

“DEEMED DISTRIBUTION”: HOW TALKING ABOUT MUSIC CAN VIOLATE COPYRIGHT LAW

*Peter Bartoszek**

I. INTRODUCTION

Imagine that you are having a conversation with a friend in your home about the latest album by your favorite artist. You tell your friend that you have the album, show her the album case, and the friend thinks to herself that she would like to have it. You leave the room for a moment and then return. The next day, you find in your mailbox a complaint filed in federal court by the record label that published the album, alleging that because you let someone know that you had the album, and made it fairly easy for that person to access it, you have distributed it in violation of their exclusive right to do so. Although this may seem far-fetched, the court in *Elektra v. Barker* tacitly accepted a strikingly similar argument advanced by the recording industry: offering to distribute a file on an online sharing service may be tantamount to the illegal distribution of that file, even absent proof that the file was ever downloaded.¹ This interpretation of the law is far from universally accepted, as other courts have not rushed to accept the recording industry’s construction of what distribution really is.² The concept in *Barker* can be thought of as the deemed distribution doctrine, because a court will deem that distribution has occurred when a file is available for download, even though it has not technically been distributed to anyone.

This Note analyzes the validity of that proposition, starting with an overview of how file-sharing works and why it presents this unique opportunity for courts to extend the concept of distribution. This Note then examines the evolution of the deemed distribution doctrine, and shows that it does not work properly in the file-sharing framework, nor does it comport with

* J.D., University of Illinois College of Law, 2009; B.A., Law, Letters, and Society; Political Science, University of Chicago, 2006.

1. *Elektra Entm’t Group, Inc. v. Barker*, 551 F. Supp. 2d 234, 243–45 (S.D.N.Y. 2008) (stating that the distribution right guaranteed by the Copyright Act may be infringed by an offer to distribute, although merely alleging that files were made available, without alleging that they were also distributed, is not enough to state a claim).

2. *Atl. Recording Corp. v. Howell (Howell II)*, 554 F. Supp. 2d 976, 985 (D. Ariz. 2008) (expressly holding that record companies must prove actual distribution under copyright law).

the traditional decisions from which it was derived. For these reasons, courts should not allow proof of distribution without requiring proof that some file transfer occurred; a mere showing that some files appeared to be shared because they were in a shared files folder should not be enough to create liability. Deciding this issue is especially poignant now, as a defendant who was recently hit with a landmark judgment in a file-sharing case was granted a new trial, with the court expressing concern over the viability of the deemed distribution theory.³ Courts have also expressed an interest in understanding file-sharing technology as it applies to offers to distribute;⁴ this Note aims to provide that understanding.

II. BACKGROUND

This section provides background in three areas that are of fundamental importance to the issue of deemed distribution: copyright law and file-sharing, the evolution of the deemed distribution doctrine, and the existing legal support for discrediting the deemed distribution theory.

A. Copyright Law and File-Sharing

Musical works are afforded protection by the same copyright statute that applies to written works, and the protections they receive are generally the same.⁵ These rights include the right to reproduce the work, the right to create derivatives based on the work, the right to distribute the work, the right to perform the work publicly, and the right to display the work publicly.⁶ Sound recordings receive an additional protection: copyright holders have the exclusive right “to perform the copyrighted work publicly by means of a digital audio transmission.”⁷ This Note focuses on the interplay between the distribution section of the copyright statute, 17 U.S.C. § 106(3), and the Recording Industry Association of America’s argument that file-sharers on the Internet are engaged in distributing copyrighted works. The relevant statutory language states that:

[T]he owner of copyright under this title has the exclusive rights to do and to authorize any of the following: . . . to distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership, or by rental, lease, or lending⁸

To understand the arguments for and against the idea that file-sharing is akin to distribution of copyrighted works, it is important to first understand

3. Capitol Records, Inc. v. Thomas, No. 06-1497, 2008 WL 4405282, at *10 (D. Minn. Sept. 24, 2008).

4. Interscope Records v. Duty, No. 05-CV-3744-PHX-FJM, 2006 WL 988086, at *2 n.3 (D. Ariz. Apr. 14, 2006).

5. For a review of the applicability of copyright law to recorded music and the implications of that law in a digital world, see generally Michael W. Carroll, *A Primer on U.S. Intellectual Property Rights Applicable to Music Information Retrieval Systems*, 2003 U. ILL. J.L. TECH. & POL’Y 313, 316–21 (2003).

6. 17 U.S.C. § 106 (2006); see also Carroll, *supra* note 5, at 316–17 (describing the rights granted by a copyright).

7. 17 U.S.C. § 106(6).

8. *Id.* § 106(3).

how file-sharing works.⁹ The most common type of file-sharing involves peer-to-peer (“P2P”) technology.¹⁰ As the name implies, such a network provides a way for file transfers between “peers” on a network, as opposed to the more traditional notion of transfers between a client and a server.¹¹ The backbone of file-sharing, therefore, is the user who provides the content; and without the horizontal connections between users and the content they provide, the network cannot exist.¹² Users that share files (intentionally or otherwise) have files in a directory that is monitored by the P2P software that it then “shares” with other users.¹³

P2P software essentially provides other users of the network a listing of file names that are available for download from a certain computer, and it then enables a transfer of those files between the computers.¹⁴ Some software is capable of identifying files with similar names, sizes, and other characteristics that are available on numerous computers throughout a network, and then distributing the transfer among the computers so that each sharing computer provides a portion of the file.¹⁵ The way that this information is provided to other users differs based on the type of network. Most decentralized networks do not keep an index of the files that users make available; rather, computers running file-sharing software respond to search requests on an individual basis, checking to see if any files that it is sharing match a particular query.¹⁶ On these types of networks, a user that downloads files may appear to be searching a catalog of available files; however, this catalog is created only after a search is run, and it does not exist as an entity or file anywhere; it is temporal and

9. The information presented here is intended as a basic primer. More detailed information is available in such sources as: Jesse M. Feder, *Is Betamax Obsolete?* Sony Corp. of America v. Universal City Studios, Inc. in the Age of Napster, 37 CREIGHTON L. REV. 859 (2004); Seth A. Miller, *Peer-to-Peer File Distribution: An Analysis of Design, Liability, Litigation, and Potential Solutions*, 25 REV. LITIG. 181 (2006); Stephanos Androutsellis-Theotokis & Diomidis Spinellis, *A Survey of Peer-to-Peer Content Distribution Technologies*, 36 ACM COMPUTING SURV. 335 (2004), available at <http://portal.acm.org/citation.cfm?doid=1041680.1041681>.

10. See Javed I. Khan & Adam Wierzbicki, *Guest Editors' Introduction: Foundation of Peer-to-Peer Computing*, 31 COMPUTER COMM. 187, 187 (2008) (“Millions of users now participate in these systems, and the user bases are spreading like wildfire.”).

11. DETLEF SCHODER ET AL., CORE CONCEPTS IN PEER-TO-PEER NETWORKING 2–3 (2005), available at <http://www.idea-group.com/downloads/excerpts/Subramanian01.pdf>.

12. This has been the traditional implementation of P2P file-sharing systems, and modern systems retain the key trait that content comes from other users, rather than a centralized location; some vendors have leveraged the distributed content aspect of P2P systems to provide a novel way for distributing all types of content. See, e.g., BitTorrent DNA – Technology, <http://www.bittorrent.com/dna/technology> (last visited Nov. 19, 2008) (“Our proprietary transport technology leverages the full available network capacity of all paths without disrupting other applications. By detecting the presence of other applications, computers, and devices sharing the consumer’s broadband connection, BitTorrent DNA automatically moderates its use of the network to ensure that web browsing, voice over IP (VoIP), Internet gaming, and other applications are not disrupted.”).

13. Miller, *supra* note 9, at 190–91.

14. The actual mechanics of how the list of available files is published is significantly more complex and varies widely from network to network. See STEFAN SAROIU ET AL., A MEASUREMENT STUDY OF PEER-TO-PEER FILE-SHARING SYSTEMS 3, <http://www.cs.ucsb.edu/~almeroth/classes/F02.276/papers/p2p-measure.pdf> (last visited Dec. 6, 2008).

15. Marek Ciglan et al., *Striped Replication from Multiple Sites in the Grid Environment*, in ADVANCES IN GRID COMPUTING – EGC 2005 778, 780 (Peter M.A. Sloot et al., eds. 2005), available at <http://www.springerlink.com/content/1hfym8tprh7reg1a/fulltext.pdf>.

16. SAROIU ET AL., *supra* note 14, at 3.

constantly changing.¹⁷ Moreover, a search does not report that a user certainly possesses a file that assuredly contains the desired contents; it only reports which files match the specified characteristics.¹⁸

Once a file is located, actually downloading the file is ordinarily a trivial task, requiring nothing more than a click of the mouse.¹⁹ However, until a user takes this step, no content flows between the downloader and the user sharing the files.²⁰ In other words, until another user actually downloads the file through the network, the sharing user's computer does not transmit the actual file; only information describing the file is disseminated. This point is critical because some P2P software distinguishes between the concept of sharing generally and the actual uploading of files to other users' computers.²¹ Users of some software packages are able to prevent anyone from actually downloading a file from them, even if a search indicates that a download would be possible.²² Other software allows users to set the maximum amount of network bandwidth devoted to uploading files to zero, effectively disallowing the transfer of files even if they appear in a search.²³ Additionally, users who appear to be sharing files when other users perform a search may be behind a firewall that allows a search to happen, but would not allow an actual file transfer.²⁴

A user's shared files may also be encrypted using Digital Rights Management ("DRM") technology, so that the file will only play on the original user's computer.²⁵ In such a scenario, although these files would

17. *But see* JIAN LIANG ET AL., THE KAZAA OVERLAY: A MEASUREMENT STUDY 4 (2004), available at <http://cis.poly.edu/~ross/papers/KazaaOverlay.pdf>. Some P2P programs, such as KaZaA, use distributed computers that act as "nodes," which automatically poll computers connected to the network to see whether they are still sharing files. *Id.* Not every P2P network functions in this way, however.

18. Ciglan et al., *supra* note 15, at 780.

19. Kazaa, Quick Guide, <http://www.kazaa.com/us/help/quickstart.htm> (last visited Dec. 6, 2008).

20. *Id.* This information is presented elsewhere in KaZaA's information for parents, which states: "If you identify material you believe to be illegal or inappropriate, do not download that material." Kazaa, The Guide – Information for Parents, http://www.kazaa.com/us/help/new_parentsguide.htm (last visited Dec. 6, 2008). This implies that there is an interim step between finding the material and actually downloading it to a computer.

21. *See* NSIT: Disabling Peer-to-Peer File-Sharing, <http://nsit.uchicago.edu/services/safecomputing/disableptp/> (last visited Dec. 6, 2008) (explaining that instructions for disabling the uploading of files to other users in various software packages indicates that some packages (e.g. Gnutella, BearShare) provide specific options for limiting file uploads that exist separately from the option to stop monitoring a shared folder for files that could be reported in a search); P2P Guide: Bearshare – ResNet, <http://www.montana.edu/resnet/bearshare.php> (last visited Dec. 6, 2008) (same).

22. NSIT: Disabling Peer-to-Peer File-Sharing, *supra* note 21; P2P Guide: Bearshare – ResNet, *supra* note 21.

23. Montclair State University, Disabling File-Sharing, <http://oit.montclair.edu/resnet/disable.html> (last visited Dec. 6, 2008).

24. For example, on KaZaA, search traffic that conveys information about what files are available is encrypted in a proprietary way that makes it more difficult for firewalls and other filtering techniques to block than the actual transfer of files that are not encrypted. LIANG ET AL., *supra* note 17, at 5; *see also* Marc Grote, *Configuring the ISA Server 2006 HTTP Filter*, ISAServer.org, Feb. 8, 2007, <http://www.isaserver.org/tutorials/Configuring-ISA-Server-2006-HTTP-Filter.html> (describing a method for network administrators to implement software that can detect KaZaA traffic and effectively discard it); FileSharingZ.com, Kazaa Lite Resurrection FAQ, http://www.filessharingz.com/faqs/kazaa_lite/index.php (last visited Dec. 6, 2008) (describing the presence of "red crosses" in search results when a file is probably unable to be downloaded).

25. *Q&A: What Is DRM?*, BBC NEWS, Apr. 2, 2007, <http://news.bbc.co.uk/2/hi/technology/6337781.stm>.

appear in search results, a user who downloaded them would be unable to access the content.

B. The Evolution of the Deemed Distribution Doctrine

The origins of the deemed distribution doctrine advanced in *Barker* can be traced back as far as *Hotaling v. Church of Jesus Christ of Latter-Day Saints*. *Hotaling* states the three factors that constitute distribution to the public by a library:

- i. The copyrighted work is added to the collection;
- ii. The work is added to the index;
- iii. The work is made available to the browsing public.²⁶

The basic reasoning behind this doctrine seems fairly simple:

When a public library adds a work to its collection, lists the work in its index or catalog system, and makes the work available to the borrowing or browsing public, it has completed all the steps necessary for distribution to the public. At that point, members of the public can visit the library and use the work.²⁷

How this applies to the online world is less clear. The argument is seemingly that once the user has made known to another user that a work is available, the remaining step of actually distributing the work is immaterial; the offer to distribute itself infringes an owner’s copyright. This is the deemed distribution doctrine.

The next step in the evolution of this doctrine came in *A&M Records, Inc. v. Napster, Inc.* Citing no authority, but referencing the logic adopted in *Hotaling*, the court stated: “Napster users who upload file names to the search index for others to copy violate plaintiffs’ distribution rights.”²⁸ This statement was cited and approved of in *Interscope Records v. Duty*.²⁹

The concept appears again in *Perfect 10, Inc. v. Amazon.com, Inc.*, where the court found that Google did not distribute Perfect 10’s images because:

Google does not own a collection of Perfect 10’s full-size images and does not communicate these images to the computers of people using Google’s search engine. Though Google indexes these images, it does not have a collection of stored full-size images it makes available to the public. Google therefore cannot be deemed to distribute copies of these images under the reasoning of *Napster* or *Hotaling*.³⁰

The *Perfect 10* case highlights two tensions extant within the deemed distribution concept: (1) Google was not held liable because they do not actually possess the images, and (2) Google was not held liable because they

26. *Hotaling v. Church of Jesus Christ of Latter-Day Saints*, 118 F.3d 199, 203 (4th Cir. 1997).

27. *Id.*

28. *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1014 (9th Cir. 2001).

29. *Interscope Records v. Duty*, No. 05-CV-3744-PHX-FJM, 2006 WL 988086, at *2 (D. Ariz. Apr. 14, 2006).

30. *Perfect 10, Inc. v. Amazon.com, Inc.*, 487 F.3d 701, 719 (9th Cir. 2007).

did not communicate the images.³¹

Barker draws a subtle distinction between making a file available and offering it for distribution. The *Barker* court held that the record company plaintiffs needed to plead within the text of the copyright statute, which does not have an express “making available” requirement.³² The court used this distinction to explain why it refused to accept the plaintiff’s “making available” theory expressly; in fact, the *Hotaling* court’s decision was criticized in *Barker*.³³ Nevertheless, the court did hold that the distribution right may be infringed by an offer to distribute.³⁴ In other words, although the *Barker* court did not give copyright holders the full bundle of rights that they sought, the fundamental concept behind the deemed distribution theory remains in play.

C. Support for Discrediting Deemed Distribution

The clearest statement against deemed distribution comes from a case where the court originally accepted the potential validity of the theory. In *Atlantic Recording Corp. v. Howell*, the court examined the concept of deemed distribution and found that an offer to distribute a file, including offers that would fit into the recording industry’s “making available” theory, did not constitute distribution.³⁵ The court opined that “[i]t is untenable that the definition of a different word in a different section of the statute was meant to expand the meaning of ‘distribution’ and liability under § 106(3) to include offers to distribute.”³⁶

There is other precedent for expressly disavowing the use of the deemed distribution doctrine in P2P cases. In granting summary judgment to the defense, a California district court held in *In re Napster, Inc. Copyright Litigation*:

[P]laintiffs’ “indexing” theory falls well short of meeting the requirements for establishing direct copyright infringement. Rather than requiring proof of the actual dissemination of a copyrighted work or an offer to distribute that work for the purpose of its further distribution or public performance, plaintiffs’ theory is premised on the assumption that any offer to distribute a copyrighted work violates section 106(3). This is not sufficient to satisfy plaintiffs’ burden of proving that Napster or its users directly infringed their copyrighted musical compositions and sound recordings, as they must do if they are to hold defendants secondarily liable for that infringement.³⁷

Strangely, this California court took this position after the Napster

31. *Id.*

32. *Elektra Entm’t Group, Inc. v. Barker*, 551 F. Supp. 2d 234, 244–45 (S.D.N.Y. 2008).

33. *Id.* at 243.

34. *Id.* at 245.

35. *Atl. Recording Corp. v. Howell (Howell II)*, 554 F. Supp. 2d 976, 984–85 (D. Ariz. 2008).

36. *Id.* at 985.

37. *In re Napster, Inc. Copyright Litig.*, 377 F. Supp. 2d 796, 805 (N.D. Cal. 2005).

litigation that seemingly accepted deemed distribution conceptually.³⁸ This position is in accordance with one taken in the Eighth Circuit in 1993, before Napster and other P2P software became popular among end users.³⁹ There, the court held that “actual dissemination” was required in order for copyright to be violated.⁴⁰ This case was about software, and not about media content, but the general concept of “actual dissemination” when dealing with an electronic medium is the important element to consider.

III. ANALYSIS

This Note makes four main arguments: (1) a legal framework that can apply to file-sharing cases already exists in *Perfect 10*, (2) using *Hoteling* in a file sharing context is inappropriate, (3) attempting to apply a system that will guarantee a file is copyrighted before it is downloaded might be impossible, and (4) treating partially transferred files as full copyright violations is unfair.

A. *Perfect 10 Provides One Correct Analytical Framework that Rejects the Viability of Deemed Distribution as a Way of Analyzing File-Sharing Cases*

The *Perfect 10* court found that Google had not violated copyright protections for two reasons that can be applied in a similar way to file-sharing cases where the copyright holder attempts to use the deemed distribution doctrine to prove liability: (1) search results themselves do not reveal whether one actually possesses the copyrighted work, and (2) search results are not communications of the actual copyrighted work.⁴¹

1. *Actual Possession of Copyrighted Works*

The court in *Perfect 10* held that the deemed distribution doctrine did not apply to Google because Google “does not have a collection of stored full-size images it makes available to the public.”⁴² The court ruled this way because it found that the way Google’s technology operated meant that it never actually stored any version of the copyrighted work.⁴³ If the deemed distribution doctrine relies on the knowledge that the alleged violator possessed the

38. *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1014 (9th Cir. 2001).

39. Napster, the first popular music P2P program, was developed in 1998–1999. Brad King, *The Day the Napster Died*, WIRED.COM, May 15, 2002, <http://www.wired.com/gadgets/portablemusic/news/2002/05/52540>.

40. *Nat’l Car Rental Sys. v. Computer Assocs. Int’l, Inc.*, 991 F.2d 426, 434 (8th Cir. 1993); see also *Typical Claims and Counterclaims in Peer to Peer Litigation*, ELECTRONIC FRONTIER FOUNDATION, http://w2.eff.org/IP/P2P/RIAA_v_ThePeople/P2P_typical.pdf (last visited Nov. 11, 2008) (summarizing the court’s decision in *National Car Rental System*).

41. See discussion *infra* Parts III.A.1–2. (discussing both the actual possession of copyrighted and the communication of copyrighted works prongs identified in *Perfect 10*). See generally *Perfect 10, Inc. v. Amazon.com, Inc.*, 487 F.3d 701 (9th Cir. 2007) (concluding that *Perfect 10* was unlikely to succeed in its direct infringement claim).

42. *Perfect 10, Inc.* 487 F.3d at 719; see also discussion *supra* Part II.B (describing how the deemed distribution doctrine did not apply to Google because Google did not communicate the images to other computers).

43. *Perfect 10, Inc.* 487 F.3d at 719.

copyrighted works, it follows that when knowledge of possession is uncertain, the doctrine ought not to be applied. In the P2P scenario, one cannot tell if a user is sharing a work that is copyrighted or merely sharing a file that has a similar name to and similar attributes of a copyrighted work.⁴⁴

However, some courts have ignored this simple logical step from *Perfect 10* to the world of P2P sharing. The exact wording of the deemed distribution doctrine as applied to P2P music sharing in *Warner v. Payne* is: "Listing unauthorized copies of sound recordings using an online file-sharing system constitutes an offer to distribute those works, thereby violating a copyright owner's exclusive right of distribution."⁴⁵ In this opinion, the *Warner* court denied a summary judgment motion based on the defendant's argument that listing the works was not tantamount to distribution.⁴⁶ The district court held that this argument was not dispositive, and the court refused to rule out the "making available" theory as possible grounds for liability.⁴⁷

Neither the plaintiffs in a copyright distribution case nor any court can know for certain whether the works being offered would violate copyright law just by looking at search results. Without actually downloading the files in question, it is impossible to tell for certain that they are what they purport to be.⁴⁸ For example, one could create a text file consisting of one letter repeated over and over thousands of times, name it "The Beatles – Eight Days a Week.mp3," and put it in a shared folder. Assuming that a P2P program reported in its search results only the name of a file, it would appear that this user is offering a copyrighted work for distribution. By the logic of the deemed distribution doctrine, this action would violate copyright law, even though the copyright holder's work was at no point anywhere close to being distributed; the works may not even be available for distribution.⁴⁹

The deemed distribution doctrine is fundamentally important to the record industry's cases against consumers, because it provides them an easy way of proving infringement.⁵⁰ By merely showing that some works are available but

44. See SAROIU ET AL., *supra* note 14, at 4 (discussing how the Napster P2P architecture performs searches by file name).

45. Warner Bros. Records, Inc. v. Payne, No. W-06-CA-051, 2006 WL 2844415, at *3 (W.D. Tex. July 17, 2006).

46. *Id.* at *2.

47. *Id.* at *4.

48. Software products exist that claim to be able to detect music files that are not what they purport to be; such software would not exist if it was assured that every file shared on a network was indeed the file it claimed to be. See Fake MP3 Detector 2.2, http://www.sharewareplaza.com/Fake-MP3-Detector-download_32834.html (last visited Dec. 6, 2008) ("Fake MP3 detector can scan your collection of music and detect corrupt/fake MP3 songs, particularly useful if you download songs from file sharing networks and websites."). It is important to note that this software requires a file to actually be downloaded from the network before it can ascertain whether it is actually a "fake" MP3.

49. This example is perhaps oversimplified, since most P2P sharing programs are able to gather specific attributes about shared media files, such as their file size, length, and the bitrate at which they are encoded. Rebecca Viksnins, *Kazaa 2.0 Document Management Reviews – CNET Reviews*, CNET.COM, http://reviews.cnet.com/document-management/kazaa-2-0/4505-3522_7-20456501.html (last visited Oct. 21, 2008). However, one could record an audio file where these attributes are identical to those of a copyrighted music file, and a standard P2P program would be unable to tell the difference.

50. The courts in both *Howell I* and *Payne* allowed distribution to be shown by providing evidence that files were available for download. Warner Bros. Records, Inc. v. Payne, No. W-06-CA-051, 2006 WL

not that they actually were copyrighted works, the record industry plaintiffs do not show that copyrighted works were distributed.⁵¹

The only thing that plaintiffs prove without actually downloading the file⁵² is that the user has a file on his or her computer in a shared directory that has a similar name and possibly similar attributes to that of a copyrighted work. This is not the same as showing that the sharing user actually possessed copyrighted work, and applying the *Perfect 10* analysis would lead a court to conclude that the defendants were not liable for infringement since plaintiffs could not show that defendants actually possessed the works.

2. Communication of Copyrighted Works

Perfect 10 would also require the record industry’s arguments to fail because they did not prove that an actual communication of the copyrighted works ever took place.⁵³ Because Google did “not communicate these images to the computers of people using Google’s search engine,” the *Perfect 10* court held that the deemed distribution doctrine did not apply to Google.⁵⁴ This is the central argument behind disfavoring the deemed distribution doctrine in P2P cases. Understanding the *Perfect 10* court’s view of Google’s image search technology is essential to recognizing why consumers who have not actually communicated copyrighted files to other users of P2P networks should not be held liable for illegal distribution of the works.

“Google Image Search provides search results as a webpage of small images called ‘thumbnails,’ which are stored in Google’s servers. The thumbnail images are reduced, lower-resolution versions of full-sized images

2844415, at *3–4 (W.D. Tex. July 17, 2006) (“However, the Court finds that offering to distribute a music file by listing it on an online file-sharing system contemplates ‘further distribution.’ Making an unauthorized copy of a sound recording available to countless users of a peer-to-peer system for free certainly contemplates and encourages further distribution, both on the Internet and elsewhere.”). In ruling on the record industry’s motion to dismiss in *Howell I*, the district court accepted as evidence of distribution the screenshots that plaintiffs provided, showing that certain files may have been available for download. *Atl. Recording Corp. v. Howell (Howell I)*, No. CV06-02076-PHX-NVW, 2007 WL 2409549, at *3 (D. Ariz. Aug. 20, 2007) (“[T]he mere presence of copyrighted works in a shared folder is enough to trigger liability.”). This decision was later reversed, however, by that same court on reconsideration. *Atl. Recording Corp. v. Howell (Howell II)*, 554 F. Supp. 2d 976, 982–86 (D. Ariz. 2008). The record industry continues to argue that evidence of an actual download “is unnecessary to prove a violation of their distribution rights under § 106(3).” *Id.* at 981.

51. An example of the types of evidence presented by the record industry in such cases can be found in the evidentiary exhibits in the complaint in *Elektra Entertainment Group, Inc. v. Barker*. Complaint at Exhibit B, Part 1, *Elektra Entm’t Group, Inc. v. Barker*, 551 F. Supp. 2d 234 (S.D.N.Y. 2008) (No. 05-CV-7340) (listing username of the user sharing the file, the name of the file being shared, the size of the file, the type of the file, and the artist of the file as read by the P2P software from the “tag” within the file).

52. In *Howell II*, the plaintiffs actually downloaded twelve files that they confirmed to be copyrighted music files. *Howell II*, 554 F.Supp.2d at 983. The court points out, however, that the copyright statute is not violated “unless the defendant has actually distributed an unauthorized copy of the work to a member of the public.” *Id.* It is unclear whether the court is including the copyright holder in its definition of the public; at least one court has held that a record industry plaintiff that used an investigator to download copyrighted files from other users could use those actual downloads as proof of infringement. *Capitol Records, Inc. v. Thomas*, No. 06-1497, 2008 WL 4405282, at *11 (D. Minn. Sept. 24, 2008).

53. *Perfect 10, Inc. v. Amazon.com, Inc.*, 487 F.3d 701, 719 (9th Cir. 2007).

54. *Id.* Note that the court in *Perfect 10* acknowledged the *Hotelling* decision and distinguished it because Google proved that they did not at any time possess a collection of the images in question. *Id.* This distinction can be extended to P2P sharing because of the plaintiff’s burden of proving a copyright violation.

stored on third-party computers.”⁵⁵ These thumbnail images serve essentially as descriptors to users of what the full-size image actually is. Much like the title, file size, and other information about a music file are all specific traits that those looking for a copyrighted file can use to determine whether a file is what they are looking for, the thumbnail image provided by Google lets users know whether they have found the image that they are looking for.⁵⁶

Google then makes it easy for users to get access to the full-size images. “By following the HTML instructions to access the third-party webpage, the user’s browser connects to the website publisher’s computer, downloads the full-size image, and makes the image appear at the bottom of the window on the user’s screen.”⁵⁷ This step of actually accessing the full, copyrighted image is what the *Perfect 10* court required to prove infringement.⁵⁸ Google’s liability as a provider of the thumbnails was limited, because “Google does not store the images that fill this lower part of the window and does not communicate the images to the user; Google simply provides HTML instructions directing a user’s browser to access a third-party website.”⁵⁹ As discussed earlier, whether a user actually possesses the copyrighted works that he or she is purporting to share is unknown,⁶⁰ and the files themselves are not transmitted to a user who does not download them.⁶¹ By providing thumbnail images, Google enables the last step before infringement can occur; in the same way that publishing a list of works that can be downloaded does. Google transmits a small thumbnail version of the images that could function in the same way that a song title does in alerting a user to what the content of a file should be.⁶² The *Perfect 10* court found that the transmission and storage of these thumbnail images does not violate copyright.⁶³

Breaking this analogy down further, it is evident that the *Perfect 10* framework works quite well in the P2P context.

Instead of communicating a copy of the image, Google provides HTML instructions that direct a user’s browser to a website publisher’s computer that stores the full-size photographic image. Providing these HTML instructions is not equivalent to showing a copy. First, the HTML instructions are lines of text, not a photographic image. Second, HTML instructions do not themselves cause infringing images to appear on the user’s computer screen. The HTML merely gives the address of the image to the user’s browser.

55. *Id.* at 711.

56. *Id.*

57. *Id.*

58. *Id.* at 716.

59. *Id.* at 711.

60. See discussion *supra* Part III.A.1 (noting one cannot tell whether a user is sharing a copyrighted work or a work with a similar file name, or similar attributes of a copyrighted work).

61. See *supra* Part II.A.

62. *Perfect 10*, 487 F.3d at 716–17.

63. The court attaches some weight to the fact that Google does not store a copy of the full images, but this distinction seems trivial in light of the problems that can interfere with actual transmission of the file. *Id.* at 717; see also *infra* Part III.B.3 (discussing the difficulty with determining whether a file exists and can be downloaded when presented only with search results).

The browser then interacts with the computer that stores the infringing image. It is this interaction that causes an infringing image to appear on the user’s computer screen.⁶⁴

A user who is sharing files via P2P software is in some ways doing the same thing that Google did in *Perfect 10*; the user is providing instructions on how to obtain a file. Instead of a link coupled with a thumbnail image, the P2P user is providing instructions that are represented by textual descriptions of files and computer commands, and not copyrighted songs.⁶⁵ These distinctions apply directly to the situations in *Barker*. Even though Google does not own the images that its search results include, it gives the user who wants the copyrighted work an easy opportunity to obtain that work, in the same way that someone who runs P2P software is essentially providing others with a list of works available for download.⁶⁶ However, in *Perfect 10*, the ease of accomplishing the infringing step was not more relevant to the court’s decision that Google did not violate copyright than the fact that Google had nothing to do with taking the infringing step.⁶⁷ By analogy, simply because one might be able to easily download a file after running a search, actually downloading the file may prove difficult, if not impossible, and without knowing if this is the case, plaintiffs cannot show that the next step of infringement was attainable.

The deemed distribution doctrine can be taken to the extreme, as proposed in the introduction of this Note.⁶⁸ Telling someone that you own an album and then making it easy for them to take it from you constitutes distribution under the deemed distribution doctrine. Loaning a friend a CD could also constitute deemed distribution, if not accompanied with a warning not to copy it, since making a copy of the songs on a CD is simple.⁶⁹

The deemed distribution doctrine also could create numerous false positives, where file-sharing is inadvertent or temporal before the user realizes they are sharing copyrighted works. For example, a user may have just installed the file-sharing program when a plaintiff runs a search and observes that the user is sharing seemingly copyrighted files. Basing a copyright violation on the appearance of a file in search results, as the deemed distribution doctrine would do, would create liability even if the alleged violator had turned off file-sharing or removed the files in question from the shared directory, making them unavailable.⁷⁰ Although this example by itself may not seem highly likely, it illustrates the fundamental unfairness of the

64. *Perfect 10*, 487 F.3d at 717.

65. See *supra* Part II.A (discussing the difference between search results and actual transmission of a file).

66. Again, there is a distinction drawn by the court in that Google does not own the copyrighted images. *Perfect 10*, 487 F.3d at 719. Although this nuance carries some weight, when taken with the arguments discussed *infra*, the distinction becomes less important. See *infra* Part III (describing the difficulty in determining whether someone who is sharing a work actually has a copy of that work without downloading it).

67. *Perfect 10*, 487 F.3d at 717.

68. See *supra* Part I (noting possible violations when making friends aware that you own an album).

69. See Windows Vista Help: Rip Music from a CD, <http://windowshelp.microsoft.com/Windows/en-US/Help/dcee3798-8afb-4f55-b17d-fed1c85da6531033.msp> (last visited Oct. 21, 2008) (outlining the procedure for copying music from a CD to a computer).

70. See *infra* Part III.B.3 (discussing more ways that actual distribution might not occur).

deemed distribution doctrine as the *Barker* court developed it.⁷¹

B. *The Analogy to Hotaling Is Fatally Flawed*

Similar issues exist when attempting to apply the public library rule of *Hotaling* to a system where individual users control the content. Whereas a library could reasonably be expected to catalog its works properly (although even this claim may not be entirely true), this expectation cannot be validly applied to individuals who may or may not catalog works correctly.⁷² In other words, a user may have mislabeled certain files that were not copyrighted, and even though they might appear in a search for copyrighted works, they are not actually files that violate copyright.⁷³ The screenshots used as evidence in *Howell* only show that files had the name of certain copyrighted works and some attributes of those copyrighted works, but without more, it is impossible to tell whether these are actually the works they claim to be.⁷⁴ The only way to tell whether a user actually possesses the copyrighted work is to download the work from the user and listen to or watch it. Copyright holders themselves, however, could not use such downloaded files to show infringement because “[i]t is well-established that the lawful owner of a copyright cannot infringe its own copyright.”⁷⁵ Some courts have held, however, that downloads performed by an investigator hired by a copyright holder can be used to show infringement.⁷⁶

Where a file is not actually downloaded, myriad issues exist with the *Hotaling* construct, which fails to account for the possibility that the shared files may not be downloadable, or are encrypted.⁷⁷ If the file is encrypted, one

71. *Elektra Entm't Group, Inc. v. Barker*, 551 F. Supp. 2d 234 (S.D.N.Y. 2008).

72. The information describing files that is returned in search results is encoded in the music files, and it can be initially provided by the user, or it can be provided by an online database service. Some of these services contain user-contributed information, which can be incorrect. See freedb.org, About freedb.org, http://www.freedb.org/en/about_freedb.org.2.html (last visited Dec. 6, 2008) (explaining that users submit data, but the site cannot guarantee that all data submitted is correct).

73. Some music organizing software attempts to fill in missing information about files from an online database. Apple – iTunes – iTunes Jukebox – Importing, <http://www.apple.com/au/itunes/jukebox/importing.html> (last visited Nov. 19, 2008) (describing Apple's information importing procedures); see also Windows Vista Help: Add or Edit Media Information in Windows Media Player, <http://windowshelp.microsoft.com/Windows/en-US/Help/4da102ff-2cb3-4949-ab3f-afd8828a4be41033.msp> (last visited Oct. 21, 2008) (explaining that Windows Media Player can download information and add it automatically). This information is based on specific traits about the file, and the automated process can incorrectly identify certain songs; numerous anecdotes abound about this feature incorrectly relabeling some files. Posting of neekap to Gizmodo: Complete Zune Guide – Complete Guide to Zune 2's Software and Firmware, <http://gizmodo.com/gadgets/complete-zune-guide/complete-guide-to-zune-2s-software-and-firmware-321605.php> (Nov. 12, 2007, 19:55) (claiming that software incorrectly labeled files); Posting of dickrhee to Gizmodo: Complete Zune Guide – Complete Guide to Zune 2's Software and Firmware, <http://gizmodo.com/gadgets/complete-zune-guide/complete-guide-to-zune-2s-software-and-firmware-321605.php> (Nov. 13, 2007, 02:09) (claiming software replaced tags with incorrect information).

74. See Complaint at Exhibit B, Part 1, *Elektra Entm't Group, Inc. v. Barker*, 551 F. Supp. 2d 234 (S.D.N.Y. 2008) (No. 05-CV-7340) (containing exhibits offered by record industry plaintiffs from *Elektra*).

75. *Olan Mills, Inc. v. Linn Photo Co.*, 23 F.3d 1345, 1348 (8th Cir. 1994).

76. *Capitol Records, Inc. v. Thomas*, No. 06-1497, 2008 WL 4405282, at *11–12 (D. Minn. Sept. 24, 2008).

77. See *supra* Part II.A (noting that until the user actually downloads the file, no content flows between the downloader and the individual sharing the file).

could characterize the situation as one where a user has looked up a book in the catalog and found the book on the shelf, but once they try to open it, they find an unbreakable lock that prevents them from accessing the content inside.⁷⁸ In both situations, the content that is protected by copyright law has not been distributed to the user. The source of information about where that content exists should not be held liable for copyright infringement simply because it appeared that actual distribution was imminent.

Turning again to the three-prong test established in *Hotaling*, a work is deemed to have been distributed if the following factors are true:

- i. The copyrighted work is added to the collection;
- ii. The work is added to the index;
- iii. The work is made available to the browsing public.⁷⁹

Even if this test and the deemed distribution doctrine were not fundamentally incompatible with file-sharing schemes, as detailed above, it is difficult to see how the test could indicate deemed distribution in the online file-sharing context. What follows is an analysis of the three factors of the *Hotaling* test as applied to file-sharing. None of these factors is analogous to the file-sharing paradigm, and they illustrate both why this test should not be used and why the logic behind it cannot apply in the online world.

1. *The Copyrighted Work Is Not Added to a Collection*

The first part of the test requires that a copyrighted work be added to a collection. In the context of a library, it is fairly simple to understand how this process would work: a library would purchase a copyrighted work such as a book or music recording, record information about the work and enter it into a catalog system, and then put the work in an accessible place within the library.⁸⁰

The analogous behavior in an online file-sharing system would be a user ripping copyrighted content to their hard drive, properly labeling the content with biographical information about the content, and then putting the file in a shared directory that is monitored and shared by a P2P sharing program.⁸¹

Only the first step in the file-sharing process bears a significant resemblance to what occurs in a library. The user must indeed acquire some form of the copyrighted work, just as the library must acquire a work.⁸² After

78. See *supra* Part II.A (noting that a search only reports files that match specified characteristics, but without downloading a file it is difficult to determine whether the file is what it purports to be).

79. *Hotaling v. Church of Jesus Christ of Latter-Day Saints*, 118 F.3d 199, 203 (4th Cir. 1997).

80. See generally JOHN COTTON DANA, A LIBRARY PRIMER 99–102 (3rd ed.1903), available at <http://www.gutenberg.org/files/15327/15327-h/15327-h.htm> (providing a framework for libraries to follow from their implementation).

81. See discussion *supra* Part II.A (describing P2P networks).

82. Users who “acquire” the media files may have to go through several steps to acquire them. Users may simply download them from some other user, or they may “rip” the content from physical media onto a digital format on their computer. This process is fairly easy for audio files, but it is more complicated for movie files that feature built-in safeguards against copying. Windows Vista Help: Rip Music from a CD, *supra* note 69. See generally Jeffrey A. Bloom et al., *Copy Protection for DVD Video*, 87 Proc. of the IEEE 1267 (1999), available at <http://ieeexplore.ieee.org/iel5/5/16709/00771077.pdf>. (describing the various copy-

this step, however, the similarities end.

As discussed earlier, there is no certainty that files labeled in a certain way actually are the files that they purport to be.⁸³ Unlike the library that is expected to catalog its works reliably and uniformly so that patrons are able to find the works that they are looking for among the thousands available, individual users of P2P file-sharing services are not charged with a similar responsibility.⁸⁴ Although the search features of the P2P software programs can collect biographical information about files that are being shared (in fact, if search results are to provide anything more than merely a file name, they rely on this information),⁸⁵ generally it relies on the information that has already been input, and cannot gauge the accuracy of this information.⁸⁶

ID3 tags, a very popular form of biographical information that is encoded into audio files shared on the Internet, can be changed to read anything and do not have to correspond to the reality of the files that contain them.⁸⁷ Other descriptive information about files—such as their size, length, and encoding type—can be read by some P2P file programs, but this information in and of itself cannot definitively determine whether a work is copyrighted.⁸⁸ Until there is a reliable way of telling whether a given work is actually the copyrighted work in question before it is downloaded, the vast expanse of various types of files cannot legitimately be analogized to the collection of copyrighted works in a library.

Finally, the works on a P2P system should not be thought of as a

protection schemes in place for DVDs).

83. See discussion *supra* Part III.A.1 (noting that until the user actually downloads the file, no content flows between the downloader and the individual sharing the file).

84. See NEW YORK PUBLIC LIBRARY, 2006 ANNUAL REPORT 74–75 (2005), available at <http://www.nysl.org/pr/objects/pdf/2006AnnualReport.pdf>. The New York Public Library system has over 50 million items in its collection, with thousands of employees working in the libraries serving millions of patrons. This illustrates the sheer magnitude of the task of cataloging work in a library and helping make it available to library patrons.

85. This information has a different name depending on the format of the media in question. The most common audio files, MP3 files, use a system called ID3 tags. ID3.org, ID3v2Easy, <http://www.id3.org/ID3v2Easy> (last visited Oct. 21, 2008). ID3 tagging encodes information in an MP3 audio file about the artist, song title, album, date, and other information in a standardized format that playback software and some P2P file-sharing software can use to organize media. *Id.*

86. *But see* Fake Mp3 Detector 2.2, *supra* note 48. Software programs exist that will examine a downloaded file and determine whether it is the file that it purports to be. However, such a utility operates on the side of one who is actually doing the downloading and, therefore, requires the transfer of at least some of the file from the sender to the downloader. In other words, the search results themselves would not allow such a program to test whether the file is indeed the copyrighted work that it describes itself as; at least some part of the download has to occur first.

87. *Id.*

88. Some controversy exists as to whether there is technology available to filter copyrighted works at the level of the P2P service provider. Andy Sullivan, *Kazaa Could Filter Copyrighted Music, Critics Say*, USATODAY.COM, Jan. 14, 2004, http://www.usatoday.com/tech/webguide/music/2004-01-14-peerfilter_x.htm. Such technology could determine from some kind of digital “signature” whether a work was copyrighted or not. *Id.* It is generally thought that a successful implementation of such technologies would require some form of code written to run at the client level. If such technology were available, files that were copyrighted would then either be not-shared, or they could not be downloaded, even if they appeared in search results. In any event, these concerns apply more to P2P provider liability than they do to individual liability, except that if they could successfully be implemented, and if one could tell that a work was copyrighted before actually downloading, the analogy to *Hotelling* would be considerably more valid.

collection in the same sense as a library. A library is well known as a repository for works that can be relied upon to contain what it purports to contain.⁸⁹ In other words, if a library says that it has a copy of Jane Austen’s *Pride and Prejudice*, one could reasonably expect that it does, and that the work available is the actual book, not merely a dust cover from the book encasing a copy of the telephone directory. The P2P world does not have this same legitimacy; there is no reliable way to know what kind of file users will get when they attempt a download.⁹⁰

2. *Copyrighted Works Are Not Actually Added to an Index in Modern File-Sharing Systems*

To avoid the fate of the original iteration of Napster (which kept a centralized index of files available at any given time on its P2P network), most current P2P sharing services employ a decentralized structure that allows them to exist without any centralized indexing system.⁹¹ This means that, unlike a catalog system in a library that is created and maintained with the express purpose of providing accurate information about the works contained within the institution keeping the catalog, the search results that are returned when one looks for a file on a P2P network are constantly changing.⁹² Works that exist in the search results may no longer be available seconds after it appears that they are.⁹³

This element in the P2P world bears the greatest similarity to its analogue in the library world. The search results do essentially allow a searcher to find a work that exists somewhere, and they provide the searcher with the “location” of the work. However, the other problems with search results that have been discussed earlier make the index comparison imperfect at best.

3. *Copyrighted Works Are Not Necessarily Available to the Browsing Public on a P2P System*

This argument echoes one found earlier in this Note regarding the actual communication of copyrighted works: works that are not actually communicated because of some technical barrier such as a firewall or setting in the P2P software are not “available to the public” in the same way that they

89. See generally DANA, *supra* note 80, at 78 (discussing the importance of cataloging works within a library, even in the smallest of libraries).

90. But see *infra* Part III.C (describing potential methods to identify works that appear in search results).

91. See *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1023–24 (9th Cir. 2001) (discussing the way that Napster could have policed its system because of the centralized, indexed nature that characterized its operations). For a discussion of decentralization, see *supra* Part II.A.

92. For example, the International Federation of Library Associations and Institutions propagates numerous standards for the cataloging of works within a library. See INT’L FED’N OF LIBRARY ASSOC’NS & INSTS., ISBD(G): GENERAL INTERNATIONAL STANDARD BIBLIOGRAPHIC DESCRIPTION, <http://www.ifla.org/VII/s13/pubs/isbdg.htm> (last visited Dec. 6, 2008) (describing the standard bibliographical descriptions to be used).

93. See S.H. Kwok & K.Y. Chan, *An Enhanced Gnutella P2P Protocol: A Search Perspective*, IEEE EXPLORE, at § 1 (2004), <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01283975> (discussing the fact that a disconnected node can no longer upload parts of a file).

would be in a library.⁹⁴ To better understand this distinction, posit a library that offered a catalog of books and their locations, but when one went to the shelf to retrieve the books, some of them were glued to the shelf, unable to be removed or opened to reveal the works inside. In a library, such a situation or something similar to it would be highly unlikely; in the P2P file-sharing world, as discussed earlier, such barriers are relatively commonplace.⁹⁵

C. Systems that Determine Unique Hashes Are Difficult to Apply

One possible way that the plaintiffs in P2P cases may be able to show that a work is copyrighted before it is downloaded is to have a way of determining the content of the work from other elements besides or in addition to the file name and the available ID3 tag information. A file on a computer can be compared to another file to determine whether the two are bit-for-bit identical to one another, but this requires that the files be available on the same computer; that is, the files need to have been downloaded.⁹⁶ However, there is a way to calculate whether two files are identical without directly comparing them, and without needing to have both files available.

The tool used for this purpose is called a hash function.⁹⁷ “A hash function takes binary data, called the message, and produces a condensed representation, called the message digest.”⁹⁸ Traditionally, this tool is used to verify that data has not been damaged or tampered with during its transmission, whether that transmission is over the Internet or through some other medium.⁹⁹

For instance, in order to verify that a large file has been received correctly over an unreliable network, it suffices to transmit the hash value of the received file to the sending party. If the returned hash value matches the hash value of the original file, then there is almost complete certainty that the file is correctly received. The only remaining uncertainty is related the probability of a *collision*: i.e., two different originals may have the same hash value. A carefully designed hash function minimizes the probability of collision.¹⁰⁰

Essentially, a hash function takes a string of data and outputs a shorter element of data that represents the input string uniquely.¹⁰¹ Changing even one

94. See discussion *supra* Part III.A.2 (noting that after finding a file in a search, the act of downloading the file may be difficult or prohibited).

95. See discussion *supra* Part III.A.2.

96. Most modern operating systems ship with some file comparison tool. Microsoft Windows Vista, for example, ships with the command “fc.exe,” which can compare files on a bit for bit basis to determine whether they are identical. See Microsoft Help and Support, How To Use the Windiff.exe Utility, <http://support.microsoft.com/kb/159214> (last visited Dec. 6, 2008) (describing how to use the utility to compare two files).

97. National Institute of Standards and Technology, Cryptograph Hash Project, <http://www.csrc.nist.gov/groups/ST/hash/index.html> (last visited Dec. 6, 2008).

98. *Id.*

99. JAAP HAITSMA ET AL., ROBUST AUDIO HASHING FOR CONTENT IDENTIFICATION 1 (2001), available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.16.2893&rep=rep1&type=pdf>.

100. *Id.*

101. *Id.*

bit of data in the original input can change the hash output dramatically.¹⁰² Hash functions are thus uniquely suited to identify content and determine whether a file is what it purports to be, regardless of what the file name says it should be.

Traditional hash functions, such as MD5, are not affected by what format the source and comparison files are in; they simply examine the binary structure of the file and output a hash based on that structure.¹⁰³ The problem with using this kind of system in the P2P world is that the media files that are being shared are often delivered in any number of different formats and qualities, and even a minor change in the ID3 tag information would fundamentally alter a basic hash.¹⁰⁴ One way to solve this problem would be to implement a more robust method of determining whether a file is a copyrighted work by looking at its content more generally. In the same way that a human is able to distinguish between a sound recording by his or her favorite band and a sound recording of a busy street, even if one of those recordings is somewhat muffled or of lower quality, sophisticated software could also look at the underlying content of an audio file, rather than looking at the bit-for-bit digital representation of that content.¹⁰⁵ When it comes to audio applications alone, this concept is known as the robust audio hash.

A robust audio hash is a function that associates to every basic time-unit of audio content a short semi-unique bit-sequence that is continuous with respect to content similarity as perceived by the [human auditory system]. In other words, if the [human auditory system] identifies two audio signals as being very similar, the associated hash values should also be very similar, i.e. the bit patterns of the respective hash values should be similar (but not necessarily identical). In particular, if we compute hash values for original content and for an MP3 compressed version of that content, the hash values should be similar. On the other hand, if two signals really represent different content, the robust hash should be able to distinguish the two signals (semi-unique). This is similar to the collision requirement for classical cryptographic hashes.¹⁰⁶

Put in layman's terms, the robust audio hash examines a file and “listens” to the sound it would produce if it were played. It then creates a unique identifier for that sound signature. Another file that contained the same musical work would yield a very similar “signature,” regardless of how it is

102. *Id.*

103. Hans Dobbertin, *The Status of MD5 After a Recent Attack*, 2.2 CRYPTOBYTES 1, Summer 1996, available at <ftp://ftp.rsasecurity.com/pub/cryptobytes/crypto2n2.pdf>.

104. HAITSMAN ET AL., *supra* note 99, at 1.

105. Logan Decker, *Do Higher MP3 Bit Rates Pay Off?*, MAXIMUM PC, Apr. 19, 2007, http://www.maximumpc.com/article/do_higher_mp3_bit_rates_pay_off (examining the impact of different bit rates on the subjective quality of music). The study showed that the different bit rates often had little impact on the perceived quality of the music, which implies that the different bit rates did not affect the ability of the listeners to determine what the music was. *Id.* Because each file was encoded differently, however, and since even the smallest change to a digital file will result in a different hash, it is readily presumed that the different files would have different hashes. *Id.*

106. HAITSMAN ET AL., *supra* note 99, at 2.

encoded or what other information is stored in the music file itself.¹⁰⁷ By looking at these hashes, one could determine without actually downloading the work and without actually listening to it whether it was a copyrighted work or not.¹⁰⁸

The problem with this application, however, is that it must be implemented as a part of the sharing software. One could not calculate the hashes just by looking at the data reported by the file-sharing program in a search; the hash must be calculated for each file when it is added to the shared library.¹⁰⁹ The search engine would then have to return this hash value to the searcher, and then the searcher could see if the work was copyrighted (if they had access to the original copyrighted work and had created their own hash). Also, the robust audio hash applies only to audio. Other copyrighted works traded on P2P networks, such as video, would require a different kind of hash, one that is far more complex because of the far larger amount of data in a video.¹¹⁰ Still, it is possible that future implementations of P2P software could use robust audio hashes to provide users with a greater guarantee that they are downloading the file that they actually want.

In fact, the simpler form of hash creation is already implemented in many P2P programs that allow for a user to download a file from multiple other users at once.¹¹¹ In such implementations, the P2P software on the computer of the user who is sharing the file calculates a hash based purely on the digital content of the file and reports this hash in search results.¹¹² The calculation necessarily relies only on the digital content and not on the auditory “signature” because for the searching user to end up with one complete file, all the various source files must be identical. If this were not the case, the resultant downloaded file would be an amalgam of different quality source files and would look or sound incorrect. Because a slight variance in the source files (such as in the ID3 tag or in the file’s bitrate) would yield a different hash, this method is not reliable for discovering on a large scale

107. *Id.*

108. This type of technology has already been successfully implemented to allow users with a large music library to properly identify and catalog their songs based on the sound signatures of the files that they possess. See About Shazam, <http://www.shazam.com/music/web/pages/about.html> (last visited Dec. 6, 2008). Some mobile phone providers allow a user to record a song that they hear on the radio or somewhere else and upload a portion of it to their servers. *Verizon Wireless Offers V Cast Song ID*, MOBILE TECH NEWS, May, 22, 2007, <http://www.mobiletechnews.com/info/2007/05/22/071546.html>. An analysis of the sound signature is then made, and the user is provided with the artist and song title, and the opportunity to purchase the full version of the song. *Id.*

109. One must have the entire file available in order to create a hash, and without downloading the file, one would not have the entire file.

110. But see Google Video Upload Program FAQ, https://upload.video.google.com/video_faq.html (last visited Dec. 6, 2008). (“We do a preliminary review on uploaded videos through both a manual and automated process. Although we try our best to detect and remove videos that violate our policy guidelines, our review process is not bulletproof.”). Nevertheless, Google is able to only conduct this process once the entire video has been uploaded, and this technology would have to be implemented on the “server” side of a P2P transaction.

111. Christopher Rohrs, *LimeWire Download Code*, LIMEWIRE, Nov. 13, 2002, http://wiki.limewire.org/index.php?title=Download_Design (last visited Dec. 6, 2008); see also MAGNET-URI Project, <http://magnet-uri.sourceforge.net/> (last visited Dec. 6, 2008) (providing an example of a simpler form of creation used in file-sharing programs that use the Gnutella network).

112. Rohrs, *supra* note 111.

whether a file is a copyrighted work or not, since those wishing to conceal copyrighted works could make a small change to the file and be immune from discovery.

However, as mentioned earlier, one of the great advantages of P2P file-sharing is that the content can be distributed among many users for the purposes of uploading it to new users.¹¹³ For this advantage to be realized, the exact same content must be spread across multiple users. In other words, for P2P to work up to its potential, some number of users must have the exact, same file. This means that the copyrighted files at issue in P2P file-sharing cases might be files that are distributed widely across the network. An alleged infringer may have obtained those files specifically because they were especially prevalent on the network. If a user were to search for a specific song, for example, the user may be presented with a number of different versions of the same song, each varying slightly in their bitrate and other information that is inconsequential for his or her purpose, but that would generate a different hash. A user looking to download a file as quickly as possible would want to download from as many peers as possible at the same time, in order to take advantage of the greatest aggregate bandwidth, and that means choosing the version of the file that is most commonly available on the network.¹¹⁴

What this means for plaintiffs is that if they, too, can obtain the most commonly available version of a file on the P2P network, and if they can generate a hash in the same way that the P2P software does to make its matches, they may theoretically be able to tell whether a file is the copyrighted work in question without downloading it, based solely upon the search results. This would eliminate the issue of whether a file being shared by a user is actually in violation of a copyright, and it would eliminate the proof problem that might occur if the sharing user’s files were mislabeled. By itself, however, this solution does not address the larger question of whether making a file available should be considered illegal distribution.

D. Partial Transfers of Files Should Not Be Treated as the Same Violation as Downloading of Complete Songs

As discussed earlier, part of the rationale behind a P2P system is the ability to download the same file from multiple peer sources at the same time, speeding the download.¹¹⁵ In such a situation, no single sharing user provides a complete copy of the file to the one user who is downloading a file. In addition, the various portions of the file that individual users provide may be

113. See discussion *supra* Part II.A (describing how P2P technology provides a way to share files between peers).

114. It stands to reason that if the user had a larger maximum download capacity than the upload capacity of the user they were downloading from, they would still have “excess” bandwidth that they could devote to downloading from another user. For a more detailed discussion of this process and empirical statistics on how it affects downloading speed, see KWOK & CHAN, *supra* note 93, at § 6.2.

115. See discussion *supra* Part II.A.

useless without the rest of the file.¹¹⁶ Even if they are discernible and usable as parts of an original copyrighted work, some point to the de minimis defense available in some jurisdictions, where a very small portion of a work that is made available is not found to violate copyright.¹¹⁷

The deemed distribution doctrine also does not take into account the capabilities of a sender, which relates to the amount of a copyrighted file that the sender is capable of actually distributing. The doctrine would allow a plaintiff to produce a list of search results of copyrighted files that could be downloaded from the users' computer, but even assuming that none of the hurdles described earlier exist¹¹⁸ and that those files could be downloaded from the user, there is no guarantee that the user has the bandwidth to realistically upload all of those files, or that the user's computer will be available on the network long enough to offer a complete work.¹¹⁹

For example, a user using the fastest possible dial-up connection (still a popular form of connection in the United States)¹²⁰ would need to spend over fourteen minutes online in order to upload an MP3 file, and over thirty-three hours online to upload a movie.¹²¹ These figures do not contemplate the user creating any other upstream traffic other than the upload of files, one at a time, and without any disconnections or other network traffic issues. It is difficult to see how providing a court with a list of files being shared by one person on a dial-up connection could translate into a reasonable offer to distribute even a small portion of those files. The point here is to illustrate the absurdity of the deemed distribution doctrine because it requires a defendant to prove that any alleged copyright violations could not have reasonably occurred.

116. The algorithms that different P2P software packages use to split a file into different downloadable parts vary, but generally, for a file to be playable by an average end user, it would have to be completed and reassembled by the P2P software. *But see* Softpedia.com, Download GOM Player 2.1.9.3752 Free Trial, <http://www.softpedia.com/get/Multimedia/Video/Video-Players/Gom-Player.shtml> (last visited Dec. 6, 2008) (describing a software media player that purports to be able to play partially downloaded media files downloaded from P2P networks).

117. 26 AM. JUR. PROOF OF FACTS 3D § 537 (2008).

118. *See* discussion *supra* Part III.A.2 (describing potential hurdles to downloading files found in a search).

119. Recently, some providers have been limiting the amount of bandwidth available on their plans, imposing caps once a certain level of activity has been reached. Eric Bangeman, *Comcast Speaks out on Bandwidth Caps, Says They Only Affect 0.01% of Users*, ARS TECHNICA, Sep. 19, 2007, <http://arstechnica.com/news.ars/post/20070919-comcast-speaks-out-on-bandwidth-caps-says-they-only-affect-0-01-of-users.html>.

120. C. Spencer Beggs, *Reports of Death of Dial-up Internet Greatly Exaggerated*, FOX NEWS.COM, June 27, 2006, <http://www.foxnews.com/story/0,2933,200606,00.html> (reporting that 34% of Americans who go online connect through dial up services).

121. The fastest possible upstream speed would occur if the user used the V.92 protocol; the user could then upload their data at 48Kbps. *An Innovation Called V.92*, EXPRESS COMPUTER, <http://www.expresscomputeronline.com/20030127/peripheral10.shtml> (last visited Dec. 6, 2008). Dividing this number by eight yields a maximum speed of 6KBps. Marshall Brain, *How Bits and Bytes Work*, HOW STUFF WORKS, <http://computer.howstuffworks.com/bytes.htm> (last visited Dec. 6, 2008). Since there are 1,024 kilobytes in a megabyte, it would take 1024/6, or 2.84 minutes, to upload one megabyte of data. *Id.* If a song is about five megabytes, it would take 5 x 2.84, or 14.2 minutes to upload it. If a movie is 700 megabytes, it would take 700 x 2.84, or 33.3 hours, to upload it.

IV. RECOMMENDATION

The deemed distribution doctrine should not be applied in file-sharing cases. Courts should hold the record industry plaintiffs to a higher evidentiary standard, requiring them to prove that their exclusive right to distribute has been violated, instead of allowing them to prove that such a violation could occur.

This standard could be met by requiring a plaintiff to download the files in question, rather than providing a screenshot showing what could be downloaded. Requiring an actual download would eliminate the need for more technical file comparison methods or the creation of hashes: one could simply listen to or watch the file, and if it were the copyrighted work in question, then it would be very clear that the work had actually been distributed. Even this may not be enough proof of distribution, however; a defendant that is named in a suit may not have been the exclusive user of the computer, or of the network that is identified in a lawsuit.¹²² Moreover, the electronic nature of file transfers and the aforementioned uniformity of files available on P2P networks might raise questions as to what kind of documentation is required to show that a plaintiff actually downloaded the file from a defendant.¹²³

The deemed distribution doctrine as it was developed in *Hotaling* is not flawed per se; it can still function as proof of a copyright violation where there is a greater likelihood that the works being cataloged actually do infringe on copyright, such as in libraries or for Web sites that charge money for works that they do not have the right to distribute. There, the expectation that the works actually do exist cannot be overcome by the myriad issues with applying the doctrine to the P2P world. In fact, the doctrine seems especially suited to situations like the one in *Hotaling*, where there is little question that a copyright violation was probable based on the library context and the information in the catalog. Unfortunately for the plaintiffs in these cases, this rule simply does not translate well to the uncertainties of the online world.

Plaintiff copyright holders do have one viable option for legitimately applying the deemed distribution doctrine to the P2P world, but strangely enough, it relies on the capabilities of the software that enables the major efficiencies of P2P file-sharing in the first place. If plaintiffs were able to successfully match the hashes created for a certain copyrighted work that was widely available on P2P networks, they could rather definitively prove that all the individuals offering that exact file for download were actually offering a copyrighted work, and not one that had been incorrectly labeled or that did not represent the work at all. This is the best hope for plaintiffs who wish to use deemed distribution to prove the existence of the copyrighted work in a shared directory, but it still does not address the problems that the works have not been transmitted or communicated, and that they may in fact be unavailable for transmission or communication due to some technical barrier.¹²⁴

122. *Atl. Recording Corp. v. Howell (Howell II)*, 554 F.Supp. 2d 976, 986 (D. Ariz. 2008).

123. See discussion *supra* Part II.A.

124. See discussion *supra* Part II.A (noting possible hurdles to downloading files found in a search).

In the interim, before deemed distribution is cast aside as a valid doctrine in the P2P file-sharing world, defendants ought to consider the varied defenses available to them in such a case. If they can show that the files were not accessible to download because of their presence behind a firewall, they can show that distribution was not contemplated. If defendants can show that plaintiffs could not have obtained the entire work from them because of their bandwidth limitations, whether established in their P2P software or because of the physical limitations of their connection, they may be able to show that any possible distribution would be of a minimal component of the work. If they could show in some way that the files that appeared to be copyrighted files were actually mislabeled, or were perhaps some parody created by the sharing user, they may show that any possible distribution did not involve copyrighted works. Perhaps better than anything else, these defenses illustrate the problems with deemed distribution as a way of proving copyright violation.

V. CONCLUSION

Courts should require more evidence that actual distribution occurred before imposing liability on consumers in P2P file-sharing cases. Logic like that used in *Barker* would make it too easy to impose liability on individuals who may have done nothing wrong. The court in *Howell* put it best: putting a file in a shared folder “only shows that the defendant attempted to distribute the copy, and there is no basis for attempt liability in the [copyright] statute, no matter how desirable such liability may be as a matter of policy.”¹²⁵

125. *Howell II*, 554 F.Supp.2d at 984.