Multimed Tools Appl (2015) 74:271–286 DOI 10.1007/s11042-013-1760-x

Methodology and implementation for tracking the file sharers using BitTorrent

Sooyoung Park • Hyunji Chung • Changhoon Lee • Sangjin Lee • Kyungho Lee

Published online: 10 December 2013 © The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Sharing copyright protected content without the copyright holder's permission is illegal in many countries. Regardless, the number of illegal file sharing using BitTorrent continues to grow and most of file sharers and downloader are unconcerned legal action to transfer copywrite-protected files. However, it is difficult to gather enough probative evidence to prosecute illegal file sharers in criminal court and/or sued for damages in civil court. Further, there is a lack of research on investigation techniques to reveal illegal BitTorrent sharers. This is because the role of the server in BitTorrent networks has been changed compared to servers in conventional P2P networks. As a result, it is difficult to apply previous investigation processes for investigation of conventional P2P networks to the investigation of suspected illegal file sharers using BitTorrent. This paper proposes a methodology for the investigation of illegal file sharers using BitTorrent networks through the use of a P2P digital investigation process.

Keywords BitTorrent investigation \cdot Illegal file sharing \cdot Digital investigation process \cdot Packet analysis \cdot Client-side disk forensics

1 Introduction

Intellectual Property (IP) is a legal concept that refers to creations of the mind for which exclusive rights are recognized [6]. Therefore, sharing copyrighted material without a

S. Park · H. Chung · S. Lee · K. Lee (🖂)

Center for Imformation Security Technologies(CIST), Korea University, Anam Campus, Anam-ding 5-ga, Seongbuk-gu, Seoul, Korea e-mail: kevinlee@korea.ac.kr

S. Park e-mail: sooyoung011@korea.ac.kr

H. Chung e-mail: foryou7187@korea.ac.kr

S. Lee e-mail: sangjin@korea.ac.kr

C. Lee

Department of Computer Science and Engineering, Seoul National University of Science and Technology(SeoulTech), Seoul, Korea e-mail: chlee@seoultech.ac.kr

copyright holder's permission is considered illegal in many jurisdictions. The act of uploading copyrighted work without a copyright holder's permission infringes the right of reproduction and transmission of the copyright holder. The act of downloading copyrighted work without copyright holder's permission infringes on their right to reproduction. There are various laws that have been imposed to punish illegal file sharers in many countries. However, despite implemented law, illegal file sharing through P2P networks persists.

There are lots of technologies that enabling efficient data sharing in different environment. Liu suggested a method for adjusting traffic in Wireless Mesh Networks [7]. And, Kun suggested a method for designing a secure and efficient mix network [11]. P2P is one of technology for efficient data sharing. P2P allows users to download media files such as music, movies, and games using a P2P software client that searches for other computers connected to the same logical network [10]. There are many kinds of P2P network protocols, such as Opennap, Gnutella, Grokster, Freenet, eDonkey2000, and BitTorrent. Although P2P is a useful technology that makes data transmission fast and reliable, it has also been abused for copyrighted file sharing purposes. There are many attempts to make the data sharing process more efficient. Kim [5] suggested a methodology to enable a fault-tolerant scheme for stable streaming services over P2P networks. However, currently BitTorrent is the most widely used peer-to-peer file sharing protocol [1].

Conventional P2P software consists of a central server-based model. Where a central server manages and links each client by indexing all of the current users and searching their computers for shared data. This mechanism allowed communication and sharing of files between clients through the server. BitTorrent, however, is different from conventional P2P networks. BitTorrent links each client without a central server, and shares fixed size data pieces of a shared file to make sharing fast and fault tolerant. A client on a BitTorrent network can connect to multiple peers without a central server, and download small pieces of a file from multiple peers at the same time. The use of multiple peers allows the data to potentially be downloaded faster than if connected to just a single peer. Further, a client can download data from any peer that has the requested piece, even if that peer has not received the entire file.

In case of copyrighted material sharing using conventional P2P networks, U.S. cours judged that P2P website operators can be held legally responsible for data their users share. Many P2P sharing sites were shut down, and order to stop offering file-sharing services [8]. As mentioned before, BitTorrent sites provide different types of services compared to conventional P2P networks. Since a BitTorrent service is not based on central server, a BitTorrent client can directly connect to other clients and share files using a torrent file or magnet link. A torrent file is a meta file that includes the shared file's name, cryptographic hash value, length of each piece, etc. File sharing is made possible by opening a torrent file with a BitTorrent client program. A magnet link is a web URL that enables file sharing without a torrent file. Even if the site where torrent or magnet links is taken down, this countermeasure is not enough to stop copyrighted file sharing. We submit that to eradicate copyrighted file sharing, it is important to determine who has distributed the copyrighted material. Until now, it is difficult to identify and copyrighted material distributors.

File sharers in conventional P2P networks consist of seeders (uploaders) and leechers (downloaders). However, file sharers on BitTorrent networks cannot be classified specifically as seeders or leechers. This is because both leeching and seeding among peers occurs at the same time. Therefore, identification of a peer's behavior in the file sharing process is essential for identification of illegal file sharers.

This paper targets one of the most popular BitTorrent client programs, uTorrent. We present a methodology for identification of copyrighted file sharers through analysis of the file sharing process. In addition, a local analysis is conducted of trace artifacts resulting from file sharing in a client's computer. Further, a tool – Identifier and Classifier of Illegal sharers using BitTorrent (ICIB) – to demonstrate the practicality of the suggested investigation methodology.

Section 2 introduces related works about BitTorrent investigations. In Section 3, we discuss the investigation framework for file sharing using BitTorrent. Next, a detailed methodology is given for tracking file sharers by identifying and classifying sharers, logging the information about sharing and analyzing artifacts remaining in a client's computer. Finally, the prototype ICIB tool is described for detecting file sharers using BitTorrent.

2 Related works

Lewen [9] analyzed the legal basis for punishing copyrighted material sharers on P2P networks in the context of the U.S. and Sweden. However, it is difficult to apply legal basis to copyrighted material sharers using BitTorrent because on BitTorrent networks, it is more difficult to track and categorize user activities than on conventional P2P networks.

Schrader [12] suggested a technique for detecting copyrighted material sharing traffic through monitoring of the network, and checking the 'info_hash' value in the BitTorrent packet. Hatahent [4] suggested detecting copyrighted material sharing traffic by analyzing the clients' behavioral patterns. But, monitoring every network is impossible. Further, non-BitTorrent protocols such as TCP, UDP, etc. are used in the process of file sharing through BitTorrent. But the methodology described in these works just focus on the BitTorrent protocol. In addition, they didn't describe how to investigate sharers. Therefore, this paper proposes an investigation process for copyrighted material sharing using BitTorrent.

3 Investigation of copyrighted material sharing

If a copyright holder or investigative agency detects sharing of copyrighted or contraband files and acquires the torrent file, then an investigation can be performed based on the process model given in Fig. 1.

When the torrent file is executed with a BitTorrent client program, information about peers sharing the file is automatically obtained from a tracker or Distributed Hash Table (DHT). Information will be classified and recorded on a case-by-case basis.

A tracker is a server that provides the information about sharers that compose the filesharing network. DHT is a class of decentralized distributed systems that provides a lookup service similar to a hash table [2]. Using this structure, we can obtain and classify the information about peers by file sharing participation type.

Even though an investigative agency may acquire a peer's IP address that appears to be sharing copyrighted or illicit material, there is always possibility of IP has been forged or otherwise incorrect. Because of this, it is difficult to confirm the client with only with IP address and transmitted packet. Therefore, additional data must be acquired. An investigative may need to attempt to get a warrant (if the suspect is within the agency's jurisdiction) to seize and investigation the suspect's computer.

If the ISP provides the identities of the suspect based on the IP address, next the investigator must request warrants for seizing each suspect's computer (or other digital device). After warrants are issued, the investigator could seize the suspect's computer, and examine whether the suspect was sharing copyrighted material by gathering and analyzing data from suspect's computer.

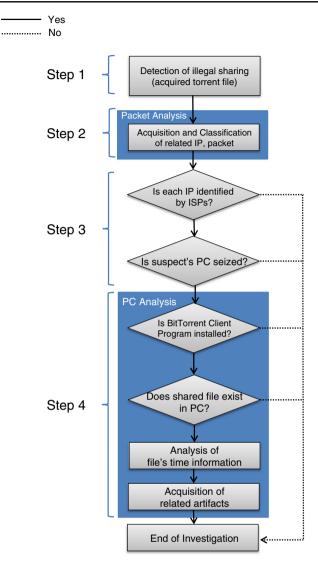


Fig. 1 Process model for the investigation of copyrighted material sharing on BitTorrent Networks

When analyzing a suspect's computer, the investigator must check for BitTorrent client software, the existence of the shared file and BitTorrent artifacts, for evidence that the suspect intentionally shared the file.

4 Investigating file sharers

In this section, a detailed methodology is given for tracking BitTorrent network file sharers that follow the previously described investigation process. This section is focused on Step 2 and Step 4 of the investigation process.

4.1 Acquiring a file sharer's IP

When a copyright holder or investigative agency detects sharing of copyrighted or contraband files, and acquires the torrent file or magnet link, they can gather copyrighted material sharer's IP and related network traffic information. Gathering a file sharer's network traffic information can be done based on the structure of the BitTorrent network and sharing process.

Figure 2 shows the process of file sharing using BitTorrent, and the structure of the sharing network. In this network peers (network clients requesting and sharing the same data) can get information about other peers that are sharing data by sending a request to the tracker. After gathering information about other peers, sharing of data can be started by requesting data from identified peers, directly. Requested data, as well as information about additional peers are shared between connected peers. This is possible through DHT. And allows each peer to gather information about other peers through the tracker or DHT. This means that to gain comprehensive information about the network, information should be gathered about peers through the tracker and DHT.

There are two methods for retrieving information about peers. The first method acquires peer information from HTTP packets that are received from the tracker as the response to HTTP GET requests that include 'info_hash'. 'info_hash' consists of a shared file's meta information. These packets have a structure as shown below in Fig. 3.

If the key peers exists in the encoded area of the HTTP packet, it is possible to get the value corresponding to the key 'peers'. This means that acquiring information about peers consisting the sharing network is possible.

The second method uses DHT, which is an embedded function in BitTorrent for exchanging information about peers. A UDP packet is sent as the response to a request message 'get_peer', and includes information about peers and nodes. The structure of these packs is shown in Fig. 4.

If the key nodes or values exist in the encoded area of the UDP packet, it is possible to get values corresponding to the key nodes or values. There is information about the nodes using DHT in the value of nodes. Further, there is information about peers in the sharing network in

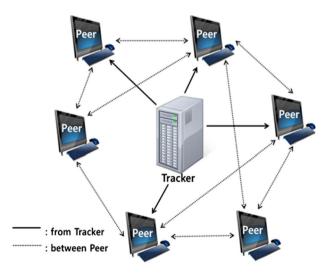


Fig. 2 Structure of a BitTorrent network for data sharing with peers on the network who receive torrent information from a centralized tracker

d •		lod	0a 64	38	3a	69 6	e 74	65	72	76	61	60	69	31	35	-	-	d8:1n
•				35 d5			5 65			32 bb		30	3a				00	00eS pe
•		be	35 67 43 b1	aa	fd	0d 4	d a3 d 73	4d	15	76 5a	fe	bd	18	01	e0)		. 5gM
peers[bytes] :		a6	75 63	3d	44	0c 2 61 4	6 9f	6e	c2	c6 c7	97	82	61	d3	77			.c.).E. .uc=D.& ^GaJ
IP(4byte)	Ports(2byte)	3c	52 ca	5d	d0	9f 3	b 8b	ad	32	5a	3d	db	25	63	2f		<f< td=""><td><r.];< td=""></r.];<></td></f<>	<r.];< td=""></r.];<>
IP(4byte)	Ports(2byte)	a3	d4 66 98 7f	12	69	15 4	7 a3	b6	89	55 ba	c5	af	6e	72	8e			f+. i.G
:		09	40 ba 0b 6e 79 17	63	af	8f b		f3	65	00 60 7b	31	26	ce	a4	ed	ĺ.		.@ 3. nc Dy.K.).
•		3c 09	af 62 0a 37	cf bb	7c bb	28 e 65 0	e a0 7 93	89 b1	ba 01	3b ba	0b 16	42 3a	c1 2f	da da	ba 71	1	<.	<.b. (.
•		3d	41 60 9e ad cf f4	c3	0a	dd 4	2 f2	7c	b9	fa c1 4c	26	ec	32	c9	e5		=.	.A`KC =B
•	e		54 79							d8			20	re	//			^

Data(Bencoded)

Fig. 3 Response packet for HTTP GET request

value of values. In some cases, there are packets that only include node information or peer information, or both. The following suggested algorithm can acquire and save information for all cases.

4.2 Classifying file sharers by action

The communication setup and data transfer process between peers is shown in Fig. 5. First, a peer identifies other peers that are able to connect to a session by attempting to send a TCP SYN packet. If a TCP session is connected, the peer checks other peers to determine if they are sharing requested data. Data checking is done during a BitTorrent handshake. After this process, peers inform requestors about data pieces they have. A sharing peer will send the message 'Have, Piece' that includes the index of each available piece. If a peer requests pieces that are shared by another peer, the peer requests pieces by sending a 'Request, Piece' message that includes the index of the requested piece. The peer that received the request will send data using TCP. A series of process can be divided into 3 sections as shown in Fig. 5. And, peers can be classified based on the section in which they are included.

Data(Bencoded)						
d .						
- ·						
•						
nodes[bytes] :						
IP(4byte)	Ports(2byte)					
IP(4byte)	Ports(2byte)					
1 - C						
•						
•						
values[bytes] :						
IP(4byte)	Ports(2byte)					
IP(4byte)	Ports(2byte)					
1						
•						
	e					

$ \begin{array}{c} 3a \ 69 \ 64 \ 32 \ 30 \ 3a \ ab \ 30 \ ap \ f3 \ 94 \ c7 \ d4 \ fb \ d5 \ 6e \\ (100 \ cm) \ c$							_										
0b db e6 90 33 53 76 2b 2a 35 3a 6e 6f 6f 55	7f																i7dl:rd2
73 32 30 38 ab 33 f7 d5 d4 1b 10 6b be 12 37 5208: 3. M.A. 73 22 30 38 ab 33 f7 d5 d4 1b 10 6b be 12 37 5208: S208: Ym7U 33 f0 bb 34 40 da 47 d4 c9 11 d7 18 d5 fe ce 65 d5 d5 d5 d4 d4 d4 c9 11 d7 18 d5 fe ce d6 d5 d6 d2 d5 d5 d4 d4 d6 d7 d6 d7 d6 d6 d7 d6 d6 d7 d8 d3 d6 d1																	:id20:.0n
c7 f4 bd bf b b) c8 c8 b) c8 c9 b) c8 c9 b) c6 c6 c6 c6 c6 c6 c6 c7 c2 b) c8 c9 b) c8 c9 c6 c7 c7 b) c6 c7 c7 b) c6 c7 c7 c7 b) c6 c6 c7 c7 c7 b) c6 c7 c7 c7 b) c6 c7 c7 c7 c7 c7 c8 c7																	
33 f0 bb 34 40 da 47 d4 c9 11 d7 18 d5 7e ce 65 3.44.6,																	\$208:.3Mk7
c7 f2 b8 6d f2 f8 eb f5 9b b3 02 20 3f 3p cd																	Ym7U
a4 91 af 75 59 64 e2 86 fb 40 fb b8 02 43 af 03																	34@.G~.e
64 a9 e1 ab 33 1d 5a b5 72 ec f6 b7 5a b4 67 c9 d3.z. rz 74 fc bf ec a7 82 56 55 51 63 24 1a e1 ab 33 5d tvu QcS 90 5f c9 d7 06 87 ce 2b f2 9b 9d 8f 2f 80 3f 56 + + c. 70 98 f0 71 85 c8 4e e4 d0 b0 ad 3f 97 37 a9 + c7 09 8f 071 85 c8 4e e4 d0 b0 ad 3f 97 37 a9 + 73 16 cf fd 06 c9 03 fd 14 f7 57 ab 32 60 68 c0 s 10 03 b0 82 80 d2 85 fe 85 0d 04 68 af d1 1f 5b 0																	mE3. ?9.
74 fc bf ec a7 82 56 51 63 24 1a e1 ab 33 5d tvu QcS 90 5f c9 d7 06 87 ce 2b f2 9b 9d 8f 2f 80 3f 56 vu QcS 90 5f c9 d7 06 87 ce 2b f2 9b 9d 8f 2f 80 3f 56																	uYd@C
90 5f c9 d7 d8 f2 g9 g4 f2 f8 g4 f8 g4 f8 g4 f8 g4 f8 g4 f8 g4 g4 g6 g4 g4 g4 g6 g4 g4 g4 g6 g4 g4 g4 g6 g4 g4 <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>d3.z. rz.g.</td></td<>																	d3.z. rz.g.
b0 b0 b0 b2 b2 c7 ab 32 8a 0e ed d2 b2 d2 ab 32 8a 0e ed d2 b2 d2 d2 ab 32 8a 0e ed d2 d2<																	tVU Qc\$3]
ec 70 98 f0 71 85 c8 e 40 b0 add 3 97 37 a9 .p.q.N 8f ab 32 17 5f cd 02 60 58 fd 74 a9 .p.q.N 73 16 cf fd 06 58 fd fd 8a 21 1a .2&.X. 73 16 cf fd 06 58 fd fd fd fd fd a .2 & X.X. 30 03 b0 82 80 d2 85 fe 85 0d 04 68 af d1 1f 50																	·_···/.?v
8f ab 32 17 5f ec d0 26 0b 58 fd 7d 8a a2 a1 1a 2&.x.}. 73 16 cf fd 06 c9 03 fd 14 f7 57 ab 32 60 68 c0 s																	" 2n.
73 16 cf fd 14 f7 57 ab 32 60 68 c0 s																	.pqN?.7.
30 03 b0 82 80 d2 85 fe 85 0d 04 68 af d1 1f 5b 0h. b6 f8 79 c8 d5 35 3a 74 6f 6b 65 6e 38 3a 96 c1y.5: to ken8 e2 f9 f3 01 c9 1b 36 3a 76 61 6c 75 65 73 6c 366: value 3a be a0 02 63 41 f4 65 65 31 3a 74 34 3a 8b 08e4 e1:t4																	2& .X.}
b6 f8 79 c8 d5 35 3a 74 6f 6b 65 6e 38 3a 96 c1																	s
e2 f9 f3 01 c9 1b 36 3a 76 61 6c 75 65 73 6c 366: value: 3a be a0 02 63 41 f4 65 65 31 3a 74 34 3a 8b 08cA.e e1:t4																	0h[
3a be a0 02 63 41 f4 65 65 31 3a 74 34 3a 8b 08 :cA.e e1:t4																	y5:t oken8:
																	6: values16
										31	Зa	74	34	3a	8b	08	:cA.e e1:t4:
14 4/ 51 58 /9 51 58 /2 05 .GI;y1:F e	14	47	31	3a	79	31	3a	72	65								.G1:y1:r e

Fig. 4 Response packet for message 'get_peer'

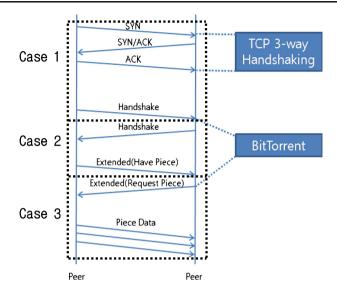


Fig. 5 TCP handshake, BitTorrent handshake and data transfer between peers on a BitTorrent network

4.2.1 First case : only a TCP session is connected

A BitTorrent client program tries connecting a session by sending a TCP SYN packet to peers that exist in the peer list. In this case, there are two possible to attempt to connect: either there are no connected sessions or there is only a TCP session. In the first case, there is either a TCP RST packet or no acknowledgement to the TCP SYN packet. In the second case, there is a connected TCP session through a TCP 3-way handshake. In both cases, peers are regarded as the same in the suggested algorithm, because they are just different from activation or non-activation of opposite peers in network at the time IP information is acquired.

The peers that belong to this case can be regarded as a suspect that had the intention of sharing the file. The reason why these peers remain in the peer list is to go through a pre-step for file sharing. It is unable to gather packets that are related to file sharing in the past using the suggested algorithm. Thus, proof of liability for this case is more difficult than other cases.

The investigator can save the information about a peer's IP and port number from the peer list.

4.2.2 Second case : exchange of BitTorrent handshake packet has occurred

A BitTorrent handshake packet includes a 'info_hash' value that consist of meta- information of the shared file. Sending this packet can be regarded as an announcement before actual file sharing, from which it can be inferred that each peer has the intention to share the file. Thus, a BitTorrent handshake packet is always exchanged before sending the file's raw data. So, just exchanging this packet assumes the will to of the peer to share the file.

After the BitTorrent handshake, peers generally send the message 'Have, Piece' to each other. The message 'Have, Piece' means "I can send the pieces you are requesting" by examining the piece list they have. Thus, peers can request pieces from other peers that send the message 'Have, Piece' at any time by sending the message 'Request, Piece'. If there is no transmission of the message 'Request, Piece' after the transmission of the message 'Have,

Piece', it means the file's data has not been transmitted. No transmission of the message 'Request, Piece' means that the requested peer does not have the pieces that are being requested. In this case the investigator can regard peers as violators that have intention to share a file, even though there was no transmission of the file's data because these peers completed the preparation for file sharing whenever peers request pieces.

An investigator can save information about the IP and port number of each peer from the peer list, BitTorrent handshake packet and information about what pieces are possessed by each peer from the packet that includes the message 'Have, Piece'.

4.2.3 Third case : file sharing actually occurred

In this case, the transmission of a file's data occurred after a peer requested the piece from another peer by sending the message 'Request, Piece' after all pre-steps for file sharing are accomplished. An investigator can save all packets that are transmitted to each peer to map the sharing environment, and exchange a file's data.

4.3 Algorithm for identification and classification of file sharers

If a copyright holder or investigator detects illegal sharing and acquires a torrent file or magnet link for illegal sharing, they can gather the sharers' IP address and network packets by using the algorithm in Fig. 6.

The given algorithm identifies whether each packet is related to illegal sharing. If packets are related to illegal sharing, suspect information will be acquired and saved to a database or log file. The algorithm can be divided into two parts: the first part is the acquisition of the IP address and network packets related with sharing, and the second part is classifying the IP address and network packets based on the roles of each peer sharing data.

If there are new packets, classification of packets can be conducted. If there is an HTTP response packet that includes information about peers as a format of encoding, this packet can

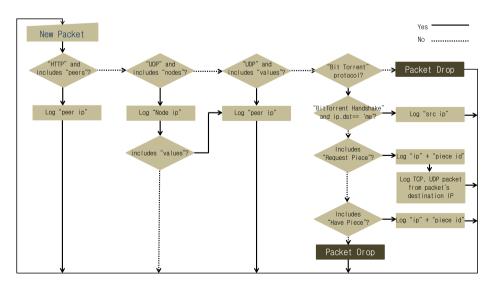


Fig. 6 Algorithm for acquisition and classification of data related with sharers

be regarded as a tracker's response. We can extract a peer's IP address and port information from this type of packet. If there is a UDP packet that includes information about peers or nodes as a format of encoding, this packet can be regarded as a response through DHT. We can extract peer and node IP addresses and port information from this type of packet.

If there is a BitTorrent packet, this packet can be seen to have a relationship to the session for the transmission of a file. We can classify theis type of packet more specifically, and log additional data. At first, if there is a BitTorrent handshake packet received from another peer, this situation belongs to 'Case 2' (Exchange of BitTorrent Handshake packet has occurred). The act of sending a BitTorrent handshake packet indicates the will to share data. In this case, we can log information of peers that sent a BitTorrent handshake packet. Secondly, if there is packet for requesting pieces that was received from another peer, this situation belongs to 'Case 3' (File sharing actually occurred). In this case, we can log information about peers that sent these packets, requested the piece ID, and all packets that include the file's data. Third, if a 'Have, Piece' packet is sent, we can log information about the peer that sent these packets, and the piece ID that is owned by that peer.

4.4 Analysis of a file sharers' PC

4.4.1 Checking whether BitTorrent client program is installed

To prove the fact that file sharing using BitTorrent has occurred on the suspect's PC, the investigator must check whether the BitTorrent client program was installed on the suspect's PC. Even though a BitTorrent client program was not installed on suspect's computer, the investigator must consider the situation that the suspect deleted the program, which can be proved by checking the Windows Registry key:

HKLM\SOFTWARE\Microsoft\Windows\CurrentVersion\Uninstall

An investigator must check whether the BitTorrent client program is included in that Windows Registry key.

4.4.2 Checking for the existence of a shared file

To prove that illegal sharing using BitTorrent has occurred on a suspect's PC, the investigator must check whether files, which were uploaded or downloaded through illegal sharing, exist on the system. Carving the file system may be needed to consider deleted data.

Additionally, it is possible that there are incomplete files in the PC. If the suspect temporarily participated in file sharing, it is possible that there are incomplete files composed of some downloaded pieces on suspect's PC. When upload or download of pieces occur with an incomplete file, laws are different in each country about whether this case should be regarded as illegal or not. But the most obvious thing is that the suspect can transmit pieces to any other peers, and help other peers to make up the complete file even though the suspect has an incomplete file. In this case, a suspect has a responsibility for the right of transmission and reproduction.

4.4.3 Analysis of a file's timestamp

If a shared file exists on a suspect's PC, the created time of the file must be considered. It is possible that a suspect obtained the file through another way, regardless of how the suspect tried sharing the file through BitTorrent at first. In this case, the investigator must remove the case from this investigation process, and an alternative investigation procedure should be started. Analysis of timestamps should determine the time when file sharing occurred. In this case, the created timestamps of shared files are after the created timestamp of the torrent file.

If a suspect obtained the file through a web site, web history might remain in logs of the web browser. The created time of web history must be between the created time of torrent file and the shared file.

4.4.4 Acquisition of BitTorrent artifacts

It is important to find additional evidence related to illegal sharing on a suspect's PC. An investigator can identify the torrent file's name that was executed recently by checking the Windows Registry in following path [3]:

 $HKEY_USERS \ SID \ Software \ Microsoft \ Windows \ Current \ Version \ Explorer \ Recent \ Docs \ torrent$

In case of uTorrent, there are torrent files the suspect had executed in following directory:

If a suspect engaged in illegal sharing using a torrent file, the torrent file may still remain in the directory mentioned above even though the suspect deleted the torrent file. There are also log files contained in this directory (Tables 1 and 2), not only torrent files.

In addition, uTorrent offers statistic information about usage. Important items among the available information are: the amount of data transmission in the last 31 days, the size of total uploaded data, the size of total downloaded data, the total run time of uTorrent, the number of torrent files that were executed, the number of uTorrent executions and the last executed time, etc.

5 Implementation

Identifier and Classifier of Illegal sharer using BitTorrent (ICIB) is a tool that identifies and classifies information about P2P file sharers using the suggested methodology. ICIB conducts analysis of P2P clients using a target torrent file as input. If a user inputs a torrent file, and clicks 'Acquire Start', the packets related to file sharing using BitTorrent would be acquired. Packets can be acquired until the user clicks 'Acquire Stop and Analysis'. Figure 7 shows the first page of ICIB, related to acquisition and analysis of related packets.

After the acquiring process is finished, analysis of the acquired packet can be conducted. Results of the analysis are shown in Fig. 8. Based on the participation degree of file sharing, sharers are classified into three groups. The list of classified IP addresses is shown in Area 1. An investigator can see detailed information about an IP address by selecting the item in the sharer list. Detailed information about a selected IP address is shown in Area 2 and Area 3. In Area 2, the whole list of packets that relate to the selected IP is shown. In Area 3, information that directly

Table 1 Directory related with uTorrent	OS version	Path
	Windows XP Windows 7	%Userprofile%\Application Data\uTorrent %Userprofile%\AppData\Roaming\uTorrent

 Table 2
 Log files in directory

Name	Detail
dht.dat	Node information organizing DHT
Settings.dat	Setting information of uTorrent and tracker information

related to transmission of file is shown. Information shown in 'Area 3' is divided into two types of information. In the 'Piece Info' tab, information about pieces that are owned by the selected IP address is shown, such as related packet number, time information, size of data that include the 'Have, Piece' message and piece ID. In the 'Shared Data Info' tab, information related to transmission of a file such as related packet number, time information and size of real data that was transmitted.

With ICIB we tested whether a file sharer's information can be loaded successfully. If some IP addresses belong to Case 1, just the packet list that is related to that IP address is shown. Figure 8 shows that the packet list of the selected IP address belongs to Case 1.

If some IP address belongs to Case 2, the packet list that is related to that IP address and the information about which pieces they have is shown. Figure 9 shows the owned data peer list that belongs to Case2.

If some IP address belongs to Case 3, a packet list that related to that IP, the information about which pieces they have, and the information that directly related to the transmission of file is shown. Figure 10 shows that the owned peer list of a selected IP belongs to Case 3. Figure 11 shows information that directly related to the transmission of the file.

6 Conclusion

Because BitTorrent is different from conventional P2P, applying a conventional investigation methodology in BitTorrent investigations is unsuitable. Furthermore, because

ICIB	×							
Select Torrent File								
C:\Users\S Y\Downloads\Sampletest.torrent								
Acquire start Acquire stop and Analysis								
======================================								
[2013-01-23 14:23:04] Acquire start [2013-01-23 14:31:25] Acquire finish								
[2013-01-23 14:31:25] Analysis start								

Fig. 7 First page of ICIB

		ICIB			
P . 12	5.177.30.196 × 59.7.201	153 141 223 238 89			
Packet No.	Time(sec) Source	Destination	Protocol	Length	Info
1404	33 163.152.1			72	SVN
1406	33 163.152.1			66	SYN/ACK
1437	83 125.177.8			66	ACK
10 126 102	33 163.152.1			54	
132.20 1439	33 125.177.3			72	
1440	33 163.152.1 33 163.152.1			122	Handshake
101.27	33 103.152.1 34 125.177.3			130	FIN/ACK
	34 125.177.3			60	ACK
227.115 1447 1381 1448	34 163.152.1			54	ACK.
114.91 1449	34 163 152 1			62	FIN/ACK
1.167 1450	34 163 152 1			54	FIN/ACK
3.140					
127 Piece Info H2+++++ Piece Info	Time(sec)	Message Size(bytes) IDX			
1.53	(interpret)	intestingle state(opties)			
186 73.139					
103					
105					
5					
2					
89					
19					
19					
88					
117					
227					
2.56					
45.111					
	Shared Data Info.				

Fig. 8 View for result of analysis in ICIB

uploading and downloading occurs simultaneously when exchanging pieces of data with other peers, it is difficult to punish illegal sharers under current law. Because of these reasons, the number of illegal sharing is still increasing. To improve his situation, we proposed a novel investigation process that considers specific, unique

Piece Info.			
Packet No.	Time(sec)	Message Size(bytes)	IDX
1609	35	101	0x1d7
1609	35	101	0x131
1609	35	101	Ox11e
1609	35	101	0x1b3
1609	35	101	0x245
1609	35	101	0x17c
1609	35	101	0x164
1609	35	101	0x66
1609	35	101	0x1bf
1609	35	101	0x30b
1609	35	101	0x7e
1609	35	101	0x39
1609	35	101	0x208
1609	35	101	0x2eb
1609	35	101	0xbe

Fig. 9 Owned pieces of selected IP belongs to Case2

Piece Info.			
Packet No.	Time(sec)	Message Size(bytes)	IDX
1593	35	101	0x155
1593	35	101	0x30d
1593	35	101	0x33
1593	35	101	0x28f
1593	35	101	0x1bd
1593	35	101	0x2d0
1593	35	101	0x60
1593	35	101	0x2b7
1593	35	101	0x1b4
1593	35	101	0x48
1593	35	101	0x201
1593	35	101	0x132
1593	35	101	0xb2
1593	35	101	0x1fb
1593	35	101	0xb5

Fig. 10 Owned pieces of selected IP belongs to Case3

characteristics of BitTorrent. However, improvement of the law surrounding P2P investigations in regards to the application of the process is needed.

In this paper, an investigation process for illegal file sharing based on characteristics of file the sharing process using BitTorrent has been suggested. By following this process, an investigator can more effectively conduct an investigation about illegal file sharing.

Shared Data I	info.	
Packet No.	Time(sec)	Data Size(bytes)
187516	57	20
189947	57	20
426021	81	334
452607	86	20
452702	114	20
566464	114	25
567119	117	20
577685	118	1438
595352	120	1438
597932	122	1438
615967	123	1438
639351	125	1438
650406	126	1438
668238	127	1438
681196	128	1438

Fig. 11 Shared data information of selected IP

Acknowledgments This research was supported by the Public Welfare & Safety Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2012M3A2A1051106)

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- 1. BitTorrent, URL : http://en.wilipedia.org/wiki/BitTorrent
- 2. DHT, URL : http://en.wikipedia.org/wiki/Distributed_hash_table
- Harjinder Singh Lallie (2011) Windows 7 registry forensic evidence created by three popular BitTorrent clients. Digit Investig 7(3–4):127–134
- Hatahet S (2010) BiTIT: Throttling BitTorrent Illegal Traffic. Computers and Communications. 2010 I.E. Symp pp.708–713
- 5. Hyunjoo K (2006) Server selection schemes considering node status for a fault-tolerant streaming service on a peer-to-peer network. J Inf Process Syst 2: No.1
- 6. Intellectual Property, URL : http://en.wikipedia.org/wiki/Intellectual_property
- Liu J (2013) An efficient load balancing scheme for multi-gateways in Wireless Mesh Networks. J Inf Process Syst 9(3):365–378
- 8. MGM studios (2005) Inc. v. Grokster, Ltd., 545 U.S.913
- 9. MUlric ML (2008) Internet file-sharing: Swedish pirates challenge the U.S., 16 Cardozo J Int'l Comp L. pp.173
- 10. P2P, URL : http://en.wilipedia.org/wili/P2P
- Peng K. (2012) Attack and correction: how to design a secure and efficient Mix Network. J Inf Process Syst 8(1):175–190
- Schrader K (2009) Tracking contraband files transmitted using BitTorrent. Advances in Digital Forensics V, IFIP AICT 306, pp.159–173



Sooyoung Park received her Master's degree in Information Security, Korea University. She has performed projects related to Document Forensics, Network Forensics, and Database Forensics. Her research interests are Social Network Analysis, Cloud Forensics and Cyber Forensics.



Hyunji Chung received her Master's degree in Information Security, Korea University. She is now studying doctor course in Graduate School of Information Security, Korea University. She is currently working for national election commission. She has performed projects related to Cloud Forensics, Document Forensics, and Social Network Analysis. Her research interests are Social Network Analysis, Cloud Forensics and Cyber Forensics.



Changhoon Lee received his Ph.D. degree in Graduate School of Information Management and Security (GSIMS) from Korea University, Korea. He is now a professor at the Department of Computer Science and Engineering, Seoul National University of Science and Technology (SeoulTech), Korea. He has been serving not only as chairs, program committee, or organizing committee chair for many international conferences and workshops but also as a (guest) editor for international journals by some publishers. His research interests include information security, cryptography, digital forensics, smart grid security, computer theory etc. He is currently a member of the IEEE, IEEE Computer Society, IEEE Communications, IACR, KIISC, KIPS, KITCS, KMMS, KONI, and KIIT societies.



Sangjin Lee received his Ph.D. degree from Korea University. He is now a Professor in Graduate School of Information Security at Korea University and the head of Digital Forensic Research Center in Korea University since 2008. He has published many research papers in international journals and conferences. He has been serving as chairs, program committee members, or organizing committee chair for many domestic conferences and workshops. His research interests include digital forensic, steganography, cryptography and cryptanalysis.



Kyungho Lee received his Ph.D. degree from Korea University. He is now a Professor in Graduate School of Information Security at Korea University, and leading the Risk management Laboratory in Korea University since 2012. He has a high level of theoretical principles as well as on-site experience. He was a former CISO in NHN corporation, and now he takes as the CEO of SecuBase corporation. His research interests include information security management system (ISMS), risk management, information security consulting, privacy policy, and privacy impact assessment (PIA).