

JavaScript による **End-to-End** セキュリティ

第 4 回 データの真正性・本人確認のためのテクニック 編

栗原 淳

2019 年 10 月 31 日

はじめに

はじめに

第 1,2,3 回では

- End-to-End (E2E) セキュリティの原則と必要性
- JavaScript で AES を使った暗号化のお作法
- JavaScript で公開鍵暗号 (RSA/楕円曲線) を使った暗号化のお作法

を勉強した。

今回は、第 3 回の最後に懸案事項だった

「データのやり取りしてる相手って本当に正しい相手？」
を保証する方法を学んでいく。

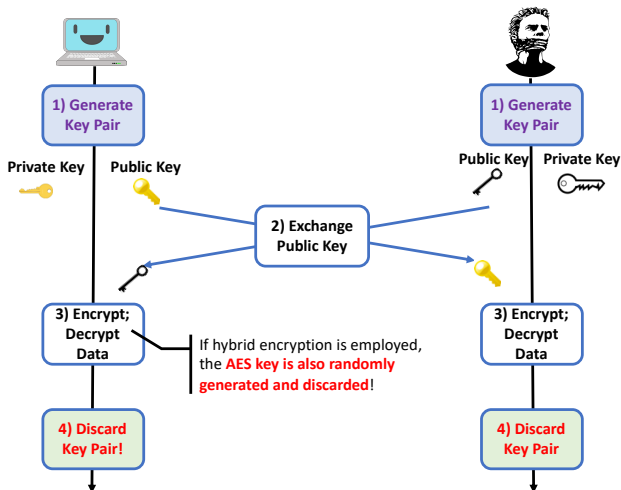
第3回のおさらい: Ephemeral Scheme

公開鍵暗号化の Ephemeral Scheme での運用

公開鍵・秘密鍵ペアを都度生成、1 回限りで使い捨てることで、**Perfect Forward Secrecy**¹ を担保する運用方法。

Perfect Forward Secrecy を守り、End-to-End 暗号化の強固な運用を。

¹長期的に保存されているマスター秘密鍵の漏洩や、一部の暗号化データがクラックされたとしても、**それ以外の過去に暗号化されたデータは復号されてしまうことはない**という概念。



Ephemeral Scheme のイメージ

まずはじめに、「送られてきた Ephemeral な公開鍵は、本当に自分がやりとりしたい相手の公開鍵か？」の確認が必須。²

²意図しない相手の公開鍵で暗号化して機密データを漏らさぬように、ということ。

というわけで、「E2E で安全にデータをやり取りする」ための基礎部分の最後のピースを今日は学ぶ。

この講義で最終的に学びたいこと

- 本人確認やデータの改ざん防止を担保する方法
 - データ毎に固有の指紋を生成する「ハッシュ」
 - 「共通鍵」を使った改ざん防止方法「MAC」³
 - 「公開鍵」を使った本人確認・改ざん防止方法「電子署名」⁴
- そしてその具体的な JavaScript での実装方法・お作法

細かい話もするが、数式は使わない。

「イメージ」と「コードの流れ&その流れの必要性」をつかめるようにする。

³HMAC (RFC2104 <https://tools.ietf.org/html/rfc2104>)

⁴RSASSA PKCS#1-v1.5/PSS (PKCS#1 RFC8017 <https://tools.ietf.org/html/rfc8017>), ECDSA (FIPS PUB186-4 <https://csrc.nist.gov/publications/detail/fips/186/4/final>)

栗原 淳 (Jun Kurihara)

- (株) ゼタント 主任研究員
(株) 国際電気通信基礎技術研究所 (ATR) 連携研究員
- 博士 (工学),
専門: セキュリティ、応用数学、システムアーキテクチャとか
- Web システム (フロントエンド・バックエンド) を作ったり、
論文他のアルゴリズムを実装したり、研究して論文書いたり、
セキュリティ技術中心に手広くやっています。
- GitHub: <https://github.com/junkurihara>
LinkedIn: <https://www.linkedin.com/in/junkurihara>

この講義の対象と事前準備

対象:

- 暗号・セキュリティ技術に興味がある初学者
- Web に暗号技術を導入したい Web 系のエンジニア

必須ではないが触って楽しむのには必要な事前準備:

- Bash, Git が使えるようになっていること
- Node.js, npm, yarn が使えるようになっていること
- Google Chrome 系ブラウザ and/or Firefox が利用可能なこと

今後の予定 (暫定)

- 1 導入&JS の暗号化コードを触ってみる
- 2 AES を正しく・安全に暗号化するには？
- 3 公開鍵暗号はどうやって使う？その使い方のコツは？
- 4 ハッシュ・MAC・署名、それぞれの使い所と使い方は？ ← 今日はココ
- 5 RFC にまつわるあれこれ（証明書・鍵フォーマット・etc... 未定。）

「こういうのを知りたい」というリクエストがあれば是非。

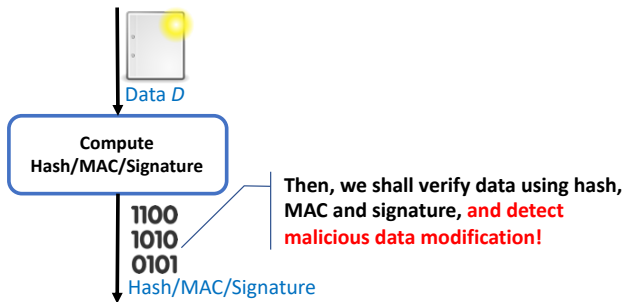
セカンドシーズンも検討中。⁵

⁵場所等変えてもっと来やすい場所へ。。。

サンプルコードの準備

準備

説明を聞きつつ手を動かすため、まず環境準備。今回は、JavaScript (Node.js) を使って手元でデータの Hash/MAC/署名をいじってみる。そしてその効果を実感する。



※サンプルコードはブラウザでも動く。

src/commands-browser.html を開くとこれから Node.JS で試すデモが開発者コンソールで実行される。適宜試したり比較すると良い。

※前回のコードの公開鍵に署名をつけたりして Ephemeral Scheme を作ってみると良い。

以下の環境が前提:

- Node.js (> v10) がインストール済。yarn が使えること。⁶
- ブラウザとして、Google Chrome (系ブラウザ)、もしくは Firefox がインストール済み
- Visual Studio Code や WebStorm などの統合開発環境がセットアップ済みだとなお良い。

⁶インストールコマンド: `npm i -g yarn`

JavaScript プロジェクトの準備

1 プロジェクトの GitHub リポジトリ⁷ を Clone

```
$ git clone https://github.com/zettant/e2e-security-04  
$ cd e2e-security-04/sample
```

2 依存パッケージのインストール

```
$ yarn install
```

3 ライブラリのビルド

```
$ yarn build
```

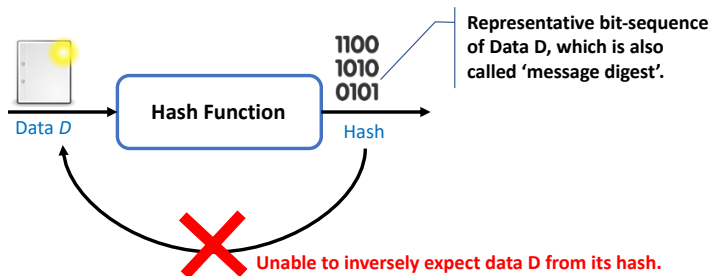
⁷<https://github.com/zettant/e2e-security-04>

データの指紋: Hash

Hash および Hash 関数とは

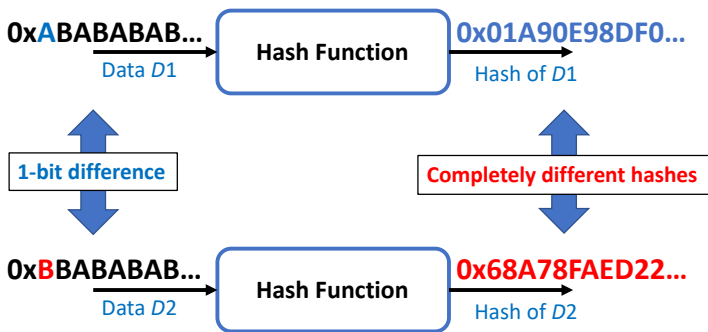
Hash および Hash 関数

あるデータに対し、そのデータを「代表するビット列」を計算する不可逆の関数を「Hash 関数」。導出したビット列を「Hash」⁸と呼ぶ。



⁸あるいは Hash 値、Message Digest

1 ビットでもデータが異なれば、全く違う Hash が導出される。



Hash および Hash 関数の役割

同じくデータ固有のビット列を導出する Checksum と似ているが、その用途はより強力で多岐にわたる。

■ Checksum

- 通信路上などでのデータの (偶発的な) エラー検知

⇒ データから一意に導ける値・高速な処理が可能なのが必須

■ Hash

- データのエラー・改ざん検知
- Hash をデータ実態の代替として署名を生成
- 多数のデータの索引作成⁹
- データの重複検出

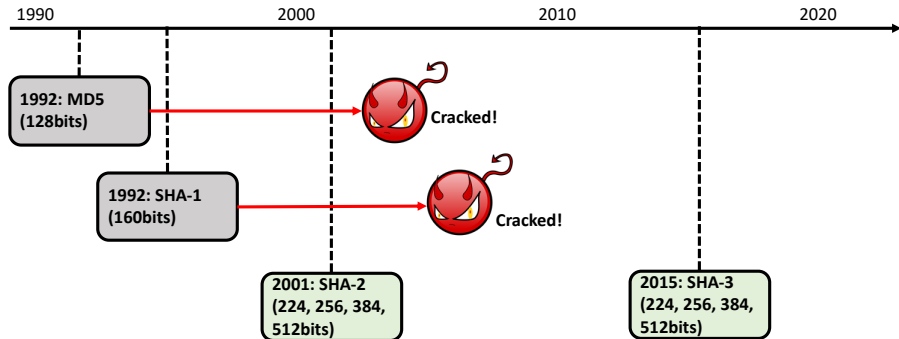
⇒ 別のデータ同士で同じ Hash を得ることが困難なのが必須

Checksum \subseteq Hash と言える。

⁹Hash Table

Hash 関数の種類

MD5, SHA-1, SHA-2 (SHA-256, 384, 512), SHA-3 という Hash 関数がよく知られている。



MD5、SHA-1 は、「**同じ Hash(指紋) を生成するデータが割と簡単に見つけられる¹⁰**」という致命的な欠陥が発見されている。

¹⁰「衝突」と呼ぶ。MD5 の場合は、 2^{20} 程度の計算量でクラック可能。

Hash 関数の選択について

- 理由がなければ SHA-2 シリーズ以降のものを選択する。
 - bit 長は長いほど、衝突するデータが見つけづらい (=強固)
 - ただし、bit 長が長いほど、計算が重くなる
- SHA-1/MD5 は、基本的に互換性の担保のためだけに利用する。但し、Checksum として使う分には概ね問題ない。何が何でも使うな、というわけではない。

IE/Edge こぼれ話

X.509 の公開鍵証明書などはまだ SHA-1 が利用されている場合が多々ある。しかし、IE/Edge では互換性の担保を全て無視して SHA-1 のネイティブサポートを全打ち切りしているので、X.509 公開鍵証明書などを JavaScript からネイティブ API を通して扱えない。

JavaScript でデータの **Hash** を生成してみる。

Hash 生成のコードはこんな感じ。

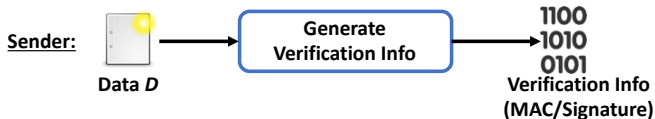
本人確認の技術

「正しい相手から正しく送信されてきたデータか」？

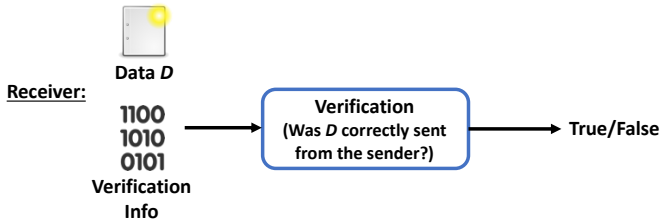
いわゆる「データの真正性と送信元の確認方法」には、大まかに 2 つの方法がある。

- Message Authentication Code (MAC)
- 署名 (電子署名)

両者とも、送信するデータから MAC/署名という検証用データを生成、元データに付与する形で送信。



受信したデータと、MAC/署名とを突合して、「送信元は意図している相手か?」「データは改ざんされていないか?」を検証。

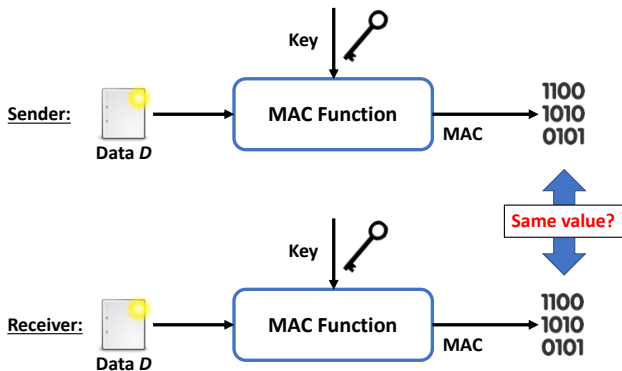


MAC・署名の中身に突っ込む前に、それぞれのざっくりとした定義と pros/cons を説明する。

Message Authentication Code (MAC)

MAC によるデータ真正性と送信者の確認

- 送信側・受信側で共有する鍵を使ってデータ・鍵固有のバイナリ (MAC) を生成する方法。
- 受信側で、送信側と同一の MAC が作れるかどうかをチェック。



MAC の特徴:

- 同じ鍵でも、データが異なれば出力される MAC も異なる。
- 同じデータでも、鍵が異なれば出力される MAC も異なる。



すなわち、受信側で同一の MAC が作れることを確認できれば、

- 鍵を共有する相手から
- 途中の改ざんなしで送られたデータであること

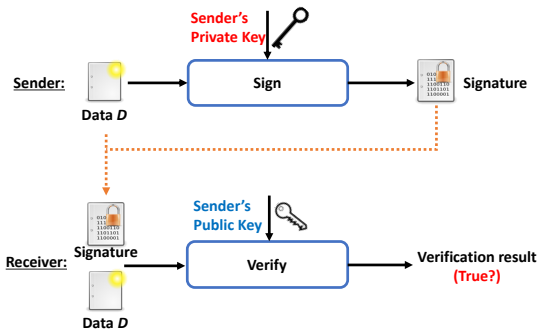
が保証される。

署名 (電子署名)

公開鍵・秘密鍵ペアをベースとした技術¹¹：

署名によるデータ真正性と送信者の確認

- 送信側は公開鍵・秘密鍵ペアを保有。
- 送信側は、データと自分の秘密鍵から署名を生成。
- 受信側は、受信データ、署名と公開鍵の間の一貫性をチェック。



¹¹ ここでいう公開鍵・秘密鍵ペアは、公開鍵暗号化に使うものと全く一緒の概念。

署名の特徴:

- データが改ざんされていたら、検証が失敗 (false が出力)。
- 意図する相手の秘密鍵¹² で署名が作られていなければ、検証が失敗。
- **MAC と違って、秘密の情報 (=鍵) を事前共有しなくて良い**



すなわち、署名技術は、

- 意図する送信者から
- 途中の改ざんなしで送られたデータなことを
- **事前の秘密情報の共有なしで**

保証する。¹³

¹² 自分が入手している公開鍵の対となる秘密鍵

¹³ 検証用の公開鍵は、信頼できる手段で入手済み、あるいはプリインストールされていると仮定

MAC と署名の pros/cons

じゃあ署名だけで MAC は不要では？…そういうわけにはいかない。

	Pros	Cons
MAC	<ul style="list-style-type: none">・ 一般的に 高速¹⁴・ 生成する MAC サイズは小さい¹⁵	<ul style="list-style-type: none">・ 鍵の 事前共有が必要
署名	<ul style="list-style-type: none">・ 鍵の 事前共有が不要	<ul style="list-style-type: none">・ 一般的に 非常に遅い・重い・ 生成する署名サイズは一般的に 大きい¹⁶

⇒ AES/公開鍵暗号の関係と全く一緒に、使い所を考えて組み合わせて使う、もしくは場合に応じて使い分ける。

¹⁴ AES (CMAC) とか Hash (HMAC) とかを構成要素としているため。

¹⁵ 通常 128–512bits 程度。

¹⁶ ECDSA は小さく、256–512bits 程度。RSA 系は非常に大きく通常 2048bits 以上。

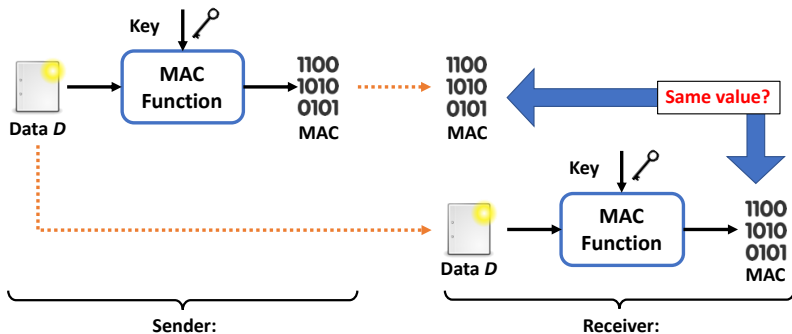
共通鍵を使った改ざん検知・本人確認: MAC

Message Authentication Code (MAC) 事始め

MAC を使った改ざん検知&本人確認手続き

送信側、受信側で秘密の鍵 (バイナリ列) を共有。

- 1 送信側はデータと一緒に、データと鍵から生成した MAC を送信。
- 2 受信側は、鍵と受信したデータから、受け取った MAC と同じものが作れるかどうかをチェック。



MAC を作る標準手法のバリエーション。

- HMAC; Hash-based Message Authentication Code
- CMAC; Cipher-based Message Authentication Code
- GMAC; Galois Message Authentication Code
- etc.

今回は、JS で一番使いやすいと思われる HMAC を取り上げる。

HMAC; Hash-based MAC

HMAC (RFC2104)¹⁷

- 鍵付き Hash¹⁸ と呼ばれる、Hash 関数ベースの MAC 生成方法。
- HDKF (RFC5869) などの標準技術や、AWS Signature v4¹⁹ 等、各所で利用されている。

¹⁷<https://tools.ietf.org/html/rfc2104>

¹⁸Keyed Hash

¹⁹AWS S3 にクライアントから REST API 経由でアップロードする時に一時的に生成する MAC

「鍵」と「データ」をまとめて Hash 関数に入れる、と考えると、

- 鍵・データ両者が正しくないと、正しい Hash も生成不能 (=MAC 検証失敗)。
- MAC から鍵・データの情報を逆算することはできない。

という特徴をイメージしやすい。

JavaScript で HMAC を実行してみる

コードの中身はこんな感じ。

その他の MAC (JS じゃビミョー…)

CMAC; Cipher-based MAC (NIST SP800-38B²⁰)

共通鍵暗号 (e.g., AES) の CBC モードを Hash 関数がわりに使用して MAC を計算する。「前のブロックの暗号文を使って次のブロックを暗号化する」という特徴を応用。

GMAC; Galois MAC (NIST SP800-38D²¹)

共通鍵暗号 (e.g., AES) の Galois Counter Mode (GCM) で暗号化と同時に生成される MAC。高速に計算できる代数演算²²を Hash 関数がわりに使用して MAC を計算する。GMAC 単独で利用可。

※ CMAC/GMAC 共々暗号化と同時に計算されることが多く、単独で利用するケースはあまり見かけない。

²⁰ <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-38b.pdf>

²¹ <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-38d.pdf>

²² $\mathbb{F}[x]/(x^{128} + x^7 + x^2 + x + 1) = \mathbb{F}_{2^{128}}$ 上の乗算

と、「標準技術」で「広く利用されている」MAC アルゴリズムはあるが、JS のネイティブ API²³ でサポートされている MAC は、現状 HMAC のみ…

CMAC, GMAC が使いたかったら自力実装 or npmjs.com で見つけて利用する。

²³WebCrypto API, Node.js Crypto

公開鍵を使った改ざん検知・本人確認: 署名

署名 事始め

署名 (電子署名) を使った改ざん検知&本人確認手続き

受信側は、送信側の公開鍵を予めプリインストール。

■ 送信側の処理:

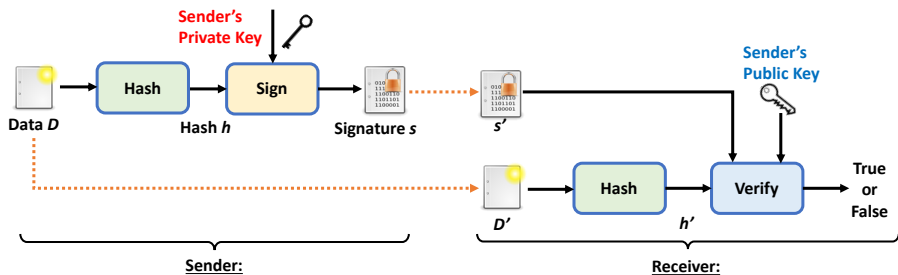
- 1 データ D を Hash 関数で短縮²⁴、hash $h = \text{Hash}(D)$ を導出。
- 2 hash h に対して秘密鍵 SK で署名 $s = \text{Sign}(h, SK)$ を生成。
データ D と署名 s とを一緒に受信側へ送付。

■ 受信側の処理:

- 1 データ D の hash $h = \text{Hash}(D)$ を導出。
- 2 hash h と署名 s の一貫性を、公開鍵 PK で検証。
 $\text{Verify}(h, s, PK) \in \{\text{True}, \text{False}\}$

²⁴データ D そのものに直接署名を施すのは計算量的・データ量的に大変 (e.g. 元データと同じかそれ以上の大きさの署名を作る羽目になる) なので、**データの指紋 (i.e., hash) に対して署名を施す。**

ざっくりフロー図。



このフローは、以下のように考えるとイメージがつきやすい²⁵

- 1 送信側は、hash h を秘密鍵で暗号化して s を生成。
- 2 受信側は、 s を公開鍵で復号して h' を入手。
命題「 $h' = \text{Hash}(D')$ 」が成立するか検証。

²⁵ 但し、常に正しい表現ではないので注意。

署名生成方式の標準方式のバリエーション。

- RSA 暗号をベースとした手法:
 - RSASSA PSS
 - RSASSA PKCS#1-v1.5
- 楕円曲線暗号をベースとした手法:
 - ECDSA
- etc.²⁶

JS で使いやすい RSASSA PSS & PKSC#1-v1.5 と ECDSA について取り上げる。

²⁶Digital Signature Algorithm; DSA (FIPS PUB 186-4
<https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.186-4.pdf>) など

細かく説明する前に…

署名において一番やばいのは、「誰かが自分を騙って署名を作ることができる」などという状況。そのため、署名においても「公開鍵サイズの正しい選択」が重要。

前回の復習

暗号化において AES-128bits と同程度の強度の公開鍵サイズは、

- RSA: $\geq 3072\text{bits}$
- ECC: $\geq 256\text{bits}$

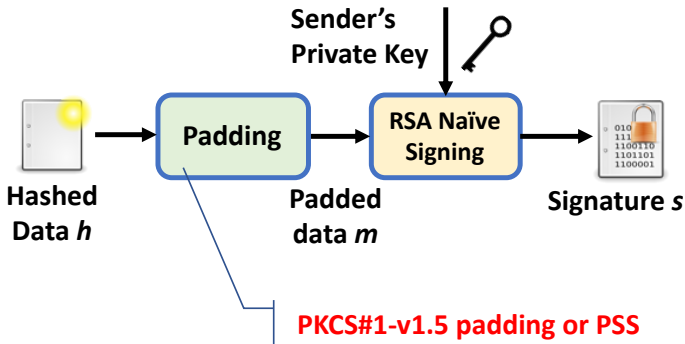
程度。

署名についても暗号化に求めるものと同程度、あるいはそれ以上の鍵サイズを担保して、公開鍵から秘密鍵を作られる (i.e., 自分を騙られる) ことを防ぐ。

RSASSA; RSA Signature Scheme with Appendix

RSASSA PKCS#1-v1.5 と、RSASSA PSS の違い

暗号化と同様に、 $h = \text{Hash}(D)$ に対して署名を作る際、鍵長に合わせたパディングが必要。そのパディングの方法が違う。



RSASSA PKCS#1-v1.5 (RFC8017²⁷)

- RSAES PKCS#1-v1.5 の署名版。PCKS#1 の v1.5 で標準化。SSL/TLS を含め、現在でも広く使われている。
- hash $h = Hash(D)$ を DER エンコードしたデータ T に、公開鍵長まで以下のようなパディングを付与。

$$m = 0x00||0x01||RandomSequence||0x00||T$$

- RSASSA PKCS#1-v1.5 は、hash ではなく生の短いデータに対して署名を生成すると、署名からその元データを戻せちゃうという脆弱性²⁸が知られている。本人確認の意味なし。
- PCKS#1 v2.2 (RFC8017) では、「仕様通りに hash 使っていればとりあえず問題ないが、今後何があるか分からないので **RSASSA-PSS を使うほうが無難**」と記載。

²⁷<https://tools.ietf.org/html/rfc8017>

²⁸Y. Desmedt et al., “A Chosen Text Attack on the RSA Cryptosystem and Some Discrete Logarithm Schemes,” in *Proc. CRYPTO 1985*, pp.516–522, 1985.、他。J. Coron (CRYPTO 1999) 等。

RSASSA-PSS (Probabilistic Signature Scheme, RFC8017²⁹)

- RSAES OAEP の署名版。RFC3447/PKCS#1 v2.1 (2003 年) で標準化策定。
- 署名対象データ D に対して、乱数成分 (RandomSalt) を加えつつ Hash 関数を 2 重掛けし、Padded data m を作っている。

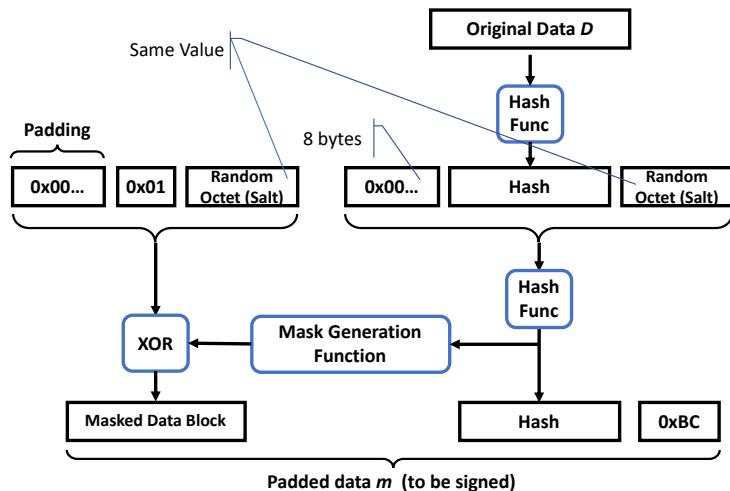
$$\begin{aligned}h &= \text{Hash}(\text{Hash}(D), \text{RandomSalt}), \\m &= \text{MaskedDataBlock} || h || 0x\text{BC},\end{aligned}$$

MaskedDataBlock はヘッダ。

- RSASSA PKCS#1-v1.5 からの乗り換えが進みつつあるものの **OAEP 同様に未対応の暗号ライブラリも多い**。ただし、RFC 標準的には、可能なら PSS を利用することを推奨。

²⁹<https://tools.ietf.org/html/rfc8017>

PSS のイメージ図



元データ D に加え、padded data m があれば Salt などを逆計算可能。

JavaScript で RSASSA-PSS を実行してみる

というわけで、RFC 推奨となっている PSS を試してみよう。

コードの中身はこんな感じ。

RSASSA-PSS は、Node.js Crypto/WebCrypto 共にネイティブサポートされている。(WebCrypto はブラウザ次第)
しかし他の環境だと、**OAEP 同様、PSS に未対応なライブラリも多く注意が必要**。(最近まで OpenSSL も未対応だった)

ECDSA; Elliptic Curve Digital Signature Algorithm

ECDSA (NIST FIPS 186-4³⁰)

- ECDH の署名版³¹。NIST FIPS 186-3 (2009 年) で標準化。
Bitcoin blockchain にも使われていて、また注目度が上がった。
- RSASSA とは異なり、事前の padding に気を使う必要がない。
アルゴリズム内部で hash h を生成・利用している。すなわち、

$$\text{Signature} = \text{SignECDSA}(D, SK),$$
$$\{\text{True}, \text{False}\} \ni \text{VerifyECDSA}(\text{Signature}, D, PK),$$

という直接の API がアルゴリズムレベルで提供される。³²

- OpenSSL をはじめほとんどの現代的な暗号ライブラリが実装をサポートしている。

³⁰<https://csrc.nist.gov/publications/detail/fips/186/4/final>

³¹正しい表現ではないが、イメージでそう捉えて欲しい。

³²RSASSA も、普通のライブラリでは直接 Sign/Verify 可能な API が提供されるが、その実態は RSA のナイーブな署名生成・検証アルゴリズムのラッパー。

RSASSA と ECDSA の比較。(一般論。実装にもよる。)

	Pros	Cons
RSASSA	<ul style="list-style-type: none">・署名検証が高速。 ECDSA と比較してもかなり速い。	<ul style="list-style-type: none">・署名サイズが大きい。 公開鍵長に等しい (e.g., 3072bits)。・署名生成がかなり低速。
ECDSA	<ul style="list-style-type: none">・署名サイズが小さい。 公開鍵長の倍 (e.g., 512bits (256bits key))。	<ul style="list-style-type: none">・署名生成・検証共に同じくらい低速。

どっちを使えばいいかは利用環境に応じて選択したほうがいい。常に ECDSA を使えばいいとかそういうわけではない。

JavaScript で ECDSA を実行してみる

MAC/署名の運用について

署名検証のブートストラップの問題

Q: 署名の検証用の公開鍵が正しいことはどうやって保証するの？

⇒ 現状ではこの問題に対して**仮定なしで OK な万能の解は未知**。
どこかに**仮定、信頼するという起点**を置いた暫定解がある感じ。

暫定解 1: Trust Anchor

End ごとに、**検証用の公開鍵をアプリケーション・端末に固定で事前に埋め込んでおく**。そこだけは起点として無条件に信頼して使う。

※ SSH とか GitHub で行う「ホスト・サービスへの公開鍵の登録」、というのはこの Trust Anchor の登録になる。

暫定解 2: PKI に頼る

さらに Verisign とかに検証用の公開鍵に署名してもらって、公開鍵証明書を作る。公開鍵証明書自身はオンラインで取得するが、その Verisign の署名の検証は Trust Anchor…(Verisign を信頼の起点にする)。

解 2 は、Verisign の鍵 1 つを Trust Anchor にすればいいので鍵を埋め込むような手間がかからないが、お金がかかる³³。

³³ 賛否あるだろうが、Let's encrypt (<https://letsencrypt.org/>) を使えば無償でいける。ただし Let's encrypt は厳密な本人確認とせずに署名つけてくれるので…

署名・MAC の使い分け

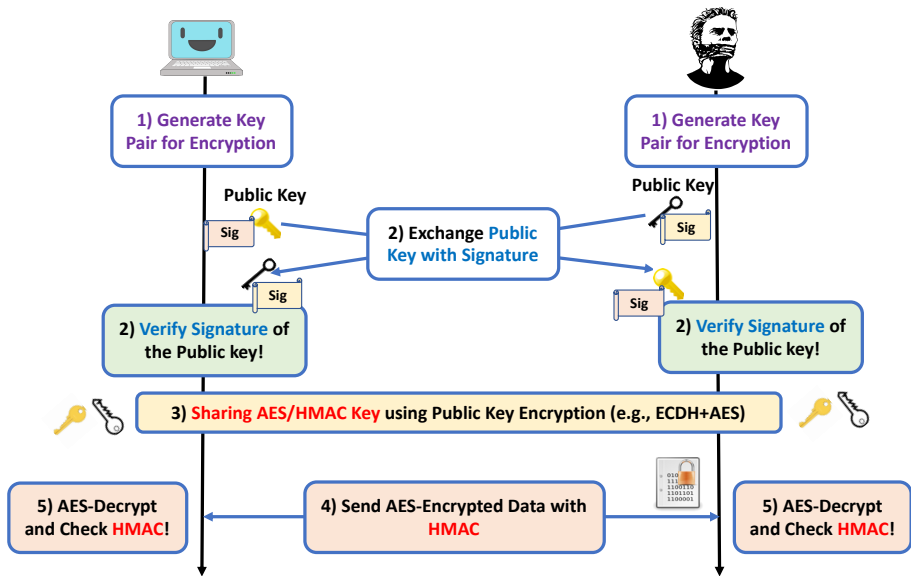
処理の重さで使い分けるのが鉄則。

- 暗号化されたデータのやり取りの手続きで、**署名はエンド間でのイニシエーション**に使う
- **MAC はエンド間でなんども繰り返すような本人確認・データ整合性確認**に使う

例えば:

- 1 署名を付与して、ECDH-ephemeral の公開鍵を交換。
- 2 ECDH-ephemeral + AES で HMAC/AES 暗号化の鍵を共有。
- 3 以降の大規模データのやり取りは AES 復号と同時に HMAC で本人確認を実施。

図にすると:



まとめ

まとめ

お疲れ様でした。

- データの真正性・本人確認のためのテクニックについて学んだ。
 - Hash: データの「指紋」。SHA-2 以降のものを使おう。
 - MAC:
 - 共通鍵ベースの本人確認手法。動作が軽い。
 - JS だととりあえず HMAC を使う。
 - 署名:
 - 公開鍵ベースの本人確認手法。MAC と比べると動作がかなり重い。
 - RSASSA: **RSASSA-PSS がおすすめ**。
 - ECDSA: 大体どの環境でも使える。
- 上記について、JavaScript のコードを実行/中身を覗いてみた。
- 署名はエンド間のやり取りイニシエーション、MAC はその後のデータのやり取りに使うと良い。

今後の予定

色々と未説明な事柄について、付録資料を作ろうと思います。

- RFC とは？NIST FIPS PUBって何。どこで何が決まってるの？
- データフォーマット (鍵とか署名とか) はどうなっているの？
PEM？DER？JWK？Web で使うにはどれがいい？
- ECDH/ECDSA の楕円曲線の選択ってどうすればいいの？

とか。

他に説明して欲しいこと等、リクエストを募集してます。

このセミナー自体はセカンドシーズンを行います (connpass 登録予定)。途中聞き逃した回があれば是非どうぞ。

宣伝: iTransfy by Zettant

簡単・安全にファイル転送ができる

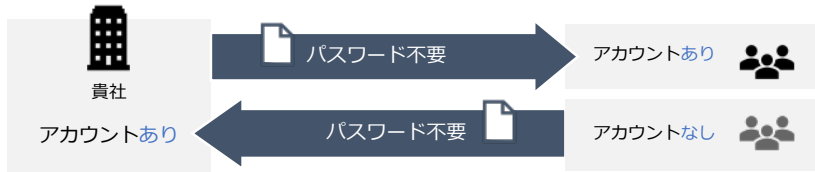


iTransfy for biz

<https://www.itransfy.com>

アカウント登録で、パスワード入力の手間が省けます

クライアント/協力会社等へファイルを送りたい、また送付してほしい時の手間を軽減



宣伝: 株式会社ゼタント



ゼタントはのミッションは、

「自分の身は自分で守ることができる世の中にする」

ことです。

共感してくれる仲間を募集しています！

問合せ先: recruit@zettant.com

会社 URL: <https://www.zettant.com>