Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Finanzdepartement EFD

Informatikstrategieorgan Bund ISB
Informatiksicherheit SEC

Rolf Oppliger 4. August 2010

Technologiebetrachtung

Kollisionsresistenz und Brechung kryptografischer Hashfunktionen

1 Einleitung

In den Medien wird immer wieder über die Brechung von kryptografischen Hashfunktionen (z.B. MD5, SHA-1, …) berichtet. Dabei ist nicht immer klar, was Gegenstand der Brechung ist bzw. was die Implikationen sind. Im Rahmen dieser Technologiebetrachtung wird kurz aufgezeigt, was eine kryptografische Hashfunktion ist, welche spezifischen Eigenschaften sie aufweist, was unter einer Brechung zu verstehen ist, was die Implikationen sind und was man aus applikatorischer Sicht dagegen tun kann.

2 Kryptografische Hashfunktionen

Kryptografische Hashfunktionen stellen Kompressionsfunktionen dar, die – in Ergänzung zu ihrer Kompressionseigenschaft – noch verschiedene andere Eigenschaften aufweisen müssen. Insbesondere muss eine solche Funktion h kollisionsresistent sein, d.h. es darf für einen Aussenstehenden berechenbar nicht möglich sein, zwei Nachrichten x und x‘ zu finden, die unter h abgebildet den gleichen Hashwert h(x) = h(x‘) aufweisen. Dabei wird üblicherweise zwischen schwacher und starker Kollisionsresistenz unterschieden.

- Bei der schwachen Kollisionsresistenz darf es berechenbar nicht möglich sein, zu einem gegebenen Hashwert h(x) eine zweite Nachricht x‘ ≠ x zu finden, die den gleichen Hashwert h(x) aufweist.
- Demgegenüber darf es bei der starken Kollisionsresistenz einmal nicht möglich sein, zwei beliebige Nachrichten x und x‘ zu finden, die unter der Hashfunktion h abgebildet den gleichen Hashwert h(x) = h(x‘) aufweisen.

Man beachte, dass es aufgrund der Kompressionseigenschaft der Hashfunktion zwar immer Kollisionen geben muss, dass es aber berechenbar nicht möglich sein darf, solche zu finden. Eine Möglichkeit, Kollisionen zu finden, stellt in jedem Fall die vollständige Suche dar, d.h. man erzeugt eine sehr grosse Zahl von zufälligen Nachrichten und sucht nach Kollisionen unter den entsprechenden Hashwerten. Bei einer Hashfunktion, die Hashwerte der Länge n Bit erzeugt, wird man nach 2^n Versuchen eine zweite Nachricht gefunden haben, die den gleichen Hashwert aufweist, wie eine vorgegebene Nachricht. Wenn man nur zwei beliebige Nachrichten sucht, die unter der Hashfunktion abgebildet den gleichen Hashwert aufweisen, wird man – aufgrund des Geburtstag-Paradoxons der Wahrscheinlichkeitstheorie – im Schnitt bereits nach $2^{n/2}$ Versuchen erfolgreich sein. Deshalb verlangt man bei kryptografischen Hashfunktionen meist eine minimale Länge der erzeugten

Hashwerte von 128 bzw. 160 Bit. Die vollständige Suche nach einer Kollision hat dann offenbar einen Aufwand von $2^{128/2} = 2^{64}$ bzw. $2^{160/2} = 2^{80}$. Ein solcher Aufwand gilt heute als untere Schranke für das für einen Angreifer kaum mehr Praktikable.

Obwohl es in der Praxis meist ausreichend wäre, schwach kollisionsresistente Hashfunktionen einzusetzen, verlangt man für solche Funktionen meist starke Kollisionsresistenz. Sobald für eine kryptografische Hashfunktion algorithmisch eine Kollision mit einem kleineren Aufwand als mit der vollständigen Suche gefunden wird, gilt die Funktion als gebrochen. Entsprechende Resultate sind für MD5 und zuletzt auch für SHA-1 publiziert. Die Resultate haben in der internationalen Forschung eine intensive Suche nach Kollisionen und entsprechende Kollisionssuchalgorithmen auch für andere Hashfunktionen ausgelöst.

3 Implikationen

Die Brechung einer kryptografischen Hashfunktion bedeutet, dass die zur Diskussion stehende Hashfunktion nicht mehr als stark kollisionsresistent angenommen werden kann. Dabei ist es unerheblich, ob der gefundene Kollisionssuchalgorithmus praktikabel ist. Im Falle von SHA-1 hat der derzeit effizienteste Algorithmus z.B. einen Aufwand von 2^{63} (statt 2^{80}) und liegt damit immer noch deutlich über dem Aufwand einer vollständigen Schlüsselsuche für DES.

Die Tatsache, dass eine kryptografische Hashfunktion nicht mehr als stark kollisionsresistent angenommen werden kann, bedeutet, dass ein Angreifer mit einem Aufwand, der kleiner ist als die vollständige Suche, zwei Nachrichten finden kann, die den gleichen Hashwert aufweisen. Er könnte dann eine Nachricht digital signieren lassen und mit der Signatur und einer zweiten Nachricht vor den Richter treten, um Gültigkeit dieser Signatur für die zweite Nachricht einzufordern. Mit Hilfe der heute bekannten Kollisionssuchalgorithmen kann der Angreifer nur zwei zufällige Nachrichten finden. Damit lässt sich der skizzierte Angriff nicht durchführen. Um diesen Angriff wirklich durchzuführen, müsste der Angreifer nämlich zwei Nachrichten erzeugen können, die beide sinnvoll sind und von denen wenigstens eine harmlos wirkt. Von solchen Möglichkeiten einer dedizierten Suche nach Nachrichten mit spezifischen Inhalten ist man aber – auch im Falle von MD5 und SHA-1 – noch weit entfernt.

Letztlich hängt die Frage, ob ein Angriff möglich und realistisch ist, auch vom applikatorischen Umfeld ab. So werden kryptografische Hashfunktionen oft in Konstruktionen eingesetzt, die solche Angriffe a priori ausschliessen. Wird z.B. eine kryptografische Hashfunktion zur Abbildung von Passwörtern und deren „sicheren“ Speicherung in Passwortdateien eingesetzt, dann stellen Kollisionen kein grosses Problem dar (genutzt wird dann primär die Einweg-Eigenschaft der Hashfunktion). Auch die HMAC-Konstruktion, die in vielen kryptografischen Sicherheitsprotokollen im Internet (z.B. IPsec, SSL/TLS, …) eingesetzt wird, um Nachrichten zu authentifizieren, gilt als resistent gegenüber Kollisionsangriffen auf die eingesetzte Hashfunktion. Schliesslich schlagen Kollisionsangriffe auch nicht durch, wenn – wie heute vielfach üblich – zwei (oder mehr) Hashfunktionen kombiniert eingesetzt werden.

4 Schlussfolgerungen und Ausblick

Aufgrund des Gesagten muss vom Einsatz von als gebrochen geltenden kryptografischen Hashfunktionen (insbesondere MD5 und SHA-1) grundsätzlich abgeraten werden. Wie kritisch der Einsatz einer solchen Hashfunktion aber wirklich ist, hängt im Einzelfall von der Applikation ab. Falls man die Möglichkeit hat, ist der Einsatz alternativer Hashfunktionen oder der gleichzeitige Einsatz zweier (oder auch mehrerer) Hashfunktionen angebracht. Als Alternativen zu MD5 und SHA-1 bieten sich insbesondere die Hashfunktionen

aus der SHA-2-Familie an (insbesondere SHA-224, SHA-256, SHA-384 und SHA-512, wobei sich die Zahl hinter der Abkürzung SHA auf die jeweilige Länge des erzeugten Hashwertes bezieht). Weil diese Funktionen längere Hashwerte erzeugen, sind sie auch resistenter gegenüber Kollisionsangriffen. Der gleichzeitige Einsatz verschiedener Hashfunktionen ist in vielen Applikationen und Standards vorgesehen. Die Wahrscheinlichkeit, dass ein Angreifer Nachrichten findet, die unter allen Hashfunktionen Kollisionen verursacht, ist hinreichend klein und kann meistens vernachlässigt werden.

Zur Zeit führt das U.S. amerikanische NIST einen offenen Wettbewerb durch, um – unter dem Arbeitstitel SHA-3 – einen Nachfolgestandard für SHA-1 zu küren. Die Benennung des Standards ist für 2012 vorgesehen. Ab diesem Zeitpunkt ist ein konsequenter Wechsel auf SHA-3 angebracht.

Abkürzungen

DES	Data Encryption Standard
HAMC	Hashed MAC
IPsec	IP Security
NIST	National Institute of Standards and Technology
MAC	Message Authentication Code
MD5	Message Digest Algorithm 5
SHA	Secure Hash Algorithm
SSL	Secure Sockets Layer
TLS	Transport Layer Security