Opinion: Cryptanalysis of the Hashing Algorithms known as MD5 and SHA Is it now time for a new standard?
Opinion by Bruce Schneier, Counterpane Internet Security Inc.

AUGUST 19, 2004 (COMPUTERWORLD) - At the Crypto 2004 conference in Santa Barbara, Calif., this week, researchers announced several weaknesses in common hash functions. These results, while mathematically significant, aren't cause for alarm. But even so, it's probably time for the cryptography community to get together and create a new hash standard. One-way hash functions are a cryptographic construct used in many applications. They are used with public-key algorithms for both encryption and digital signatures. They are used in integrity checking. They are used in authentication. They have all sorts of applications in a great many different protocols. Much more than encryption algorithms, one-way hash functions are the workhorses of modern cryptography.

In 1990, Ron Rivest invented the hash function MD4. In 1992, he improved on MD4 and developed another hash function: MD5. In 1993, the National Security Agency published a hash function very similar to MD5, called the Secure Hash Algorithm (SHA). Then in 1995, citing a newly discovered weakness that it refused to elaborate on, the NSA made a change to SHA. The new algorithm was called SHA-1. Today, the most popular hash function is SHA-1, with MD5 still being used in older applications.

One-way hash functions are supposed to have two properties. One, they're one-way. This means that it's easy to take a message and compute the hash value, but it's impossible to take a hash value and re-create the original message. (By "impossible," I mean "can't be done in any reasonable amount of time.") Two, they're collision-free. This means that it's impossible to find two messages that hash to the same hash value. The cryptographic reasoning behind these two properties is subtle, and I invite curious readers to learn more in my book Applied Cryptography.

Breaking a hash function means showing that either -- or both -- of those properties aren't true. Cryptanalysis of the MD4 family of hash functions has proceeded in fits and starts over the past decade or so, with results against simplified versions of the algorithms and partial results against the whole algorithms.

This year, Eli Biham and Rafi Chen, and separately Antoine Joux, announced some pretty impressive cryptographic results against MD5 and SHA. Collisions have been demonstrated in SHA. And there are rumors, unconfirmed at this writing, of results against SHA-1.

The magnitude of these results depends on who you are. If you're a cryptographer, this is a huge deal. While not revolutionary, these results are substantial advances in the field. The techniques described by the researchers are likely to have other applications, and we'll be better able to design secure systems as a result. This is how the science of cryptography advances: We learn how to design new algorithms by breaking other algorithms. In addition, algorithms from the NSA are considered a sort of alien technology: They come from a superior race with no explanations. Any successful cryptanalysis against an NSA algorithm is an interesting data point in the eternal question of how good they really are in there.

As a user of cryptographic systems -- as I assume most readers are -- this news is important, but not particularly worrisome. MD5 and SHA aren't suddenly insecure. No one is going to be breaking digital signatures or reading encrypted messages anytime soon with these techniques. The electronic world is no less secure after these announcements than it was before.

But there's an old saying inside the NSA: "Attacks always get better; they never get worse." These techniques will continue to improve, and probably someday there will be practical attacks based on these techniques.

It's time for us all to migrate away from SHA-1.

Luckily, there are alternatives. The National Institute of Standards and Technology (NIST) already has standards for longer --and harder-to-break -- hash functions: SHA-224, SHA-256, SHA-384 and SHA-512. They're already government standards and can already be used. This is a good stopgap, but I'd like to see more.

I'd like to see NIST orchestrate a worldwide competition for a new hash function, like it did for the new encryption algorithm, Advanced Encryption Standard, to replace Data Encryption Standard. NIST should issue a call for algorithms and conduct a series of analysis rounds, where the community analyzes the various proposals with the intent of establishing a new standard.

Most of the hash functions we have and all the ones in widespread use are based on the general principles of MD4. Clearly we've learned a lot about hash functions in the past decade, and I think we can start applying that knowledge to create something even more secure.

Better to do it now, when there's no reason to panic, than years from now, when there might be.