Problems with the FIPS 140 Certification Process

Independent Security Evaluators
www.securityevaluators.com

April 15, 2007

1 Introduction

For over a decade all software and hardware cryptographic modules used by the United States government to protect unclassified data have had to pass a vigorous certification process known as the Cryptographic Module Validation Program (CMVP) [2], established by the National Institute for Standards and Technology (NIST). This program is designed to verify that a cryptographic module does indeed meet the requirements specified in a series of documents known as the Federal Information Processing Standard 140 (FIPS-140) [1]. These documents describe in detail the requirements for a module's design, implementation, testing and accompanying documentation. They provide a framework for bestowing upon a module a specific level of security (of which there are four) in 11 distinct categories.

In recent years especially, more and more products have been receiving this certification. Though it is specifically noted by the FIPS 140-2 publication that security is not guaranteed by the validation process, there is a tendency in the commercial world to believe that receiving FIPS certification does in fact provide some level of security assurance. This is not the case.

The CMVP excludes numerous areas of potential vulnerability from the review process (some of which are fundamental to the goals set out to achieve by the CMVP), including the proper usage of cryptographic algorithms within the certified module, the verifiable destruction of critical security parameters (CSPs) (i.e. keys), and the proper application of user authentication processes. Additionally, the CMVP is limited in several ways. The list of approved cryptographic algorithms and modes of operations excludes the most effective, efficient and secure algorithms available, and the process imposes several useless requirements from a practicality point of view.

This white paper describes our experiences with the FIPS certification process, from the standpoint of a security consulting firm assisting other companies to generally secure their products as they undergo the FIPS certification process and the CMVP.

2 Non-cryptographic security components

The primary drawback of the CMVP is that there is no evaluation of many security features not mentioned in FIPS 140-2. With the exception of cryptographic algorithm correctness, key management policy, integrity testing and physical security in the case of actual physical modules, there is no sound manner in which other vulnerabilities can be addressed.

Buffer overflows. Potentially one of the most serious drawbacks, is that there is no evaluation of the software or hardware's effective prevention of buffer overflow attacks or other similar avenues of compromise. These sorts of attacks are widely prevalent in the real world and are very powerful tools for an adversary. Using an attack such as this, an adversary corrupts the underlying executable while it is running to reach various ends, such as causing the module to generally misbehave, causing the module to output sensitive information such as cryptographic keys, or even compromising the underlying system on which the module is executing.

Some might argue that the CMVP does attempt to evaluate whether these vulnerabilities exist, by testing the module with various corrupted inputs, but this process merely tests whether or not the module outputs CSPs or performs unexpected actions. Also, in higher levels of certification for software modules the CMVP requires modules to be executed on certain "hardened" operating systems, which may or may not include adequate buffer overflow prevention mechanisms.

In either case, these activities are not sufficient to guarantee the resilience of the module to buffer overflow attacks, and though there is no process that can make such definitive guarantees, an extensive security analysis with the goal of eliminating this avenue of compromise should be performed.

Availability. Another critical area of security unrecognized by the CMVP is availability and the potential risk of denial of service (DoS) attacks. For all intents and purposes, DoS has little if anything to do with cryptography. However, providing the availability of a cryptographic service is often a critical security goal. And though a module's robustness against DoS attacks has no real measurable scale, or standard system for evaluation, improvements can be made through a traditional security analysis performed by security experts. The CMVP itself offers no security improvements in this area.

In the business world, DoS is often thought of as an attack on a network, but there are more direct relationships to cryptographic modules than typically realized. The ”service” denied in a DoS attack is by no means limited to network services, and can include the exhaustion of random number generation entropy sources, the saturation of CPU resources through excessive encryptions or decryptions, or the ease of crashing a cryptographic module altogether. Any one of these attacks directly affect the performance and availability of the cryptographic module itself, and as mentioned previously, the CMVP does not consider these sorts of attacks.

Information Leakage. The CMVP employs a very strict method for determining whether sensitive information is obtainable through manipulation of the cryptographic module being tested. That is, CSPs used within the module are identified, verified to never be displayed in the clear and destroyed when no longer needed. However, this method only serves to identify actual, tangible CSPs as being output, but is insufficient to identify many sources of information leakage that could lead to a security failure.

Take for example a module that encrypts messages sent in real time between two end-points. The CMVP may prove (to a reasonable extent) that an encryption key for the stream of encrypted data is never be revealed, and hence the plaintext data source can never be successfully decrypted. However, the CMVP does not take in to consideration non-cryptanalytical attacks such as the statistical timing analysis of keystrokes that could possibly reveal the entire plaintext data source without requiring the encryption key at all.

In general, any security analysis should consider many more avenues of compromise than simply determining whether or not a cryptographic key is ever disclosed. There is no formal methodology capable of taking all aspects of information leakage under consideration, instead the best that can be achieved is to thoroughly and diligently evaluate any security system from the viewpoint of an adversary, and determine exactly what information is obtainable.

And many more... We've discussed above the more common security attributes of a system that cannot be accurately assessed through the CMVP, but this list is nowhere near complete. Many security products incorporate security goals pertaining to anonymity, privacy, obfuscation, intrusion detection, virus detection/prevention, reverse-engineering protections, copy protection, covert communications, and so on... The CMVP makes no claims or attempts to evaluate these security metrics as they do not pertain directly to the cryptographic functionality of a security product. The CMVP is therefore not necessarily at fault, but it should be noted over and over again that FIPS certification does not imply the validity of any security a system employees save its cryptographic functions.

Or even less...

3 Cryptographic security components

Concerning the cryptography itself, the CMVP primarily tests a cyrptographic module for correctness of algorithm implementation. That is, the algorithms are given a set of inputs, and the values that are output are verified to be correct. However, the CMVP does not identify flaws in the module that arise from misuse of the algorithms themselves.

Initialization vector reuse. For example, the CMVP can prove beyond a reasonable doubt that an implementation of the Advanced Encryption Standard (AES) is correct, but does not take in to consideration whether or not the cryptographic module is repeatedly using the same Initialization Vectors (IV) for each encryption it performs. In many cases, IV reuse poses a tremendous risk to the security of the encrypted data. In certain FIPS 140-2 approved modes of operation, such as AES in counter mode, the repeated reuse of an IV could reveal all encrypted data to an adversary possessing a single known plaintext/ciphertext pair.

Incorrect use of primitives. In the previous example, it was a poor choice of algorithm inputs that lead to the compromise of the plaintext data being protected. Another manner in which the CMVP lacks sufficiency is in the incorrect combination of cryptographic primitives that leads to security vulnerabilities.

For example, consider a cryptographic module that takes as input a plaintext data source, and outputs both a hash of the plaintext data source and a ciphertext. The CMVP does not recognize that including an unencrypted hash of a plaintext data source along side of the ciphertext is a security vulnerability. It is widely accepted in the security community that hashes of plaintext data should be treated as plaintext data themselves, and included within the ciphertext rather than along side of it. An alternative (depending on the security requirements of the module) could be to compute the hash digest over the ciphertext, in which case the hash can exist along side of the encrypted data with no security implications.

There are many more such examples, from the blatant and common mistakes as mentioned above to more subtle, implementation specific mistakes that can completely squash the security goals set out to provide by a cryptographic module. The CMVP is not likely to catch any of these vulnerabilities in the process of validating the correctness of the various algorithms. Instead, a formal security evaluation provided by experts should be performed on any cyrptographic module during its design phase to eliminate these mistakes.

4 Choice of cryptographic algorithms

The current FIPS 140 contains relatively few approved cryptographic algorithms, while in practice there are more, and often stronger or more efficient, algorithms to choose from. This is probably due to the fact that approving and adding algorithms to the FIPS is a longer and more arduous process than the simple adoption of these algorithms in the commercial world. However, the fact remains that a cryptographic module operating in a FIPS certified mode is restricted to using the approved modes of operation and may not provide the best security the industry has to offer.

Authenticated encryption. One serious draw back to the list of approved modes of operation, is that there are no FIPS approved *authenticated* encyrption modes, such as OCB and XCBC modes. These modes of operation provide a stronger guarantee than simply providing the confidentiality of data. Authenticated encryption modes simultaneously provide the authenticity of the underlying plaintext data source. Without the availability of these authenticated encryption modes, the only option available to developers is to utilize multiple approved algorithms in tandem to create their own authentication and encryption mechanisms.

For example, in lue of OCB or XCBC modes, a cyrptographic module could encrypt a data source using AES and subsequently calculate a message authentication code (MAC) over the resulting ciphertext. This effectively provides for the authenticity and confidentiality of the underlying plaintext. However, as mentioned in the previous section the CMVP does not validate whether these two algorithms are used together correctly. Using an authenticated encryption mode provides the authenticity and confidentiality in a single step, under a single key. This reduces overhead and there is less room for error as only a single mode is required.

Provably secure algorithms. Another issue that is raised with the choice of approved algorithms is that none of the approved algorithms employ “provable security.” Provable security is a concept by which ... Regular encryption modes do not provide this guarantee.

For example, the FIPS 140-2 allows for the use of AES, 3DES, and Skipjack, including the AES-based NIST key wrapping function, none of which are *provable*, while in practice there exist a number of algorithms and modes of operation that are provable, such as ..., and .... These algorithms make stronger choices for a cryptographic algorithm that could easily be used in place of any of the unprovable modes.

5 Zeroization

The CMVP is diligent in its task of making sure no CSPs are output from the cryptographic module and that no CSPs remain after the cryptographic module has been compromised, destroyed, or released from memory in a software module. However, considering software-only modules, it is debatable whether or not this validation process is sufficient (at least for the lower certification levels).

Level 1 certification for software modules requires that all CSPs be documented, and it is checked that each of these CSPs are written over with zeroes before the module is released from memory. For example, if memory is allocated for the storage of an encryption key (or any CSP), the validation process checks that this memory location is overwritten with zeroes before it is released back to the operating system. However, it is a hazardous assumption to claim that simply writing over the only user-assigned memory location with zeroes is sufficient to destroy the CSP. Operating systems are complicated, and often times memory is swapped to disk, output as a stack-dump during a system error, or even inadvertently copied during various system calls. Studies have shown that memory is often scattered heavily throughout a computer system without any indication that this has occurred from the viewpoint of the executing application.

Higher certification levels for software modules require that the underlying operating system have a minimum EAL rating, which in some cases could provide a better guarantee that memory is kept secure. Short of this, there are options to developers for providing better security for the CSPs stored in memory. Systems can be configured to use an encrypted swap drive, so that memory swapped to disk will not be readily available to an adversary, secure memory management systems can be employed, and the exclusion of calls that could inadvertently copy data can be avoided.

6 Module Integrity Checking

Another problematic requirement of the CMVP is that all cryptographic modules have an integrity verification test. This integrity test must utilize one of the FIPS approved algorithms for providing integrity, and be executed whenever the module is initialized. For example, when a module is first powered on, it could verify an embedded DSA signature over the entire module's data as it has been stored, demonstrating that no bits of the module have been altered, and if the signature does not verify, the module will cease to operate. As far as simple error checking is concerned, these algorithms are overkill, and given the requirement that the algorithm used be a FIPS-approved cryptographic algorithm, it suggests that resistance to malleability by an adversary is the primary security goal.

In modules that are built in hardware, this policy may have more substance, as the modification of a single, critical bit using specialized attack tools may be more plausible than accurately modifying an entire portion of the module's data. However, in software only modules it is likely that an adversary capable of modifying a single bit is also capable of modifying the entire signature itself. This signature could then simply be replaced with a new signature that accepts any of the other modifications the adversary has made to the stored state of the module. Better still, an adversary could remove the integrity check altogether and always return a successful result regardless of how the module has been changed.

7 Authentication

The CMVP also demands a slew of requirements to be met when it comes to user authentication in a FIPS certified module. Some of them however provide no real security enhancement. Considering passwords, there are lower bound restrictions placed on the number of possible passwords that could possibly exist, as well as upper bound restrictions placed on the number of password entry attempts that can be made within a given time frame. The fact that these restrictions are present is all that is verified by the CMVP, whether or not they are utilized properly and securely is not.

Choosing passwords. Passwords are required to be chosen from a universe of passwords that is very large. In other words, given the available character set and minimum length requirement placed on a password, there has to be a minimum number of possible password choices. This restriction seems reasonable, however there is no restriction on the actual passwords that can be chosen. The CMVP verifies that any given password is 1 of N passwords, where N is very large, but does not verify that all passwords chosen contain sufficient entropy to make password guessing attacks unreasonable.

For example, there may be trillions of possible password choices in a module that allows 64 choices per character and requires at least 8 characters, but there is no guarantee that the module will only ever see weak passwords such as “password1,” “password2,” “password3”, etc.

Online guessing attacks. The CMVP also mandates that the number of password entry attempts to a module cannot exceed a specified threshold within a given amount of time. For example, after 3 consecutive failed password entry attempts, a system may require the user to wait 15 seconds before another attempt is possible. The attack prevented here is known as an online password guessing attack, the keyword being *online.*

In practice, it is possible to provide protection against online password guessing attacks, as well as restrict the possibility of an attack to an online scenario, however the CMVP does not guarantee these attributes.

For example, the CMVP may restrict an account log on to 3 tries per 15 seconds, but this security premise alone does not prevent online guessing attacks. The security model does not consider such adversaries whom are not interested in obtaining access to a specific account, but are instead interested in accessing *any* account. A slightly more advanced adversary could spread the guessing attack over multiple (perhaps tens of thousands) accounts simultaneously, vastly increasing the number of possible guesses per 15 second interval.

Offline guessing attacks. Still, even should online guessing attacks be entirely prevented, the security model checked by the CMVP does not include the possibility of reducing the password guessing attacks to an offline setting. For example, consider a module that provides some sort of remote log on where a password is verified through a challenge and response protocol. The server sends a challenge to the client, the client then encrypts the challenge using a hash of the password as the key and sends this value to the server, and the server verifies that the value received is indeed the correct challenge encrypted using the correct password. Even if the system allows only 3 attempts to correctly perform the procedure before locking the account, an adversary in possession of both the challenge and response could execute the guessing attack offline. In this case, the adversary would simply repeatedly encrypt the challenge under different passwords until the correct encryption is found, with no module intervention.

8 The up side

This whitepaper only serves to point out the areas where the FIPS 140 certification process and the CMVP are not sufficient to provide any sort of guarantee of security, and not to claim that the process is entirely bad or has no merit. In fact, obtaining FIPS certification forces developers down a strict path requiring detailed documentation of module design and functionality, adherence to explicit implementation guidelines, and vigorous testing to validate that the module accurately computes cryptographic algorithms. Simply undergoing the process is sure to harden the implementation itself.

Furthermore, NIST claims that in as many as 50% of the modules that have been validated, errors in the cryptographic algorithms have been caught and corrected. It isn’t clear what impact exactly this has on the security of a module, but you can pretty much guarantee that a FIPS certified module is performing its cryptographic algorithms correctly.

9 Additional Security Evaluation

In any case, having this stamp of approval does not come close to a guarantee that the security of a product is sufficient. Instead it is prudent and necessary to have an external security-centric evaluation of a product for locating vulnerabilities and design flaws that the CMVP does not identify.

As described in this white paper, there are many cases left unchecked by the CMVP, such as the presence of buffer overflows, the misuse of cryptographic algorithms in a manner that makes them weaker, the misuse of the product itself, or basically any other security issue not relating to tamper resistance or specific cryptographic algorithms. All products that provide a security component, whether certified or not should receive some sort of security evaluation by an independent 3rd party. There are numerous benefits to this activity, including having new eyes critic the choices made in designing a product, catching security flaws or vulnerabilities that would not ordinarily be brought to the forefront by the CMVP, other certification processes or by stand alone analysis tools, and most importantly the advice and input from industry experts in helping to design secure products.

References

[1] National Institute of Standards and Technology. Fips pub 140-2, 2002.

[2] National Institute of Standards and Technology. Cryptographic module validation program. http://csrc.nist.gov/cryptval/, April 2007.