Practical advice to beat piracy

The internet presents seemingly endless opportunities for bad actors to pirate and infringe copyrights and trademarks. **Jeff Parmet**, **Tom Ashley** and **Nick Ferrara** explain what investigative techniques can be deployed

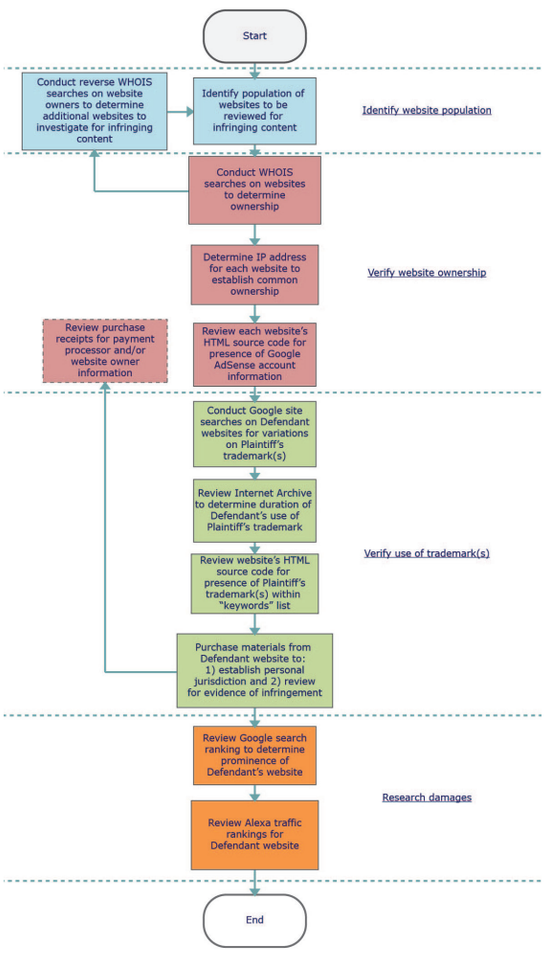
Online piracy is an increasingly important issue facing US policy makers.[1] But it is especially daunting for US copyright and trademark holders, the attorneys who represent them in combating piracy, and the forensic experts who provide support. This article offers a brief overview of some of the techniques we have successfully used to lay the necessary evidentiary foundation to support misappropriation and infringement claims. Forensic experts face a number of challenging requirements: they must collect and preserve allegedly infringing materials, identify the owners and operators of the websites distributing infringing materials, and evaluate the extent of infringement, so that the amount of damages can be determined. Fortunately, a variety of tools, reference sources, and techniques exist that facilitate tracking down and prosecuting those individuals or organisations engaged in online copyright or trademark piracy.

While much has been written about TV and film piracy on illegal file-sharing sites, this article focuses on the more mundane theft of trademarks and other online content. The following hypothetical scenario illustrates an increasingly common example of online piracy. We will use it in the balance of this article to describe the steps skilled forensic examiners take to collect and preserve evidence.

Hypothetical piracy scenario

Your client sells educational course material on its website under the trademark "AlwaysGetAnA". The courses have been developed over a period of a dozen years and contain thousands of pages of original content. The defendant registers for your client's courses, pays the tuition, and downloads course materials. However, instead of taking the courses, the defendant takes the materials, translates them into several foreign languages, sets up dozens of websites around the world, and offers the translated materials for sale under the trademark "NothingButA". Your client has identified some of these websites and wants to put a stop to this illegal behaviour. Further, your client wants to sue for damages. Your challenge is to gather the evidence needed to obtain an injunction and to quantify the extent of harm to your client.

Figure 1 – A procedure for investigating online piracy



Some of the forensic methods we use to support injunction efforts, and to quantify the harm to the client, are set forth in Figure 1 (p43). Our work is typically delivered in the form of a declaration for use by counsel in support of an injunction motion.

Identify initial population of websites to investigate

Before an online piracy investigation can proceed, the plaintiff should provide the forensic expert with an initial population of websites to "seed" the investigation. This initial population can be derived through basic online research methods such as conducting Google searches on variations of the plaintiff's trademark. Frequently, the trademark will appear directly in a website's URL. For example, under the hypothetical scenario discussed previously, websites with URLs such as www.nothingbuta.com, www.nothingbuta.net, and www.nothing-but-a.com are likely candidates for trademark infringement. As discussed further, the expert should be aware that this initial website population may expand as the investigation proceeds.

Determine whether each suspected domain is under the defendant's control

A forensic expert's typical starting point is to determine whether the defendant controls each website containing allegedly misappropriated materials. One useful technique is to use the WHOIS service to determine whether a given URL or "domain" is under a party's control. WHOIS is a service used to look up the registration information on file for a given domain. WHOIS information can be obtained through websites such as whois.domaintools.com or by using software such as the Linux utility "whois". Registration information typically includes several contacts, such as "Administrator", "Registrant", and "Billing", all of which we have relied on in prior infringement investigations to draw conclusions regarding domain ownership.

"In determining whether websites share common ownership, a forensic expert will also rely upon Google AdSense account information."

Identify the population of domains under defendant's control

Additionally, forensic experts often use "reverse WHOIS" searches to find other domains having similar registration. A reverse WHOIS lookup is a type of search that can find all domains registered to a common email address. In infringement investigations, we use an online "Reverse Whois Lookup" tool to conduct such lookups.[2] In particular, we have utilised the results of conventional WHOIS lookups, described above, to identify registration information to be included as input to Reverse WHOIS Lookup searches. Continuing with our hypothetical scenario, assume that a WHOIS lookup on the "NothingButAnA.com" website indicates that the registrant email is nothingbutana@gmail.com. The Reverse WHOIS Lookup tool may indicate that the following domains are associated with this email address: nothingbutana.net, nothingbutana.com, nothingbuta.net, nothingbuta.com, nothingbutas.net, and nothingbutas.com. The forensic expert may then infer that such domains likely share common ownership, and should be reviewed for possible evidence of infringement.

Identify IP addresses associated with each such website and establish multiple domains

Another useful technique is to determine whether an infringing website is part of a broader "ring" of websites owned by the same entity that are all used to distribute material that may be copyrighted or trademarked. In making such a determination, a forensic expert reviews the internet protocol ("IP") addresses associated with a defendant's websites. An IP address is a unique string of numbers separated by periods or colons that is assigned to computers and other digital devices and allows them to communicate with one another via the internet (eg, "74.125.22.100" and "2607:f8b0:400d:c06::66" are IP addresses for Google). Experts in the field of information technology routinely rely upon IP address information to draw conclusions regarding domain ownership. Due to a process known as "shared web hosting service", it is possible for multiple domain names to share an identical IP address. Because it is unlikely that multiple unrelated infringers would independently decide to use the same shared web hosting service to distribute allegedly infringing material, the use of shared web hosting service for multiple domains that are similar in nature is strong circumstantial evidence that those domains are owned and operated in common.

Identify websites under defendant's control by using AdSense account numbers

In determining whether websites share common ownership, a forensic expert will also rely upon Google AdSense account information. Google offers a program called "AdSense" by which a website may display Google advertisements and receive remuneration based on the number of users that view or click on the advertisements. Google identifies the entity to receive remuneration by an AdSense account number that uniquely identifies the AdSense account. The AdSense account number is included in the HTML source code of a webpage. As the account number identifies the entity to receive remuneration for publishing the advertisements, two websites using the same AdSense account number can be reasonably expected to share common ownership.

Collect information on the use of client marks in the content of each website

After researching and drawing conclusions regarding the ownership of potentially infringing websites, the forensic expert should thoroughly review each website for infringing content. We have used Google's "site search" function to collect information on the use of a plaintiff's intellectual property in the content of each website. A Google site search is performed by running a search that includes the phrase "site:domain.com", which will cause Google to search only a given website. For example, a Google search for "site:nothingbutana.net NothingButAnA" will return the number of uses of the word "NothingButAnA" solely within the nothingbutana.net domain. Additionally, Google search includes the "OR" operator to simultaneously search for multiple terms. For example, a Google search for "site:nothingbutana.net NothingButAnA OR NothingButA" will return search results for either "NothingButAnA" or "NothingButA" solely within the nothingbutana.net website.

Gather historical information pertaining to the duration and extent of misappropriated content and marks

To determine the duration of a defendant's misappropriation of a plaintiff's content, which may directly impact damages, a forensic expert will often search the internet archive's "wayback machine" to gather historical information pertaining to a defendant's websites.[3] The Internet Archive is a free online library of historical collections that exist in digital format, and it includes billions of website captures as they existed at previous points in time. The wayback machine is an application provided by the Internet Archive for searching its archive of historical websites. While the Internet Archive is neither complete nor perfect, it is voluminous, accurate, and generally reliable. In our prior investigations, we have used the wayback machine primarily to determine the earliest known date on which each defendant website

began displaying our clients' marks, and also to ascertain the presence of our clients' marks on websites that either are no longer accessible at a given domain or otherwise no longer contain substantive content. In such cases, we have made screenshots of archived pages displaying our clients' content so as to preserve each such webpage as it appeared on the date it was added to the internet archive.

Examine HTML metadata to determine whether defendants are using plaintiff's content to direct traffic to their websites

To determine whether a defendant is utilising a plaintiff's trademark to attract internet traffic, forensic experts typically review the HTML metadata in a defendant's website to determine whether variants on the plaintiff's trademark are included as keywords. Keywords are included in websites to increase the likelihood that the website will appear in search engine results for terms containing those keywords. Thus, the presence of variants of a plaintiff's trademark in a defendant website's keywords list may increase the likelihood that users will be directed to the website when searching for the plaintiff's products. For example, a defendant website's use of variants of the "AlwaysGetAnA" mark in its HTML keyword list may increase the probability that users will find the defendant's website when looking for the plaintiff's materials. Moreover, the presence of a plaintiff's trademark in a defendant's HTML metadata provides clear evidence of the defendant's intent to trade on the plaintiff's mark.

Purchase materials from defendant's website to establish personal jurisdiction and review for evidence of infringement

Before a court can consider the substantive merits of a trademark infringement action, the plaintiff must establish that the court has personal jurisdiction over the defendant. In the US, personal jurisdiction requires the existence of "sufficient minimal contacts" between the defendant and the jurisdiction in which the plaintiff litigates the dispute.[4] To demonstrate such contacts, it is often useful for the forensic expert to purchase infringing content from a location within the court's geographical jurisdiction. Additionally, such purchases allow the expert to review the purchased content for infringing material and to review the purchase receipts for payment processor information. A payment processor is a company appointed by a merchant, whether online or brick and mortar, to handle credit or debit card payments on the merchant's behalf. In our prior investigations, we purchased products from the websites we investigated and reviewed the resulting payment receipts to identify the payment processors used by the operators of such websites. In some cases, the payment receipts also indicated the entity or organisation selling products through the associated website.

Evaluate defendant websites using Google search rankings to gather evidence of a website's prominence to consumers on the internet

To determine a website's relative prominence to internet consumers, which can help establish harm to the IP owner, a forensic expert examines a website's Google search rankings. Google ranks each webpage among search results based on various criteria including keywords in the page and the rankings of other pages that link to that page. As such, a website's ranking in Google's search results relative to a specific search term can be increased by using that search term on the website frequently, and by creating a number of other secondary websites that host content related to the search term and are linked to that website. For example, a website's increased use of a term such as "NothingButAnA" will signal Google to list the website more prominently when users conduct Google searches on "NothingButAnA". More prominent websites tend to get more traffic than less prominent ones, and more traffic usually translates to more sales.

Determine the popularity of defendant's websites

To establish traffic to a website, which can assist a damages expert in quantifying damages, the forensic expert will use tools such as Alexa's "Traffic Rank" and "Site Comparison" features to determine the relative popularity of websites containing infringing content.[5] Alexa traffic rankings provide "a rough estimate of [a] site's popularity" and are "calculated using a combination of average daily visitors to [a] site and page views on this site over the past [three] months."[6] In this context, a website with a lower ranking is associated with a larger number of daily visitors and page views. Alexa's "Site Comparison" tool provides a graphical depiction of traffic rankings over time for multiple websites. Although we have found that Alexa's tools provide an estimate of website traffic and sometimes vary from the actual traffic a given website is experiencing, they provide a reasonable basis for ascertaining a website's relative degree of popularity with users. Thus, in our prior infringement investigations we have captured screenshots of Alexa traffic ranking results to depict increases in relative popularity of websites containing protected content. While these numbers produce reliable estimates, hard statistics will always be preferable. To that end, actual traffic to a website can be quantified by subpoenaing the logs of the servers hosting the infringing content, while actual sales can be quantified by subpoenaing payment processor records.

Summary

The internet presents seemingly endless opportunities for bad actors to pirate and infringe copyrights and trademarks, creating significant challenges for copyright and trademark owners seeking to protect and enforce their exclusive IP rights. To combat this behaviour, investigative techniques such as those discussed in this article, have assisted us in our role as forensic IP experts, not only in identifying and preserving allegedly infringing material, and identifying bad actors, but also in laying the proper evidentiary foundation necessary to prevail on and quantify a claim for misappropriation, and ultimately, infringement.

Footnotes

1. The Commission on the Theft of American Intellectual Property, 'The IP Commission Report', at 4 (2013) (finding that the US loses more than $300bn a year in revenue due to IP theft, millions of jobs are lost, innovation is stifled, and GDP is lessened); Id at 24 (noting that the shadow market for software alone increased by 9.2% between 2010 and 2011, with China's piracy rate an astonishing 77% and its software investment a mere 7%).
2. Reverse Whois Lookup, available at: http://viewdns.info/reversewhois/
3. Internet Archive wayback machine, available at: https://archive.org/
4. See *International Shoe Co v Washington*, 326 US 310, 316 (1945).
5. Alexa 'Site comparison' website, available at: www.alexa.com/comparison
6. Alexa 'Site overviews' website, available at: www.alexa.com/siteinfo

Authors



Jeff Parmet (left) is a consulting and testifying expert and the founder and managing partner of DisputeSoft.

Tom Ashley (middle) is an IP software expert and Nick Ferrara (right) is a software expert also at DisputeSoft.