

Re: [Full-disclosure] Re: Case ID 51560370 – Notice of ClaimedInfringement

Source: <http://www.derkeiler.com/Mailing-Lists/Full-Disclosure/2005-04/0194.html>

From: Jason (security_at_brvenik.com)

Date: 04/08/05

Date: Fri, 08 Apr 2005 13:45:51 -0400

To: Valdis.Kletnieks@vt.edu

Valdis.Kletnieks@vt.edu wrote:

> *On Fri, 08 Apr 2005 12:50:24 EDT, Jason said:*
>
>
>>*I think that entirely depends on the format the file is distributed in.*
>>*You could take a zipfile and pad it in non critical areas to change the*
>>*MD5 without creating a substantial difference in the deliverable*
>>*content. You could do the same with gzip or bzip formatted files. You*
>>*could also pad any embedded jpeg images to engineer a collision. There*
>>*are quite a few opportunities where this method could be used to twiddle*
>>*the new MD5 without materially changing the content.*
>
>
> *It's easy to tweak a file and get a different MD5. That's why Tripwire works.*
>
>
>>*Software that is ~150M in size, it gets redistributed as a new file that*
>>*is 160M in size but has a collision with your software which is also*
>>*160M in size. I imagine there would be some computational time involved*
>>*to find the appropriate collision but a lot less computational time than*
>>*finding a perfect match to the original.*
>
>
> *You're missing the point.*
>
> *Let's say we have a file A that's 150M in size, and a file B that's 160M in*
> *size. File B is *not* under our control, and has a known fixed MD5 hash.*
>
> *It's easy to take file A, and create 2 files C and D from it that happen to*
> *have the same MD5 hash as each other. What is *NOT* easy is creating a file E*
> *that has the same hash as A or B.*

I get the point just fine. Injecting files C and D results in a situation that cannot be resolved without downloading both files.

Song A = mp3 format file with valid license to BSA
Song B = mp3 format file without valid license to BSA
Song C = zip of Song A plus pad to generate MD5
Song D = zip of Song B plus pad to generate same MD5

It is now impossible to distinguish between C and D without downloading both. The content inside is still fully usable and valid but a violation cannot be confirmed without yourself violating the law.

What you might see in a DL dialog

NAME MD5 SIZE
somefile.zip ABCD321312 120M
someotherfile.zip ABCD321312 120M

You cannot remotely know that either file is in fact the content you are looking for without downloading both files. Both files may not be the content you are looking for. How can you remotely distinguish that a violation has occurred?

Full-Disclosure – We believe in it.
Charter: <http://lists.grok.org.uk/full-disclosure-charter.html>
Hosted and sponsored by Secunia – <http://secunia.com/>