BitTorrent ネットワークにおける
効率的な著作権侵害監視手法について

2015 年 4 月

株式会社クロスワープ

内容

改訂履歴

改訂日	改訂箇所	改訂内容
2015/4/2	2.3 トレントファイル検索サイトについて	トレントファイル検索サイトを修正
2015/4/2	3. P2P FINDER について	P2P FINDER の動作仕様説明についての記載ミスを修正

1. はじめに

BitTorrent ネットワーク上で通信を行うためのプロトコルは一般公開されており、μTorrent や BitComet といった BitTorrent ネットワークと互換性のあるソフトウェアが日本国内外を問わず幅広く使用されているが、BitTorrent ネットワークでは著作権者の許諾を得ることなく多くのファイルが無断で流通しており、その被害が顕在化している状況にあると考えられる。しかし、BitTorrent ネットワークそのものは非常に広範であり、これまで著作権侵害の実情を正確に把握することや実際に対策を講じることが困難な状況にあった。

そこで本レポートでは、P2P ネットワークの監視システムである P2PFINDER を用いる事によって、ユーザーがファイルを保持し、BitTorrent ネットワークにアップロードを行っているかどうかを特定する技術検証を行い、効率的な著作権侵害監視を実施する手法の提案を行う。

2. BitTorrent について

2.1. BitTorrent の概要

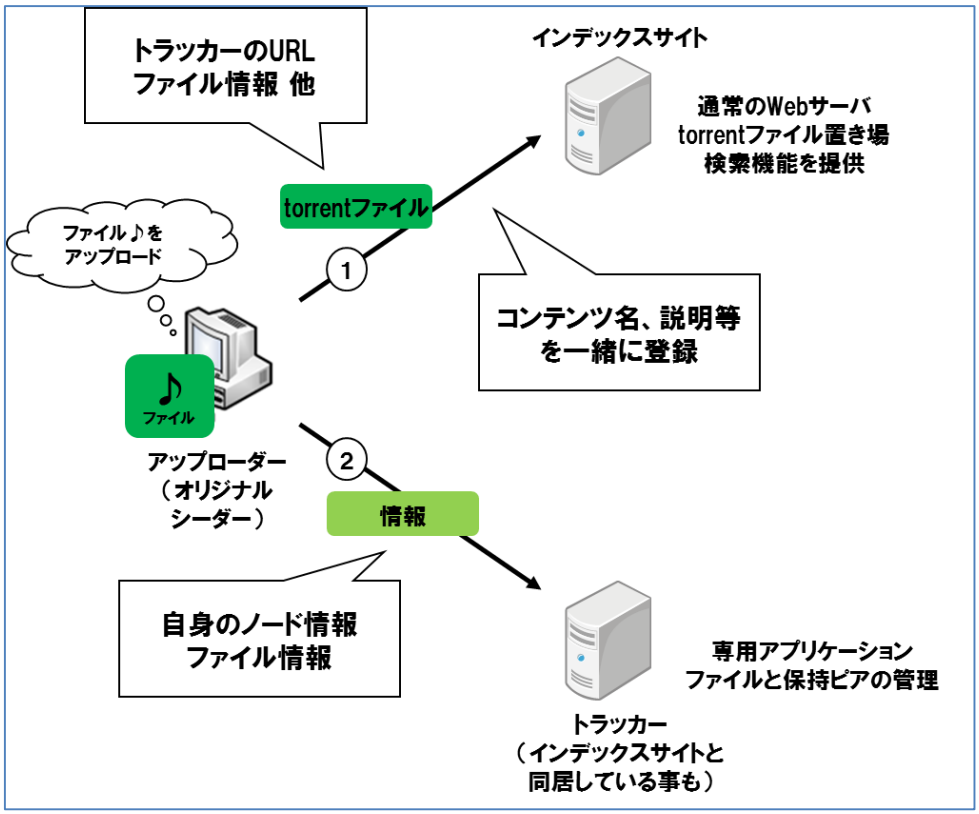
2001 年に開発されたハイブリッド型 P2P ネットワークでそのプロトコルが公開されているため、互換ソフトウェアが多い。互換ソフトウェアの例として μTorrent、BitComet、Vuze 等がある。これらの互換ソフトウェアは基本的にフリーソフトである。

BitTorrent はネットワーク上の検索機能を持たず、外部の WEB サイトに依存している。用途としてはファイル配信に特化しており、Linux の DVD イメージの配布等、合法配信にも利用されている。

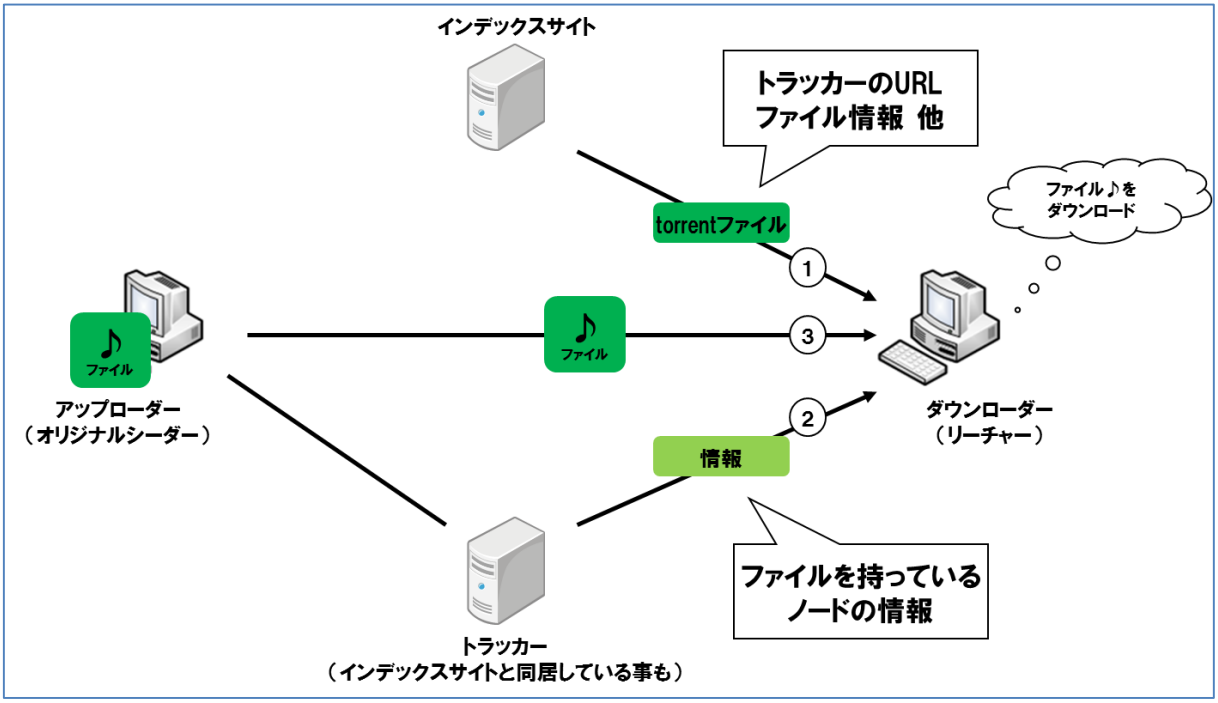
世界的に主流な P2P ネットワークの 1 つであり、日本国内でも利用者が増加傾向にある。

2.2. BitTorrent の初期ファイル流通について

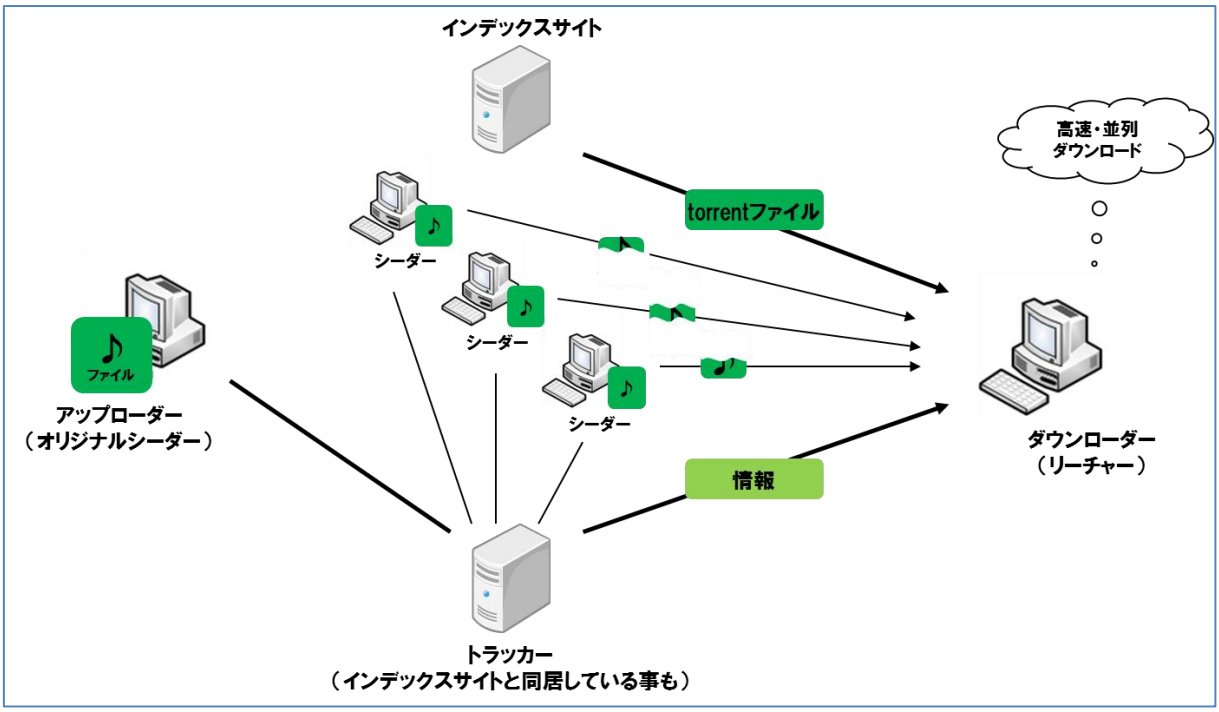
BitTorrent ネットワークはトラッカー、リーチャー、シーダーの 3 種のノードから構成されている。これらのノードに加えてトレントファイルを蔵置、配布するトレントファイル検索サイト(インデックスサイト)が多く存在する。その関係を以下に示す。



ファイルアップロード



ファイルダウンロード



並列ファイルダウンロード

- リーチャー：トレントファイルをトレントファイル検索サイト(インデックスサイト)等から取得し、実ファイルをダウンロードしようとするノード
- シーダー：実ファイルの全部を保持し、アップロードをしているノード
- ピア：リーチャーとシーダーの総称
- トラッカー：ピアからの登録を受け付け、他のピアの情報を提供する Web アプリケーションサーバ。トラッカーの情報はトレントファイル内に格納されている。次の例に示すようにトラッカーは URL の末尾に「announce」という文字列が付くことが多い。

  例）http://tracker.abcdef.net:3277/announce

2.3. トレントファイル検索サイトについて

多くのトレントファイルがアップロードされており、トレントファイルの検索、取得が可能になっている。以下に代表的なトレントファイル検索サイトを示す

サイト名	URL
The Pirate Bay	https://oldpiratebay.org/
Torrentz	http://torrentz.eu/
Tokyo Toshokan	http://tokyotosho.info/
nyaatorrents	http://www.nyaa.eu/

※2015 年 4 月現在

3. P2PFINDER について

BitTorrent 向けの P2PFINDER は「トレントファイル収集プログラム」「ノード接続プログラム」「ピースダウンロードプログラム」の 3 つからなり、それらを P2PFINDER の「管理システム」が管理する。それぞれについて以下に説明する。

3.1. トレントファイル収集プログラム

前述の「2.3 トレントファイル検索サイトについて」で記載された WEB サイト等からトレントファイルを集めるプログラム。一般的なブラウザによる WEB サイトの参照と同様に HTTP プロトコルを用いて WEB サイトを自動巡回し、トレントファイルをダウンロードする。以下に収集可能な主な項目を示す。

- トレントファイルを作成したソフトウェア名
- 実ファイルのハッシュ値
- 実ファイル名
- 実ファイルのサイズ(byte)
- 実ファイルのピース数

3.2. ノード接続プログラム

BitTorrent ネットワーク上のピアに接続して情報を収集し、テキストファイルに出力するコマンドラインプログラム。以下に収集可能な主な項目を示す。

- 接続日時
- ピアの IP アドレス
- ピアのポート番号
- ピアの保持しているピース数

3.3. ピースダウンロードプログラム

ピースダウンロードプログラムは、シーダーからファイルの断片(ピース)をダウンロードするプログラムである。ピースダウンロードプログラムはトレントファイルを元に Bittorrent ネットワークに参加してピアに接続する。その後、ピアからピースをダウンロードしつつ、そのピアがシーダーかリーチャーかを判別し、シーダーだった場合にのみ情報を管理システムに出力する。この際、ピアに直接接続することにより、ピアの IP アドレス、ポート番号、ダウンロード日時についてピースダウンロードプログラムは正確に把握することができる。以下に収集可能な主な項目を示す。

- ダウンロード日時
- ダウンロード先ピアの IP アドレス
- ダウンロード先ピアのポート番号

3.4. 管理システム

監視対象キーワード等の管理をした上で上述のトレントファイル収集プログラムとノード接続プログラムとダウンロードプログラムを連携稼働させ、収集可能な項目（キー情報）を収集、分析、保存、閲覧するシステム。

4. 検証手法

4.1. 手順概要

本技術検証は以下の手順で実施する。なお、利用する BitTorrent 互換ソフトウェアは 2012 年 4 月現在、最も有名な BitTorrent 互換ソフトウェアの一つである *µ*Torrent を用いる。

検証手順

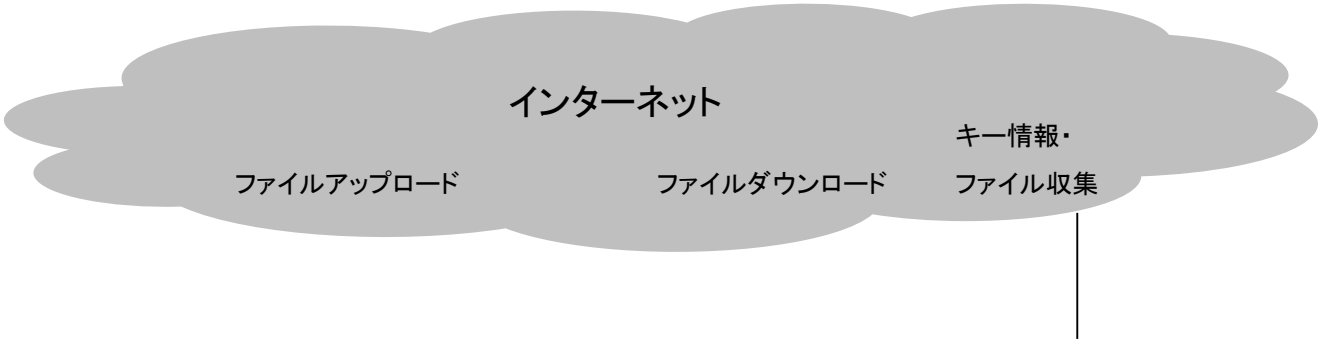
1. アップロード用 PC 端末上で *µ*Torrent を稼働させ、それぞれの端末でランダムデータを含み、ランダムな文字列で作られたファイル名を持つ 5 個のテストファイルをインデックスサイトに公開し、BitTorrent ネットワークにアップロード
2. P2PFINDER に 1 でアップロードしたテストファイルのファイル名を監視するように設定
3. P2PFINDER で BitTorrent ネットワーク上のキー情報収集および、ファイル断片(ピース)のダウンロードを実施
4. アップロード用 PC 端末とは別のダウンロード用 PC 端末から、アップロードされたテストファイルを BitTorrent ネットワークからダウンロード
5. 3 でのキー情報収集結果から、1 でのテストファイルのアップロードと 4 でのテストファイルのダウンロードがどのように検知されているかを確認
6. 3 での P2PFINDER によるダウンロードで取得したファイル断片(ピース)と 1 でアップロードしたテストファイルおよび 4 で取得したダウンロードファイルとのバイナリマッチを実施

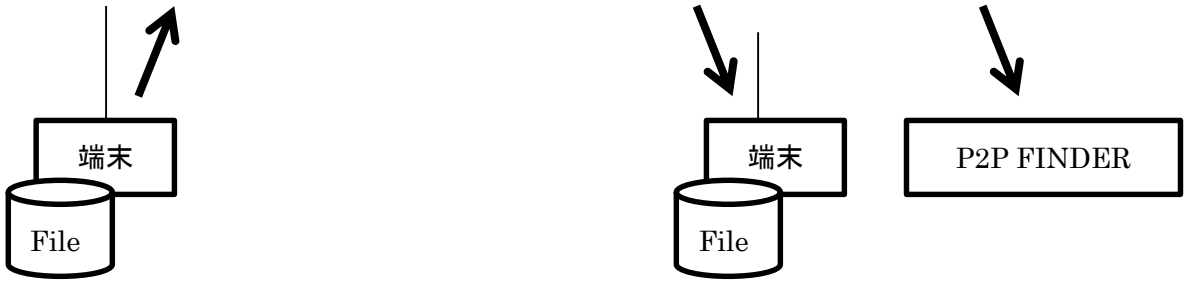
検証のポイント

- P2PFINDER で収集したキー情報を見て、テストファイルのアップロードノードをシーダーとして特定できるかどうか
- P2PFINDER によるファイル断片のダウンロード結果から、テストファイルのアップロードノードを特定できるかどうか
- テストファイルとファイル断片はバイナリー致するか（ランダムファイルなので偶然に部分一致することはありえない）

4.2. 機器構成

本技術検証は以下の構成で実施する。





BitTorrent を稼働させるアップロード用端末　　ダウンロード用端末　　監視システム

5. 検証結果

5.1. ファイルのアップロード

インターネットに接続している Windows の端末上で μTorrent(3.1.3.0)を稼働させ、以下の通り、5 つのテストファイルを BitTorrent ネットワークにアップロードした。なお、各テストファイルはランダムなバイナリデータを使用して作成した。

本技術検証の実施日時、使用した端末と各端末でアップロードしたファイル名、利用したトラッカーは以下の通りである。

アップロード期間	IP	ファイル名	利用したトラッカー
2012/6/7 17:00 から 2012/6/11 11:00	219.108.148.249	92ED65A7A0CEC040C31ACAD3322D69F01.zip	Tokyo Toshokan
	219.108.148.249	92ED65A7A0CEC040C31ACAD3322D69F02.zip	Tokyo Toshokan
	219.108.148.249	92ED65A7A0CEC040C31ACAD3322D69F03.zip	Tokyo Toshokan
	219.108.148.249	92ED65A7A0CEC040C31ACAD3322D69F04.zip	Tokyo Toshokan
	219.108.148.249	92ED65A7A0CEC040C31ACAD3322D69F05.zip	Tokyo Toshokan

5.2. ファイルのダウンロード

インターネットに接続している 1 台の Windows の端末上で μTorrent(3.1.3)を稼働させ、以下の通り、上述の 5 ファイルを BitTorrent ネットワークからダウンロードした。

ファイル名	ダウンロード期間
92ED65A7A0CEC040C31ACAD3322D69F01.zip	2012/6/8 19:15～2012/6/8 23:34
92ED65A7A0CEC040C31ACAD3322D69F02.zip	2012/6/8 19:51～2012/6/9 17:56
92ED65A7A0CEC040C31ACAD3322D69F03.zip	2012/6/8 19:15～2012/6/9 2:23
92ED65A7A0CEC040C31ACAD3322D69F04.zip	2012/6/8 19:16～2012/6/9 3:02
92ED65A7A0CEC040C31ACAD3322D69F05.zip	2012/6/8 19:16 ～2012/6/8 21:26

ダウンロードしたそれぞれのファイルとアップロードしたテストファイルについて MD5 及び SHA1 の 2 種類のファイルハッシュ値を比較したところ、両者は完全に一致した。

ファイル名	MD5 ハッシュ	SHA1 ハッシュ
92ED65A7A0CEC040C31ACAD3322D69F01.zip	1E412C06E950AAF9FA96304EBB63F4D7	7F6D82D3C07FD87941AB5E1E807711EEF05EAE63
92ED65A7A0CEC040C31ACAD3322D69F02.zip	7937A124CF57175C05FEA9EF8A927B37	0D40A86C778CC8132F232FADDF2C786F3D27586F
92ED65A7A0CEC040C31ACAD3322D69F03.zip	9255C5C62AB5B4BE097DFA20173B8029	D29F17C29DF5ED12752F6C197B73707413C6CC2C
92ED65A7A0CEC040C31ACAD3322D69F04.zip	7654B86568F9D4FA7DF3B44FE3C6BE46	30CD431A687C26A6DDD0C1804B5438122D799208
92ED65A7A0CEC040C31ACAD3322D69F05.zip	B263124E270977ABCD6A08E2AA69005C	61C011146439CBB47FA3950540686FA12CB46CA3

5.3. キー情報の収集結果

アップロードしたそれぞれのファイル名に共通して含まれるランダムな英数字を P2PFINDER システムの監視対象キーワードとして設定し、BitTorrent ネットワーク上からそのランダムな文字列がファイル名に含まれるファイルをアップロードするノード(シーダー)の情報を収集した。収集結果を以下に示す。

ファイル名	リーチャー 初検出時刻	リーチャー 初検出 IP	シーダー 初検出時刻	シーダー 初検出 IP
92ED65A7A0CEC040C31ACAD3322D69F01.zip	2012/6/7 18:06	1.163.137.66	2012/6/7 18:07	219.108.148.249
92ED65A7A0CEC040C31ACAD3322D69F02.zip	2012/6/7 17:45	220.147.77.3	2012/6/7 17:47	219.108.148.249
92ED65A7A0CEC040C31ACAD3322D69F03.zip	2012/6/7 18:46	202.53.210.230	2012/6/7 17:40	219.108.148.249
92ED65A7A0CEC040C31ACAD3322D69F04.zip	2012/6/7 18:44	202.53.210.230	2012/6/7 17:38	219.108.148.249
92ED65A7A0CEC040C31ACAD3322D69F05.zip	2012/6/7 17:39	220.102.245.226	2012/6/7 17:40	219.108.148.249

シーダーの初検出 IP アドレスは全て、アップロード用端末が利用した IP アドレス(219.108.148.249)であった。

5.4. 検知ノードからの直接ダウンロード結果

さらに収集したキー情報に基づく P2PFINDER によるダウンロードの結果から、アップロード用端末の IP アドレスからキー情報内のファイルを直接ダウンロードすることができているかどうかを確認した。

ファイル名	INFO HASH	アップロード用端末からの DL 回数
92ED65A7A0CEC040C31ACAD3322D69F01.zip	3487cc8dd59a8e0ee7b6fcc60f74f1d78da2dbb4	114
92ED65A7A0CEC040C31ACAD3322D69F02.zip	ebb9381f5f17703fb7c4d1cc330029cea7de85cd	131
92ED65A7A0CEC040C31ACAD3322D69F03.zip	e731bdb3dcef1ad55d79e9882fd1434750f1ace7	130
92ED65A7A0CEC040C31ACAD3322D69F04.zip	798bf379aa5c6439f3ec8326d6b77916f163b27c	121
92ED65A7A0CEC040C31ACAD3322D69F05.zip	a249d7d1c4713a762225ec135b37e0274a462332	127

アップロード期間 2012/6/7 17:00 から 2012/6/11 11:00 までにファイルのピースを総計 623 回に渡ってアップロード用端末の IP アドレスからダウンロードした。

この 623 回のダウンロードで収集したそれぞれのピースファイルについて、全てがアップロードしたテストファイルの一部と完全に一致した。

6. 本技術検証から得られた知見

収集したキー情報からテストファイルのアップロードノードは最初にシーダーとして検出され、その後、複数回シーダーとして検出された。今回シーダーとして最初に検出されたことは重要ではなく、アップロード用端末の IP アドレスが、確実にシーダーとして検出されたことが重要である。この事により、P2PFINDER が間違ったシーダーの IP アドレスを検出していないことが裏付けられる。また、テストファイルのアップロードノードから P2PFINDER を用いて直接ダウンロードすることも可能であった。

アップロードしたテストファイルと P2P FINDER でダウンロードされたピースファイルはバイナリレベルで一致した。

7. 監視手法の提案

本技術検証から得られた知見より、実際の監視手法として以下の方法を提案する。

1. 特定のトレントファイルを対象に BitTorrent 互換ソフトウェアを用いて実ファイルをダウンロードし、内容が著作権侵害物であることを確認する

2. 1. で内容が確認できたトレントファイルをもとに P2PFINDER で BitTorrent ネットワークをクローリングし、ファイル全体をアップロードしている可能性があるノード（シーダー）を絞り込む

3. 2. で得られたノード情報を元に P2PFINDER でファイル断片(ピース)のダウンロードを試行し、成功した場合には1. の結果得られたファイルとバイナリマッチを実施の上、著作権侵害ファイルアップロードノードとして確定する