アルゴリズムへの攻撃――ハッシュ攻撃の現状

佐々木（ささき）　悠（ゆう）　/青木（あおき）　和麻呂（かずまろ）
藤岡（ふじおか）　淳（あつし）
NTT情報流通プラットフォーム研究所

2004年にMD5の衝突が発見されて以来，ハッシュ関数の解析やその脆弱性を利用した実応用プロトコルへの攻撃法が進化し続けています．本稿では，これらの進展を解説し，次世代ハッシュSHA-3選定状況について紹介します．

本研究のねらい

NTT情報流通プラットフォーム研究所では暗号理論に関する研究を行っています．最先端の研究を行い，高い技術力をアピールするとともに，NTTグループが提供する製品の安全性向上に寄与することが目的です．本稿では，多岐にわたる暗号理論の研究テーマの中で，共通鍵暗号理論研究の進展，特に，昨今話題となっているハッシュ関数の安全性動向について解説します．

ハッシュ関数とは

ハッシュ関数は，任意長の入力情報に対し，ハッシュ値と呼ばれる比較的短い固定長の文字列を出力する関数です．ハッシュ関数は，情報通信の完全性・秘匿性を担保します．実際に利用するにあたり，安全かつ効率的に演算可能であることが求められています．

ハッシュ関数の利用例を**図1**，**2**を用いて説明します．図1は巨大なファイルの一致を検証する目的にハッシュ関数を利用することを表しています．巨大なファイルどうしの比較は時間がかかり非効率です．そこで，あらかじめファイルのハッシュ値を計算・記憶しておくことで，短いデータの比較だけでファイルの一致を検証することができます．

図2はメールサーバにアクセスするユーザのリモート認証にハッシュ関数を用いることを表しています．このようなプロトコルでは，ユーザとサーバは事前に秘密のパスワードを共有しておき，ユーザがサーバにアクセスするたび

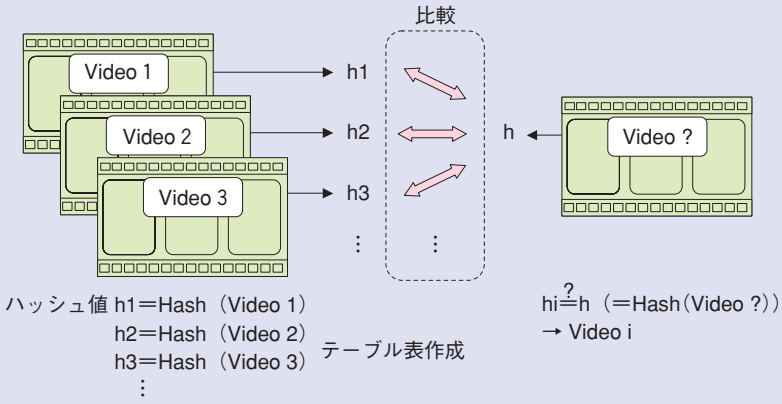


図1　ハッシュ関数の利用例1

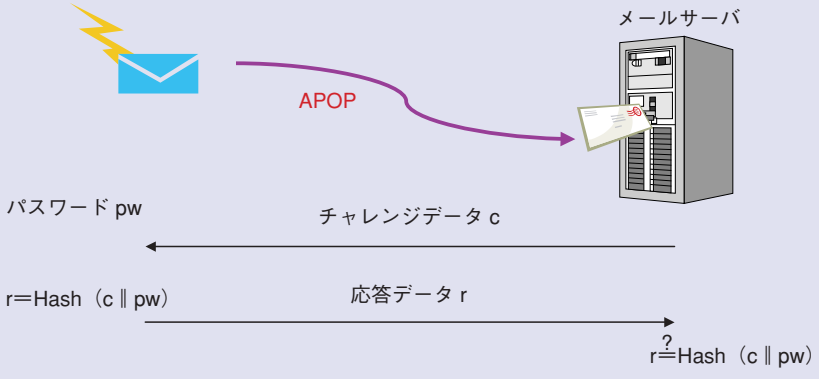


図2　ハッシュ関数の利用例2

に，パスワードを用いた行動をし，秘密のパスワードを保持していることを証明します．この際，秘密のパスワードをそのまま通信してしまうと，悪意のある第三者の盗聴によりパスワードが漏洩してしまう可能性があります．そこで，認証のたびにサーバがチャレンジと呼ばれるランダムに生成されたデータをユーザに送り，ユーザはこのチャレンジとパスワードを組み合わせたデータのハッシュ値を計算し，このハッシュ値を応答データとしてサーバに送り返します．このようにすることで，たとえ応答データを盗聴されても，パスワードを知られることがないようにすることができます．

ハッシュ関数の安全性

暗号学的ハッシュ関数は，「衝突困難性」「原像回復困難性」と呼ばれる2つの安全性を満たす必要があります．衝突困難性とは，ハッシュ値が同じとなる異なるデータを求められないことを意味し，原像回復困難性とは，ハッシュ値からその値となる元データを求められないことを意味します．これらの安全性が守られない場合，ハッシュ関数を利用した方式で問題が生じる可能性があります．図1の利用例では，ハッシュ値が同じとなる異なるファイルが存在すると，その両方が一致するファイルとして検証に合格してしまいます．図2の利用例では，盗聴された応答データから元データを求めることができると，パスワードに関する何らかの情報を取得されるおそれがあります．

これら2つの安全性を**図3**にまとめます．一般に，ハッシュ関数の出力長をnビットとしたとき，衝突困難性は$2^{n/2}$回ハッシュ値を計算する能力がある攻撃者がいると破れてしまい，また，原像回復困難性は2^n回ハッシュ値を計算する能力がある攻撃者がいると破れてしまいます．一方で，衝突困難性が破れたことによる影響を受けるプロトコルは限定されますが，原像回復困難性が破れた場合には，あらゆるプロトコルがその影響を受けます．

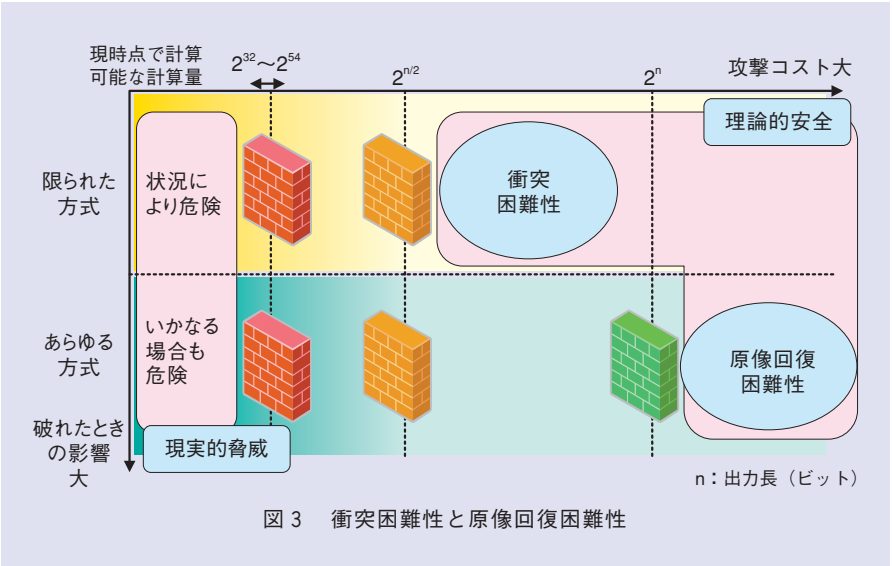


図3　衝突困難性と原像回復困難性

MD5の衝突とその応用

MD5は1990年代に標準化され，現在でも世界中で利用されているハッシュ関数です．2004年に中国の研究者によってMD5の衝突が発見されて以来，衝突が実際に利用されている製品やプロトコルに対してどのような影響を与えるかを明らかにする研究が進められてきました．その中の1つがAPOP（Authenticated Post Office Protocol）という，ユーザがメールサーバにアクセスするときの認証プロトコルに関する研究です．2007年，フランスの研究者により，MD5の衝突を用いると，攻撃者がメールサーバになりすますことでユーザのパスワードを3文字まで復元できるという攻撃が示されました（**図4**）．この攻撃では，多くのユーザがメール取得を短い間隔で自動的に行うことを利用し，ユーザに攻撃を受けていることを検知されずにサーバになりすますことができる性質を利用しています．しかし，復元可能なパスワードの文字数は，MD5の衝突攻撃に起因する性質により3文字を超えることがないという制限があったため，攻撃の影響は限定的だと思われていました．そこで，我々はMD5の衝突攻撃を抜本から見直し，APOPの攻撃に特化するように再構成することで，パスワードを31文字まで復元できることを発見し，事実上ほぼすべてのパスワードを復元できることを示しました．

さらには，SIP（Session Initiation Protocol：セッション確立プロトコル）にも同様の攻撃が適用できる可能性があり，実際にSIPを実装したいくつかの特定の端末で攻撃を確認しました．その後，これらの攻撃の存在や攻撃を防ぐ対策を関係部署に周知し，NTTが提供するサービスが第三者により攻撃されないよう手を尽くしました．

原像回復困難性

前述のとおり，2004年以降，MD5をはじめとするさまざまなハッシュ関数で衝突困難性に対する安全性が低下しましたが，原像回復困難性についても

ほとんど研究の進展がないという状況でした．最近では**図5**で示されるように，安全ではなくなったMD5の利用を停止し，より安全なハッシュ関数であるSHA-1やSHA-2を利用しようという移行が進んでいます．しかし，より安全なハッシュ関数を選択するためには，衝突困難性の評価だけではなく，原像回復困難性についての評価も必要であるため，原像回復困難性の高さを測る手法が必要でした．

そこで，我々は「中間一致型原像回復攻撃」という新しい評価方法を開発し，実際に利用されているさまざまなハッシュ関数の安全性を評価しました．中間一致型原像攻撃の概要を**図6**に示します．中間一致型攻撃では，攻撃者は入力と出力の対応を直接考えるのではなく，入力とある中間値との対応・出力とある中間値との対応を独立に求め，双方から導かれた中間値が一致する組を探すことで，全体として原像となっているデータを探します．この方法を用いると，素朴な方法であれば2^nの演算コストがかかる原像回復攻撃が，もっとも条件が良いとき，$2^{n/2}$の演算コストで実行できるようになります．

図7は我々の攻撃をさまざまなハッシュ関数に適用した結果をまとめたものです．図7に示すとおり，多様なハッシュ関数に対して，世界初の成果や既存の研究よりも優れた成果を得ることに成功しました．特に，MD5に対してそれが原像回復困難性を満たしていないことを世界で初めて示すことに成功し，暗号業界に大きなインパクトを与えました．また，SHA-1・SHA-2など，MD5からの移行が進み，今後利用が増えることが予想されるハッシュ関数に関して，もっとも有効な評価を与えることができ，NTTが最先端の評価技術を有するという点で，大きなプレゼンスを獲得しました．

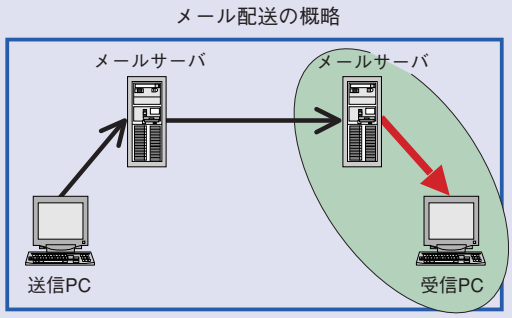


図4　APOPに対するなりすまし攻撃の設定

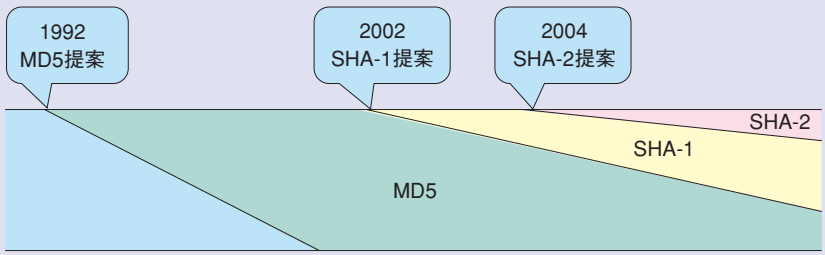


図5　利用されているハッシュ関数の移行

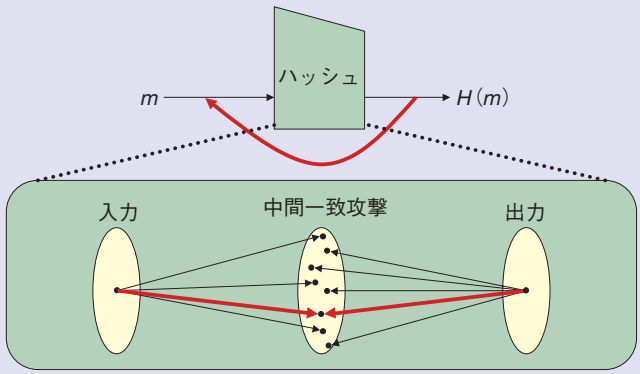


図6　中間一致型原像回復攻撃

次世代ハッシュ関数標準化動向

これまで述べたように，最近の解析技術の進歩により，過去に設計されて

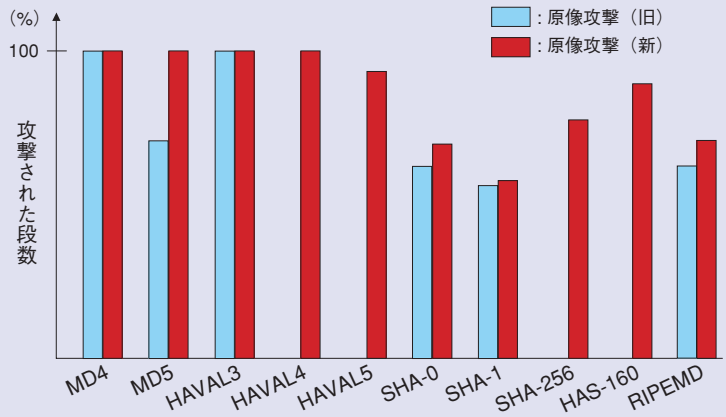


図 7　原像回復困難性の適用結果

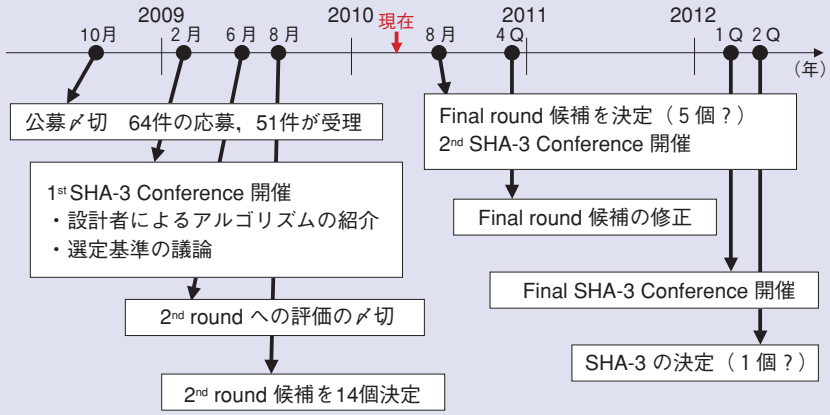


図 8　SHA-3コンテストのスケジュール

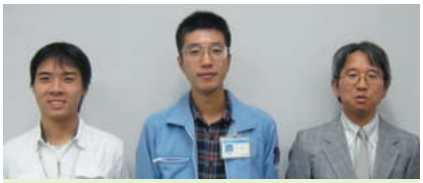
きたハッシュ関数はその安全性が危ぶまれています．MD5はもとより，2002年に提案されていたSHA-1に対してもすでに理論的な攻撃が見つかっています．また，現在移行が進んでいるSHA-2に関しては，今のところは（100%の段数を破る）攻撃方法が見つかっていませんが，SHA-2の設計方針はMD5やSHA-1の設計方針と同じであるため，遠くない未来に急に攻撃方法が見つかる可能性を否定できない状況です．

このような背景から，NIST（National Institute of Standards and Technology：米政府標準技術局）はSHA-3と呼ばれる次世代の標準ハッシュ関数を定めるプロジェクトを進めています．このプロジェクトは「SHA-3コンテスト」と呼ばれ，安全性と実装効率を兼ね備えたハッシュ関数を世界中から公募し，外部研究者による評価を交えながら5年がかりでSHA-3を選定するという大規模なプロジェクトです．SHA-3コンテストでどのハッシュ関数がSHA-3として採用されるかについては，暗号学者の間だけでなく，暗号・セキュリティ商品のベンダ・開発者などの間でも注目の的となっています．

SHA-3コンテストの進行スケジュールを**図 8** に示します．SHA-3の候補アルゴリズムは2008年10月に募集が行われ，64個のアルゴリズムが提案されました．2009年11月には，14個のアルゴリズムが候補として残っており，2012年までかけて正式なSHA-3を選定する予定となっています．NTTでは，SHA-3コンテストに独自のアルゴリズムは提出していませんが，外部からの評価というかたちで参加し，適切なアルゴリズムが次世代標準であるSHA-3として選ばれるように努力しており，これまでに，2つの候補アルゴリズムに対して攻撃が存在することを示し，このコンテストに貢献しています．

今後の展開

アルゴリズムの設計技術と攻撃技術は表裏一体の関係にあります．今後はハッシュ関数の攻撃手法の研究を継続するとともに，革新的な設計技術を開発し，NTTの高い技術力をアピールしていきます．



（左から）佐々木　悠/ 青木　和麻呂/
藤岡　淳

暗号研究は情報セキュリティの基盤です．今後も革新的な成果の創出に励んでいきます．

◆問い合わせ先
NTT情報流通プラットフォーム研究所
情報セキュリティプロジェクト
セキュリティプラットフォームグループ
TEL　0422-59-3471
FAX　0422-59-4015
E-mail　sasaki.yu@lab.ntt.co.jp