided that this was sufficient reason to permit "discovery" and ordered that several Australian ISPs turn over their records.

Proving copyright infringement

The fact that the judge accepted the possibility that the IP addresses might be being used for infringing copyright, however, does not necessarily mean that the ISP account holders identified will automatically be held liable for copyright infringement.

The judge authorised handover of IP records for three purposes:

- seeking to identify end-users using BitTorrent to download the movie
- suing end-users for infringement
- negotiating with end-users regarding their liability for infringement.

But identifying the end-user responsible for BitTorrent use to a sufficient degree of certainty may prove challenging in many cases, to an extent not clearly articulated in the judge's decision.

For instance, home Wi-Fi networks are often left "open" (not requiring a password to access the network), allowing any device within range to use the network, including for BitTorrent. That range can often extend considerably beyond the boundaries of a person's property.

It's clearly going to be a challenge to identify all the actual people responsible for accessing illegal copies of the Dallas Buyers Club.

Evading the BitTorrent monitors

There are a number of technical measures that determined pirates can use to avoid BitTorrent IP monitoring, aside from taking advantage of open Wi-Fi networks.

Virtual Private Networks (VPNs) are one such measure. They provide an encrypted "tunnel" between an Australian computer and a proxy in a country with a less conducive legal framework for copyright infringement lawsuits.

Many VPN providers take payment by near-untraceable means such as pre-paid credit cards, or Bitcoin, and claim not to keep logs tying the visible IP address from their systems to the Australian IP address at the other end of the tunnel.

Like BitTorrent itself, VPN technology has many legitimate uses, not least in providing secure remote access to corporate and governmental networks for employees. As such, banning or restricting the technology would be costly and impractical.

This story is published courtesy of The Conversation *(under Creative Commons-Attribution/No derivatives).*

Source: The Conversation

Citation: BitTorrent and the digital fingerprints we leave behind (2015, April 10) retrieved 27 April 2024 from https://phys.org/news/2015-04-bittorrent-digital-fingerprints.html