Anti-Disassembly using Cryptographic Hash Functions

John Aycock, Rennie deGraaf, and Michael Jacobson, Jr.
Department of Computer Science
University of Calgary

About the Authors

John Aycock is an assistant professor at the University of Calgary in the Department of Computer Science. He received a B.Sc. from the University of Calgary, and an M.Sc. and Ph.D. from the University of Victoria. He researches computer security and compilers, and conceived and taught the University's "Computer Viruses and Malware" and "Spam and Spyware" courses.

Rennie deGraaf is a graduate student in Computer Science at the University of Calgary, where he is researching firewalling technology. Other interests include port knocking, malware, honeypots, and the application of cryptography to solve real-world security problems.

Michael Jacobson is an assistant professor at the University of Calgary in the Department of Computer Science. He received a B.Sc. (Hon.) and an M.Sc. from the University of Manitoba and Dr. rer. nat. from the Technical University of Darmstadt. His research interests are cryptography and related applications of computational number theory.

Mailing Address: Department of Computer Science, University of Calgary, 2500 University Drive N.W., Calgary, Alberta, Canada T2N 1N4; Phone: +1 403 210 9409; E-mail: {aycock,degraaf,jacobs}@cpsc.ucalgary.ca.

Keywords

Anti-disassembly, code obfuscation, cryptography, cryptographic hash functions, botnets

Anti-Disassembly using Cryptographic Hash Functions

Abstract

Computer viruses sometimes employ coding techniques intended to make analysis difficult for anti-virus researchers; techniques to obscure code to impair static code analysis are called anti-disassembly *techniques. We present a new method of anti-disassembly based on cryptographic hash functions which is portable, hard to analyze, and can be used to target particular computers or users. Furthermore, the obscured code is not available in any analyzable form, even an encrypted form, until it successfully runs. The method's viability has been empirically confirmed. We look at possible countermeasures for the basic anti-disassembly scheme, as well as variants scaled to use massive computational power.*

Introduction

Computer viruses whose code is designed to impede analysis by anti-virus researchers are referred to as *armored* viruses.[1] Armoring can take different forms, depending on the type of analysis being evaded: dynamic analysis as the viral code runs, or static analysis of the viral code. In this paper, we focus on static analysis.

Static analysis involves the tried-and-true method of studying the code's disassembled listing. *Anti-disassembly* techniques are ones that try to prevent disassembled code from being useful. Code using these techniques will be referred to as *disassembly-resistant code* or simply *resistant code*. Although we are only considering anti-disassembly in the context of computer viruses, some of these techniques have been in use as early as the 1980s to combat software piracy (Krakowicz, c.1983).

Ideally, resistant code will not be present in its final form until run time – what can't be seen can't be analyzed. This could involve self-modifying code, which presents problems for static analysis of code (Lo, Levitt, & Olsson, 1995). It could also involve dynamic code generation such as a just-in-time compiler performs (Aycock, 2003).

In this paper, we present a new method of anti-disassembly based on dynamic code generation, which has the following properties:

- It can be targeted, so that the resistant code will only run under specific circumstances. We use the current username as a key for our running example, but any value available to the resistant code (or combinations thereof) with a large domain is suitable, like a machine's domain name. Because this key is derived from the target environment, and is not stored in the virus, our method may be thought of as environmental key generation (Riordan & Schneier, 1998).
- The dynamically generated code is not available in any form, even an encrypted one, where it can be subjected to analysis until the resistant code runs on the targeted machine. Other encryption-based anti-disassembly methods require that the resistant code be available in encrypted form (e.g., Filiol, 2005), in which case it may be subject to analysis.
- Even if the dynamically generated code were somehow known or guessed, the exact key

[1] The techniques we describe can be used by any malicious software (malware), so we use the term “computer virus” in this paper without loss of generality.

used by the resistant code is not revealed.

- It does not rely on architecture-specific trickery and is portable to any platform.

Below, we begin by explaining our anti-disassembly technique and presenting some empirical results. We then look at how the technique might be countered, along with some more entrepreneurial means of deployment.

The Idea

A cryptographic hash function is one that maps each input to a fixed-length output value, such that it is not computationally feasible to reverse the process, nor is it easy to locate two inputs with the same output (Schneier, 1996). Like regular hash functions, a cryptographic hash function is many-to-one.

Our idea for anti-disassembly is to combine a key – here, we use the current username for concreteness – with a "salt" value, and feed the result as input into a cryptographic hash function. The hash function produces a sequence of bytes, from which we extract a subsequence between bytes *lb* and *ub*, and interpret that subsequence as machine code. We will refer to this subsequence as a *run*. The salt value, for this application, is a sequence of bytes chosen by the virus writer to ensure that the desired run appears in the hash function's output when the correct key is used.

A pseudocode example of this idea is shown below: from a high-level point of view, this is what an analyst would be confronted with. The code for a cryptographic hash function is assumed to be available, likely in a system library, and `run` is the code sequence that the virus writer is trying to hide. The task of the analyst is to determine precisely what this code does when executed (the value of `run`) and what the target is (the correct value of `key`).

```
key ⇑ getusername()
hash ⇑ md5(key + salt)
run ⇑ hash_lb…ub
goto run
```

This pseudocode uses the username as a key, and MD5 as the cryptographic hash function (Rivest, 1992); + is the concatenation operator. MD5 is now known to be vulnerable to collisions (Wang, Feng, Lai, & Yu, 2004), but this is irrelevant to our technique: the analyst does not have the MD5 hash, and in any case would be interested in the exact value of the key, not just a key which produces the same hash. In addition, our anti-disassembly technique can be used with any cryptographic hash function, so a different/stronger one can be chosen if necessary.

There are three issues to consider:

1. Having the wrong key. Obviously, if the wrong key value is used, then the run is unlikely to consist of useful code. The resistant code could simply try to run it anyway, and possibly crash; this behavior is not out of the question for viruses. Another approach would catch all of the exceptions that might be raised by a bad run, so that an obvious crash is averted. A more sophisticated scheme could check the run's validity using a checksum (or re-using the cryptographic hash function), but this would give extra information to a code analyst.
2. Choosing the salt. This is the most critical aspect; we suggest a straightforward brute-force search through possible salt values. Normally, conducting a brute-force attack against a cryptographic hash function to find a particular hash value, i.e., a collision, would be out of the question because the hash functions are designed to make this computationally

prohibitive. However, we are only interested in finding a subsequence of bytes in the hash value, so our task is easier. An analysis of the expected computational effort required to find the required salt is presented in the next section.

3. Choosing *lb* and *ub*. These values are derived directly from the hash value, once the desired salt is found.

The salt search is by far the most time-consuming operation, but this need only be done once, prior to the release of the resistant code. The search time can be further reduced in three ways. First, faster machines can be used. Second, the search can be easily distributed across multiple machines, each machine checking a separate part of the search space. Third, the search can be extended to equivalent code sequences, which can either be supplied manually or generated automatically (Joshi, Nelson, & Randall, 2002; Massalin, 1987); since multiple patterns can be searched for in linear time (Aho & Corasick, 1975), this does not add to the overall time complexity of the salt search.

Analysis

In order to find a salt value, we simply compute the cryptographic hash of

```
key + salt
```

for all possible salt values until the hash output contains the required byte sequence (run). In order to speed up the search, we allow the run to begin in any position in the hash output.

Approximately half of the output bits of a cryptographic hash function change with each bit changed in the input (Schneier, 1996); effectively, we may consider the hash function's output to change randomly as the salt is changed. Given that, the probability of finding a particular *b*-bit run in a fixed position of an *n*-bit output is the ratio of the bits not in the run to the total number of bits: $2^{n-b}/2^n$, or $1/2^b$. The expected number of attempts would then be 2^{b-1}. Furthermore, because only the salt is being changed in the brute-force search, this implies that we would need *b-1* bits of salt for a *b*-bit run.

If we allow *lb*, the starting position of the run, to vary, the expected number of attempts will be reduced by a factor equal to the number of possible values of *lb*. If we index the starting position at the byte level, then there are $m = (n-b)/8$ possible starting positions. The probability of finding the *b*-bit run increases to $m/2^b$, and the expected number of attempts becomes $2^{b-1}/m$. Similarly, if we index at the bit level, there are *n-b* starting positions and the expected number of attempts reduces further to $2^{b-1}/(n-b)$.

Notice that the computational effort depends primarily on the length of the run, not the length of the hash function output. The length of the hash function only comes into play in reducing the expected number of attempts because the number of possible values for *lb*, the starting point of the run, depends on it.

We only discuss the case of single runs here, but this technique trivially extends to multiple runs, each with their own salt value. Because the salt computation for each run is independent of the others, the total effort required for multiple-run salt computation scales linearly. If the computational effort to compute the salt for one run is *X*, then the effort for one hundred runs is *100X*.

As an example of salt computation, suppose we want our run to consist of a single Intel x86 relative jump instruction. This instruction can be encoded in five bytes, so we need to find a salt that, when concatenated to the key, yields a hash value containing this five-byte run starting in any position.

The MD5 hash function has 128-bit outputs, so if we index the run at the byte level, there are 11 possible values for *lb*. The expected number of attempts to find the run is therefore

$$2^{39} / 11 < 2^{36}.$$

If we instead index at the bit level, there are 88 possible values for *lb* and the expected number of attempts is

$$2^{39} / 88 < 2^{33}.$$

Using a 160-bit hash function such as SHA-1 yields $2^{39} / 15$ and $2^{39} / 120$ when indexing *lb* at the byte and bit levels, respectively. In all cases, the computation can be done in only a few hours on a single modern desktop computer.

It is feasible to use this method to find runs slightly longer than five bytes, but the computational effort adds up very quickly. For example, to find an eight-byte run using SHA-1 and indexing *lb* at the bit level would require roughly $2^{63} / 120 > 2^{56}$ attempts. A special-purpose, massively parallel machine would likely be required to find the run in this case, as the computational effort involved is roughly equivalent to that required to break the DES block cipher, for which such hardware was also required (Electronic Frontier Foundation, 1998).

Empirical Results

To demonstrate the feasibility of this anti-disassembly technique, we searched for the run (in base 16)

```
e9 74 56 34 12
```

These five bytes correspond on the Intel x86 to a relative jump to the address 12345678_{16}, assuming the jump instruction starts at address zero.

Algorithm	*Key*	*Salt*	*Search Time (s)*
MD5 (128 bits)	aycock	55b7d9ea16	39615
	degraaf	a1ddfc1910	47650
	foo	e6500e0214	60185
	jacobs	9ac1848109	28723
	ucalgary.ca	4d21abe205	18220
	Average		44746
SHA-1 (160 bits)	aycock	07e9717a09	36584
	degraaf	0d928a260e	55424
	foo	2bc680de1e	120472
	jacobs	ca638d5e06	24958
	ucalgary.ca	585cc614	325
	Average		47552

Table 1: Brute-force salt search for a specific five-byte run

The search was run on an AMD AthlonXP 2600+ with 1 GB RAM, running Linux 2.6. We tested five different keys with one- to five-byte salts, sequentially searching through the possible salt

values.[2] Table 1 shows the results for two cryptographic hash functions, MD5 and SHA-1. For example, the salt "07e9717a09," when concatenated onto the key "aycock," yields the SHA-1 hash value

```
ef 6d f4 ed 3b a1 ba 66 27 fe e9 74 56 34 12 a2 d0 4f 48 91
```

Numbering the hash value's bytes starting at zero, our target run is present with *lb = 10* and *ub = 14*.

Another question is whether or not every possible run can be produced. Using the key "aycock," we were able to produce all possible three-byte runs with three bytes of salt, but could only produce 6% of four-byte runs with a three-byte salt. With a four-byte salt, we were able to generate four-byte runs which covered between 99.999-100% of the possible combinations – this was checked with five different keys and three different cryptographic hash functions. (Our test system did not have sufficient memory to record coverage data for five-byte runs in a reasonable amount of time.) The four-byte run data are shown in Table 2.

Algorithm	*Key*	*Runs Found*	*Runs Not Found*
MD5 (128 bits)	aycock	4294936915	30381
	degraaf	4294937044	30252
	foo	4294936921	30375
	jacobs	4294937188	30108
	ucalgary.ca	4294936946	30350
	Average	4294937003	30293
SHA-1 (160 bits)	aycock	4294966707	589
	degraaf	4294966733	563
	foo	4294966660	636
	jacobs	4294966726	570
	ucalgary.ca	4294966769	527
	Average	4294966719	577
SHA-256 (256 bits)	aycock	4294967296	0
	degraaf	4294967296	0
	foo	4294967296	0
	jacobs	4294967296	0
	ucalgary.ca	4294967296	0
	Average	4294967296	0

Table 2: Generation of possible four-byte runs using a four-byte salt

These results tend to confirm our probability estimate from the last section: *b*-bit runs need *b-1* bits of salt. Four-byte runs are of particular interest for portability reasons, because RISC instruction sets typically use instructions that are four bytes long; this means that at least one RISC instruction can be generated using our technique. One instruction may not seem significant, but it is sufficient

[2] For implementation reasons, we iterated over salt values with their bytes reversed, and didn't permit zero bytes in the salts.

to perform a jump anywhere in the address space, perform an arithmetic or logical operation, or load a constant value – potentially critical information that could be denied to an analyst.

Countermeasures

An analyst who finds some resistant code has several pieces of information immediately available. The salt, the values of *lb* and *ub*, and the key's domain (although not its value) are not hidden. The exact cryptographic hash function used can be assumed to be known to the analyst, too – in fact, resistant code could easily use cryptographic hash functions already present on most machines.

There are two pieces of information denied to an analyst:

1. The key's value. Unless the key has been chosen from a small domain of values, then this information may not be deducible. The result is that an analyst may know that a computer virus using this anti-disassembly technique targets someone or something, but would not be able to uncover specifics.
2. The run. If the run is simply being used to obscure the control flow of the resistant code, then an analyst may be able to hazard an educated guess about the run's content. Other cases would be much more difficult to guess: the run may initialize a decryption key to decrypt a larger block of code; the entire run may be a “red herring” and only contain various NOP instructions.

Note that even if the run is somehow known to an analyst, the cryptographic hash function cannot be reversed to get the original key; this is a property of the cryptographic hash function (Schneier, 1996). At best, the analyst could perform their own brute-force search to determine a set of possible keys (recall that the hash function is many-to-one). However, the analyst also knows the salt and the domain of the key, so given the run, the analyst can find the key by exhaustively testing every possible value. This underscores the point that the key domain must be sufficiently large to preclude such a brute-force analysis – our example in the last section of using usernames as keys would likely not prevent this.

Whether or not every last detail of the resistant code can be found out is a separate issue from whether or not a computer virus using resistant code can be detected. In fact, there is malware already that can automatically update itself via the Internet, like Hybris (F-Secure, 2001), so complete analysis of all malware is already impossible.

Fortunately for anti-virus software, computer viruses using the technique we describe would present a relatively large profile which could be detected with traditional defenses, including signature-based methods and heuristics (Szor, 2005). Precise detection does not require full understanding.

Enter the Botnet

What if the computing power available for a brute-force salt search were increased by five orders of magnitude over the computer we used for our experiments? Few organizations have that much computing power at their fingertips, but a few individuals do. A *botnet* is a network of malware-controlled, “zombie” machines that executes commands issued via Internet Relay Chat (IRC) channels (Cooke, Jahanian, & McPherson, 2005). These have been used for sending spam and distributed denial-of-service attacks (Cooke et al., 2005), but they may also be viewed as very large-scale distributed computing frameworks which can be used for malicious purposes.

If a virus writer wants to armor a virus using the anti-disassembly technique described here, especially for long runs with many instructions, a botnet may be used for salt computation. A naïve

salt computation on a botnet would involve partitioning the salt search space between machines, and the key and desired run would be available to each machine. Using the earlier Intel x86 relative jump example, for instance, four zombie machines in a botnet could each be given the desired key (e.g., ``aycock'') and run (e974563412) and a four-byte salt search could be divided like so:

Zombie 1 00000000…3fffffff

Zombie 2 40000000…7fffffff

Zombie 3 80000000…bfffffff

Zombie 4 c0000000…ffffffff

Having the virus writer's desired key and run on each zombie machine would not be a bad thing from an analyst's point of view, because locating any machine in the botnet would reveal all the information needed for analysis.

A more sophisticated botnet search would do three things:

1. Obscure the key. A new key, *key′*, could be used, where *key′* is the cryptographic hash of the original key. The deployed resistant code would obviously need to use *key′* too.
2. Supply disinformation. The virus writer may choose *lb* and *ub* to be larger than necessary, to mislead an analyst. Unneeded bytes in the run could be NOP instructions, or random bytes if the code is unreachable.
3. Hide the discovery of the desired run. Instead of looking for the exact run, the botnet could simply be used to narrow the search space. A weak checksum could be computed for *all* sequences of the desired length in the hash function's output, and the associated salts forwarded to the virus writer for verification if some criterion is met. For example, the discovery of our five-byte run in the "Empirical Results" section could be obliquely noted by watching for five-byte sequences whose sum is 505.

This leaves open two countermeasures to an analyst. First, record the *key′* value in an observed botnet in case the salt is collected later, after the virus writer computes and deploys it – this would reveal the run, but not the original key. Second, the analyst could subvert the botnet, and flood the virus writer with false matches to verify. The latter countermeasure could itself be countered quickly by the virus writer, however, by verifying the weak checksum or filtering out duplicate submissions; in any case, verification is a cheap operation for the virus writer.

Related Work and Conclusion

There are few examples of strong cryptographic methods being used for computer viruses – this is probably a good thing. Young and Yung (1996) discuss cryptoviruses, which use strong cryptography in a virus' payload for extortion purposes. Riordan and Schneier (1998) mention the possibility of targeting computer viruses, as does Filiol (2005).

Filiol's work is most related to ours: it uses environmental key generation to decrypt viral code which is strongly-encrypted. Neither his technique nor ours stores a decryption key in the virus, instead finding the key on the infected machine. A virus like the one Filiol proposes hides its code with strong encryption, carrying the encrypted code around with the virus. In our case, however, the code run never exists in an encrypted form; it is simply an interpretation of a cryptographic hash function's output. Our technique is different in the sense that the ciphertext is not available for

analysis.

The dearth of strong cryptography in computer viruses is unlikely to last forever, and preparing for such threats is a prudent precaution. In this particular case of anti-disassembly, traditional defenses will still hold in terms of detection, but full analysis of computer viruses may be a luxury of the past. For more sophisticated virus writers employing botnets to find salt values and longer runs, proactive intelligence gathering is the recommended defense strategy.

Acknowledgments

The first and third authors' research is supported in part by grants from the Natural Sciences and Engineering Research Council of Canada. Karel Bergmann made helpful comments on an early version of this paper, as did the anonymous referees.

References

Aho, A. V., and Corasick, M. J. (1975). Efficient string matching: an aid to bibliographic search. Communications of the ACM, 18(6), 333–340.

Aycock, J. (2003). A brief history of just-in-time. ACM Computing Surveys, 35(2), 97–113.

Cooke, E., Jahanian, F., & McPherson, D. (2005). The zombie roundup: Understanding, detecting, and disrupting botnets. In USENIX SRUTI Workshop.

Electronic Frontier Foundation (1998). Cracking DES: Secrets of Encryption Research, Wiretap Politics, and Chip Design. O'Reilly.

F-Secure (2001). F-Secure virus descriptions: Hybris. Retrieved 5 January, 2006, from http://www.f-secure.com/v-descs/hybris.shtml.

Filiol, E. (2005). Strong cryptography armoured computer viruses forbidding code analysis: The Bradley virus. In Proceedings of the 14th Annual EICAR Conference, pages 216–227.

Joshi, R., Nelson, G., & Randall, K. (2002). Denali: a goal-directed superoptimizer. In Proceedings of the ACM SIGPLAN 2002 Conference on Programming Language Design and Implementation, pages 304–314.

Krakowicz (c. 1983). Krakowicz's kracking korner: The basics of kracking II. Retrieved 5 January, 2006, from http://www.skepticfiles.org/cowtext/100/krckwczt.htm.

Lo, R. W., Levitt, K. N., & Olsson, R. A. (1995). MCF: a malicious code filter. Computers & Security, 14, 541–566.

Massalin, H. (1987). Superoptimizer: a look at the smallest program. In Proceedings of the Second International Conference on Architectural Support for Programming Languages and Operating Systems, pages 122–126.

Riordan, J., & Schneier, B. (1998). Environmental key generation towards clueless agents. In Mobile Agents and Security (LNCS 1419), pages 15–24.

Rivest, R. (1992). The MD5 message-digest algorithm. RFC 1321.

Schneier, B. (1996). Applied Cryptography. Wiley, 2nd edition.

Szor, P. (2005). The Art of Computer Virus Research and Defense. Addison-Wesley.

Wang, X., Feng, D., Lai, X., & Yu, H. (2004). Collisions for hash functions MD4, MD5, HAVAL-128 and RIPEMD. Cryptology ePrint Archive, Report 2004/199. Retrieved 5 January, 2006, from http://eprint.iacr.org/.

Young, A., & Yung, M. (1996). Cryptovirology: Extortion-based security threats and countermeasures. In IEEE Symposium on Security and Privacy, pages 129–141.