12 A cryptographer and a binarista walk into a bar

by Ange Albertini, Binarista
and Maria Eichlseder, Cryptographer

So you meet a stingy schizophrenic genie, who grants you just one wish, and that wish is a single hash collision, with a bunch of nasty restrictions. In the following story, cleverness wins over stinginess, as it does, in a classic fairy-tale way! —PML

SHA-1 uses four constants internally. `0x5a827999`, `0x6ed9eba1`, `0x8f1bbcd` and `0xca62c1d6` are the square roots of 2, 3, 5, and 10 respectively. These nothing-up-my-sleeve numbers are supposedly innocent, but nobody knows why they were chosen, rather than any other constants. It's a common practice in embedded devices to use known checksum algorithms such as SHA-1 but with different internal parameters: it gives you a proprietary algorithm based on a robust model.

What could go wrong?

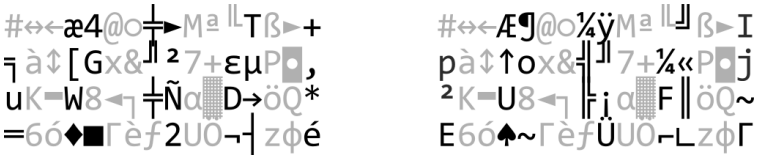
Aumasson et al.[23] show how to find practical collisions for such modified SHA-1 when the attacker can control these constants.

From a high-level perspective, finding a collision pair is a bit of an involved process. It roughly involves the following, but you should read the paper for full details.

1. Feeding the difference pattern (explained below) and the fixed bits (w.r.t. to the pattern) to an optimized automatic search algorithm.
2. Experimenting with the parameters until a few reasonable-looking candidates emerge, aborting if none do.
3. Feeding those candidates to a similar search algorithm with a similar parameter set.
4. Waiting a day or two for completion, maybe eliminating the less promising candidates successively.

Let's consider the consequences from a non-cryptographic perspective.

You have a colliding pair of pseudo-random blocks. They took between fifteen and thirty hours to compute, on eighty cores. They have the same SHA-1 checksum (`e033efe8e6e74d75c6d0bbaf2f2eba8d163f70b5`) if the internal constants are `0x5a827999`, `0x88e8ea68`, `0x578059de`, `0x54324a39` instead of the original ones. You're happy, you win.



If you look at these blocks as a normal person, you probably think, "This is just colliding random garbage. Big deal!" They just don't seem that scary. It would be far more useful if you had colliding files using a standard binary format.

Here are the rules of the game, from the binary perspective.

- You have two different blocks of `0x40` bytes, at offset `0`, that yield colliding hashes. You can append the same content to both, of course, and the overall hashes would still collide.
- Certain positions in these blocks are occupied by the same bytes, while bytes in other positions differ. We call the bitwise pattern of the differences a *difference pattern* and call the bytes/bits that must be the same in both blocks *fixed* and the rest *"random"*. Only a handful of such patterns exist that still have practical attack complexity.

[23] Albertini A., Aumasson J.-Ph., Eichlseder M., Mendel F., Schlaeffer M. Malicious Hashing: Eve's Variant of SHA-1. In: Joux, A. (ed.) Selected Areas in Cryptography 2014, LNCS, Springer (to appear)

- All available patterns have at most three consecutive bytes without a difference. Typically, in every double word, only the middle two bytes have no differences.
- A few more bits can be set to fixed values on top of a difference pattern, but the majority of the remaining bits will need to be "random". Typically, the more bits you fix, the higher the computational attack complexity. Fixing between 32 and 48 of the 512 bits in the first block usually works fine.
- All available patterns have a difference in the higher nybble of the last byte, and one pattern has no difference in the first three bytes.

This means that you can't have a magic signature of four bytes in a row in both blocks, nor four `00` bytes in a row, so you already know that you can't have two files of the same type with a classic four-byte magic value at offset zero.

You must either somehow skip over the randomness or deal with it. We will now discuss various ways to do so.

12.1 Skipping over the Randomness

Shell Scripts

You can see that our two blocks start with a hash and contain no carriage-return characters. That pattern is treated as a comment in many scripting languages, and thus ignored as unneeded data. Appended to two differing but colliding comment blocks, the same scripting code could check for some difference and produce different results accordingly. This will result in two colliding scripts.

```
0000000: 231d 1b91 3440 09d8 104d a6d3 54e1 102b  #...4@...M..T..+
0000010: b885 125b 4778 26bd fd37 2bee e650 082c  ...[Gx&..7+..P.,
0000020: 754b 1657 3811 bfd8 a5e0 b244 1a94 512a  uK.W8......D..Q*
0000030: cd36 a204 fee2 8a9f 3255 99aa b47a ed82  .6......2U...z..
0000040: 0a0a 6966 205b 2060 6f64 202d 7420 7831  ..if [ `od -t x1
0000050: 202d 6a33 202d 4e31 202d 416e 2022 247b   -j3 -N1 -An "${
0000060: 307d 2260 202d 6571 2022 3931 2220 5d3b  0}"` -eq "91" ];
0000070: 2074 6865 6e20 0a20 2065 6368 6f20 2220   then .  echo "
0000080: 2020 2020 2020 2020 285f 5f29 5c6e 2020          (__)\n
0000090: 2020 2020 2020 2028 6f6f 295c 6e20 202f         (oo)\n  /
00000a0: 2d2d 2d2d 2d2d 2d5c 5c2f 5c6e 202f 207c  -------\\/\n / |
00000b0: 2020 2020 207c 7c5c 6e2a 2020 7c7c 2d2d       ||\n*  ||--
00000c0: 2d2d 7c7c 5c6e 2020 205e 5e20 2020 205e  --||\n   ^^    ^
00000d0: 5e22 3b0a 656c 7365 0a20 2065 6368 6f20  ^";.else.  echo
00000e0: 2248 656c 6c6f 2057 6f72 6c64 2e22 3b0a  "Hello World.";.
00000f0: 6669 0a                                  fi.
```

```
$ sh eve1.sh
        (__)
        (oo)
  /-------\/
 / |     ||
*  ||----||
   ^^    ^^
```

```
0000000: 231d 1b92 1440 09ac 984d a6d3 bce1 1049  #....@...M.....I
0000010: 7085 1218 6f78 26b9 bd37 2bac ae50 086a  p...ox&..7+..P.j
0000020: fd4b 1655 3811 bfcc ade0 b246 ba94 517e  .K.U8......F..Q~
0000030: 4536 a206 7ee2 8a9f 9a55 99a9 1c7a ede2  E6..~....U...z..
0000040: 0a0a 6966 205b 2060 6f64 202d 7420 7831  ..if [ `od -t x1
0000050: 202d 6a33 202d 4e31 202d 416e 2022 247b   -j3 -N1 -An "${
0000060: 307d 2260 202d 6571 2022 3931 2220 5d3b  0}"` -eq "91" ];
0000070: 2074 6865 6e20 0a20 2065 6368 6f20 2220   then .  echo "
0000080: 2020 2020 2020 2020 285f 5f29 5c6e 2020          (__)\n
0000090: 2020 2020 2020 2028 6f6f 295c 6e20 202f         (oo)\n  /
00000a0: 2d2d 2d2d 2d2d 2d5c 5c2f 5c6e 202f 207c  -------\\/\n / |
00000b0: 2020 2020 207c 7c5c 6e2a 2020 7c7c 2d2d       ||\n*  ||--
00000c0: 2d2d 7c7c 5c6e 2020 205e 5e20 2020 205e  --||\n   ^^    ^
00000d0: 5e22 3b0a 656c 7365 0a20 2065 6368 6f20  ^";.else.  echo
00000e0: 2248 656c 6c6f 2057 6f72 6c64 2e22 3b0a  "Hello World.";.
00000f0: 6669 0a                                  fi.
```

```
$ sh eve2.sh
Hello World.
```

MBR & COM

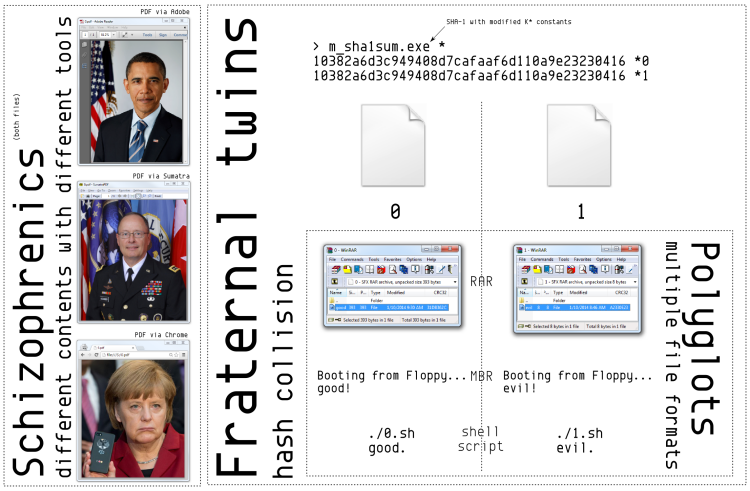
Another possibility is to use one of the header-less file formats, such as an MBR boot sector or a COM executable. Encode some jumps in the constant part, with the relative offset in the differing part. Execution will land in different offsets, where you can have two different stubs of code.

7 Zip & Rar

Archives that are parsed sequentially, such as 7 Zip and Rar, simply scan for their respective signatures at any offset. So to create an archive collision, simply concatenate two archives and remove the first byte of the top archive. Then you have to make sure that one block of the colliding pair ends with the missing byte

of the signature. This block will restore the signature of the top archive, whereas the other block will keep it disabled, thus enabling the bottom archive.

Note that these are not exclusive. With a bit of perseverance, you can have a Rar-MBR-Shell colliding polyglot. And append a schizophrenic PDF, too! Why not? ;)



12.2 Dealing with Randomness

A JPEG file is made of segments. Each segment is defined by its first two bytes: first `0xff`, then an extra marker byte (but never `0x00`). For example, a JPEG should start with a Start-of-Image segment, marked `0xff 0xd8`.

Most segments then encode a length on two bytes (which is handy because it won't get out of control if it's random), and then the content of the segment.

A weird property of the JPEG format is that even though these markers are either constant-sized or encode their length, you can still insert random data between two segments.

How does the parser know where a new segment starts? It looks for an `0xff` byte that is followed by a non-null. Thus, if your JPEG encoder outputs an `0xff`, it should also output an extra `0x00` afterwards to avoid problems.

This is very handy for us, particularly as several contiguous segments with a length and value (APPx `0xe?` and COM `0xfe`) will be ignored.

12.2.1 Crafting our Colliding Pair

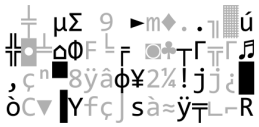
First, our blocks should be valid JPEGs. They must start with `0xff 0xd8`, which we can control. Then we need one last byte we can fully control, `0xff`, to start a segment. Then comes the fourth byte, which we'll set to `0xe?`. With luck, both cases will give us a valid+ignored segment start. Lastly comes the size of the segment, which we can't fully control, but which will not be too large as it's encoded in two bytes.

So, if we're lucky enough that the blocks are not too small, end after the `0x40` byte block, and their ends are not too close to each other, we just have to place the segments of two different JPEG pictures where these segments are ending.

Now we just have to hope that none of our random bytes creates an `0xff` byte. If we can't create the `0xff` sequence right after the signature, then we could retry later in the file, as other random data will be okay as long as no `0xff` appears.

We now have two valid JPEG start markers, and starting at the same offset two dummy segments of different lengths. All that is needed now is to start a comment segment right after the end of the smaller dummy segment, to comment out the first image's segment that will be placed immediately following the longest dummy segment. After the comment segment, we place the segment of the second image.

In one block, the dummy segment is longer; right after it come the segments of a valid JPEG image. In the other block, the dummy segment is shorter; it is directly followed by a comment segment that covers the rest of the longer dummy chunk and the chunks of the first valid image. Right after this comment segment come the segments of the second JPEG image.



JPEG signature — Chunk marker: - ff e5 in block 1, - ff e6 in block 2 — Chunk length: - c4 00 in block 1, - e4 00 in block 2

```
00000: ff d8 ff e? ?4 00 39 54 ?? 6d 04 2e ?? b7 b2 ??
       ?? 08 cf ?? ?? 46 d4 ?? ?? 0a 05 ?? ?? cb e2 ??  (contains no 0xff)
       ?? 87 fc ?? 38 98 83 ?? ?? 32 ac ?? ?? 6a a8 ??
       ?? 43 1f ?? ?? 66 87 f5 ?? 85 f7 ?? ?? 1c a9 ??


0c404: ff fe b5 e9     <COMment chunk covering Image 1>

0e404: ff e0           <start of Image 1>

       ...
       ff d9           <end of Image 1> <end of comment>

179ed: ff e0           <start of Image 2>

1b0d7: ff d9           <end of Image 2>
```

So now we have two blocks that can integrate any pair of standard JPEG files, provided they're not too big, and also a Rar archive collision, as one of the blocks ends with an 'R'. Why not, when we get the Rar for free?

12.3 And a Failure

The PE file format starts with an obsolete DOS header that is `0x40` bytes long (exactly the size of our block!), for which the only relevant elements nowadays are as follows:

- The '`MZ`' signature, at offset `0`.
- A pointer to the PE header, `e_lfanew`, aligned on four bytes at offset `0x3c`

As mentioned before, we know that the pointer will be different between the two blocks, as it is four bytes long. The problem is that the pointer in one of the two blocks will have a bit of its highest nybble set, thus that pointer will be greater than `0x1000000` (that's greater than 16 Gb). By manually crafting a

PE, the greatest value of `e_lfanew` that was found to be functional is `0xffffff0`, which is smaller than the lowest limit, yet very big. That PE itself is 268,435,904 bytes!

Thus, creating colliding PEs doesn't seem possible with this technique.

12.4 Conclusion

Having two different pictures with the same checksum that you can open in any image viewer is way more impressive than having two random colliding blocks—especially if you can freely use any picture for your final PoCs.

There are more than purely artistic reasons for studying polyglot collisions. When the attacker controls the constants as the hash function is initially specified, he only gets a single collision, a single pair of colliding blocks, for free. Finding more different collisions is as hard as finding one for the original SHA-1. So, if you want to have some freedom in using your collisions in practice, all target file formats must already be supported by your one colliding block.

In order to save significant time and heartache, a script was created that simulated all necessary conditions (generate two fully random blocks, set some bytes according to your rules, then check that they work). This script helped considerably to determine in advance the actual rules to feed the crunching cluster and then to be sure that you have working collisions at the end, rather than waiting a day or two to get the block pairs, which would likely fail to support the intended formats, and be forced to repeat this time-consuming and random process.

That makes two people happy: the cryptographer has a sexy new PoC, while the binarista has a nifty solution to an unusual challenge. Ain't that neighborly?