小特集
IEICE Communications Society Magazine
Special Issue

編集チームリーダ　木下雅之
Masayuki Kinoshita

DX 時代に求められるセキュリティ課題と対策

近年，私たちの社会は急速なディジタル化とともに，サイバー空間におけるリスクと隣り合わせの時代に突入しつつあります．DX（ディジタルトランスフォーメーション）が進展し，IoT，AI，生体認証といった技術が社会基盤に深く組み込まれる中，それらを標的とするサイバー攻撃や情報改ざんの手法もまた高度化・多様化しています．セキュリティ対策は一部の専門領域にとどまるものではなく，あらゆる分野における不可欠な基盤となりつつあります．本小特集では，こうした現代的な脅威とそれに対抗するセキュリティ技術・制度について「セキュリティ法規」，「ぜい弱性」，「暗号」，「生体認証」，「ディープフェイク」の視点から多角的に掘り下げ，5 件の解説記事を専門家の皆様に執筆頂きました．

初めに，セキュリティ法規の観点から，自動車及びコネクティッド・ディジタル製品におけるサイバーセキュリティの法規要件と管理対策について紹介頂きます．次にサイバー攻撃の起点となるぜい弱性について取り上げ，その発生原因や攻撃例，そしてそれらへの対策について解説頂きます．また，私たちの身近で利用されているセキュリティ技術として，暗号化技術と生体認証について焦点を当てます．セキュリティ実現における中核的な技術の一つである暗号については，攻撃手法と安全性評価の重要性について解説頂きます．顔認証や指紋認証といった生体認証は利便性に優れ，PC やスマートフォンのロック解除や本人認証などに利用される一方，そのリスクと対策について御紹介頂きます．更に，AI の発展に伴い，新しい脅威として浮上しているディープフェイク音声に関する最新の研究成果や動向，課題について紹介頂きます．

小特集記事では，セキュリティの基礎的な知識と応用の両面に触れる解説記事をそろえており，初学者にも親しみやすい内容となっています．本小特集が，御覧の皆様のセキュリティ分野への関心を高めるきっかけになりましたら幸いです．

最後に，本小特集の発行にあたり，御執筆頂いた皆様，査読や校閲に御協力頂いた皆様，編集作業に御協力頂いた皆様へ感謝申し上げます．

小特集編集チーム
木下雅之，伊藤敬義，小南大智，中西孝行，松室尭之，村上靖宜

世界のセキュリティ法規と求められる管理・技術対策

竹森敬祐　Keisuke Takemori　DNV ビジネス・アシュアランス・ジャパン

1　はじめに

製品ユーザをサイバー攻撃から保護することを目的として，2021 年 1 月に国連欧州経済委員会（UN/ECE）から自動車のサイバーセキュリティ（CS）規則として UN Regulation No.155（UN-R155）[(1)] が公表され，2022 年頃から各国の法規として施行が始まった．2024 年 4 月に英国から消費者向けコネクティッド製品を対象とした CS 法規として Product Security and Telecommunications Infrastructure（PSTI）[(2)] が，2024 年 12 月に欧州連合（EU）の執行機関である欧州委員会からディジタル製品に対して CS 要件を課す Cyber Resilience Act（CRA）[(3)] がそれぞれ発効され，各国で施行へと進みつつある．

これらの法規は，自動車，コネクティッド・ディジタル製品の製造者が開発フェーズから CS に真摯に取り組み一定の CS レベルを確保すること，自動車やコネクティッド・ディジタル製品としてサイバー攻撃への耐性を高めることや，ぜい弱性やインシデントへの迅速な対応を求めている．違反すると，生産・販売の停止や大きな違反金が課される可能性がある．これまで自主努力で進めていた技術的 CS 対策が，各国において法規化へと移行する中で，自動車，コネクティッド・ディジタル製品の製造者にとって，組織の管理策として Cyber Security Management System（CSMS）の構築と技術的 CS 対策の理解は重要である．

本稿では，UN-R155，PSTI，CRA の三つの CS 法規の要件を概観し，自動車やコネクティッド・ディジタル製品の製造者が整備すべき管理策と，製品の CS を確保するために必要となる技術的 CS 対策の一例を紹介する．

以下，**2.** で三つの法規要件を概観する．**3.** で組織としてのルール・プロセスを整えて運用するための CSMS の管理策を整理し，**4.** で CS リスク分析手法や暗号技術の一例を紹介する．最後に**5.** でまとめる．

2　法規要件の概説

自動車の CS 規則である UN-R155 は，事業者として CSMS を整備して，サイバー攻撃からの損害を防ぐための管理策と技術的 CS 対策を考慮した車両の開発・生産・保守を推進する．車両の CS 耐性を確保することでサイバー攻撃からの損害を防ぐことができ，道路ユーザの安全やプライバシー保護などを達成する，学びの多い規則となっている．まず初めに，UN-R155 の要件を整理する．その後，英国で施行された消費者向けコネクティッド製品の CS 対応を求める PSTI と，欧州委員会から発効されたコネクティッド・ディジタル製品に対する CS 対応を求める CRA の要件を整理していく．

なお，誌面の都合上，全ての法規要件を解説できないことに御留意を頂きたい．詳細は，文献（1）～（3）に示した原文を参考にされたい．

2．1　UN-R155 の要件

UN-R155 の概形を表 1 に示す．UN-R155 は，国連欧州経済委員会（UN/ECE）の傘下にある自動車基準調

表 1　UN-R155 の概形

項目	内容
制定主体	国連欧州経済委員会（UN/ECE）WP.29
公開日	2021 年 3 月 4 日
施行日	2022 年頃から順次，1958 年協定加盟国で法規施行 [(4)]
目的	車両の開発・生産・保守フェーズの CSMS の構築と車両の CS 強化
対象	自動車，車載電子制御システム，これらと相互作用のある車両外コンポーネント
主な要件	Cyber Security Management System（CSMS），車両 CS 型式

表2　UN-R155 7.2 節　CSMS 要件の概要

項目	要件の概要
7.2.2.1	開発・生産・生産後フェーズへの CSMS 適用
7.2.2.2	(a)組織としての CSMS の構築，ルールの整備 (b)リスク特定のプロセス (c)リスク評価，対策導出のプロセス (e)テスト実施のプロセス (f)リスク評価を最新に保つプロセス (g)ぜい弱性や攻撃を監視し対応するプロセス (h)攻撃関連のデータを提供するプロセス
7.2.2.3	迅速にぜい弱性へ対応するプロセス
7.2.2.4	(a)ぜい弱性／サイバー攻撃を監視するプロセス (b)ぜい弱性／サイバー攻撃を分析するプロセス
7.2.2.5	CS に関する責任分担，サプライチェーン管理

表3　UN-R155 7.3 節　CS 技術要件の概要

項目	要件の概要
7.3.1	CSMS 適合証明書の取得
7.3.2	サプライヤ関連のリスクの特定と管理
7.3.3	リスク評価の実施
7.3.4	対策の実装
7.3.5	第三者製品から CS 影響を受けない車両対策
7.3.6	テストの実施
7.3.7	(a)攻撃の検知と防止 (b)監視能力のサポート (c)データフォレンジックの確保
7.3.8	コンセンサス規格に準拠した暗号技術の採用

和フォーラム（WP.29）内の自動運転／CS 分科会で議論され，2021 年 3 月に公開された．自動車の型式認定を国際的に調和させるための「1958 年協定」[4] に加盟する国において，2022 年頃から順次，自動車 CS 法規として施行が始まっている．車両並びに車載電子制御システム，及びそれらと相互作用のある車両外のコンポーネントを対象とした法規である．自動車製造者に課される法規であるが，要件の中に部品サプライヤの管理や技術文書の説明責任を課しており，部品サプライヤへも間接的に適用される．CS 要件としては，組織とルールを整えて，サイバー攻撃に対して車両の耐性を確保する CS 開発と生産を推進すること，及び生産後のぜい弱性・インシデントの監視と対応を推進する CSMS の構築と運用を求めている．また，CS リスク分析に基づく技術的 CS 対策を織り込んだ車両開発と生産を求めている．CSMS に関する要件の概要を表 2 に，CS 技術要件の概要を表 3 に示す．

UN-R155 7.2 節は，車両のライフサイクルとしての開発・生産・生産後フェーズに CSMS の管理策を適用することを求めている．CSMS とは，組織として CS ルールやプロセスを整え，役割と責任を割り当て，運用と監査による改善を推進する管理策である．開発フェーズに対して，リスクの特定・評価，対策の導出，実装した技術的 CS 対策のテストに関するプロセスの整備を求めている．また，生産後フェーズに対して，発覚するぜい弱性や車両へのサイバー攻撃を監視して対応を進めるプロセスや，監視の結果を開発フェーズにフィードバックすることでリスク評価を最新に保つプロセスの整備を求めている．ぜい弱性やサイバー攻撃の監視においては，原因や攻撃経路を分析する能力が求められ，迅速に対応するためのプロセスを整備する必要がある．UN-R155 は，車両製造者に課される規則であるが，その中には部品サプライヤとの CS に関する責任分担を定めるサプライチェーン管理の要件があり，サプライチェーンにも間接的に適用される．車両製造者は，CSMS を適切に構築，運用できていることを，各国が定める認定機関，日本の場合は自動車技術総合機構に申請し，CSMS 試験を受審する[5]．合格すると，CSMS 適合証明書を取得できる．

UN-R155 7.3 節は，CS に関する車両等の型式認定[4] を取得するために，取得した CSMS 適合証明書に基づく CS 開発やサイバー攻撃監視を行うことを求めている．CS 開発においては，サプライヤ開発において特定されたリスクを把握し管理する必要がある．車両製造者としても，CS リスク評価に基づく設計を行い，対策の実装とテストが求められている．また，車両製造者の管理が及ばない第三者製品のぜい弱性に起因して，車両が CS 影響を受けない仕組みの実装も必要とされている．生産後フェーズである市場へ提供された車両に対するサイバー攻撃を監視・検知し，必要に応じて対策を図る必要がある[6], [7]．採用する暗号技術は，各国の暗号評価プロジェクトなどが認めたコンセンサス規格に準拠した堅ろう性の高いものを選択しなければならない．こ

表4　PSTI の概形

項目	内容
制定主体	英国
施行日	2024 年 4 月 29 日
目的	コネクティッド製品の CS 強化
対象	消費者向けコネクティッド製品
主な要件	・コネクティッド製品への CS 要件の導入 ・ぜい弱性情報を受け付ける窓口の設置 ・サポート並びにサポート期間の公表

表5　CRA の概形

項目	内容
制定主体	欧州委員会
発効日	2024 年 12 月 11 日
施行日	2026 年 9 月 11 日 ぜい弱性の報告
目的	2027 年 12 月 11 日 CRA の全体
対象	ディジタル製品の CS 強化
主な要件	・CS リスク分析に基づく CS 設計・開発 ・ぜい弱性やインシデントの当局への報告 ・サポート並びにサポート期間の公表

れにより，サイバー攻撃に耐性のある車両が開発・生産され，新たなサイバー攻撃にも迅速に対処できることが期待される．

2．2　PSTI の要件

PSTI の概形を表 4 に示す．PSTI は，英国において 2024 年 4 月から施行された，インターネットに接続可能な消費者向け製品に対する技術的 CS 対策を要求する法規である．技術的 CS 対策として，製品ごとに異なるパスワードを設定するか，初回起動時にユーザに変更させることなどを求めている．また，製品のぜい弱性の発見者から情報を受け取る窓口の設置や，セキュリティ更新プログラムの提供やそのサポート期間を公表することなどを製造者に求めている．UN-R155 のような CS リスク分析に基づく製品開発を明示的には求めておらず，表 4 に示す主な要件から構成されている．

2．3　CRA の要件

CRA の概形を表 5 に示す．CRA は，欧州委員会から 2024 年 12 月に発効された．欧州域の市場へこれまで提供されてきた（以降，上市と呼ぶ．）コネクティッド・ディジタル製品に発覚するぜい弱性について，2026 年 9 月から報告義務が課される．その後，2027 年 12 月から CRA の全ての要件が施行される．対象は，ディジタル要素を含む OS，ブラウザ，ネットワーク管理システムなどの製品や，これらを統合した製品である．なお，様々な製品に水平的に適用される OS，ブラウザ，ネットワーク管理システムなどは，重要なディジタル製品に分類され，OS 内の仮想実行環境やファイアウォールなどの CS 製品については，重要なディジタル製品クラス II に分類され，クラスごとに決められた適合性評価を受審する必要がある．CRA の主な要件として，CS リスク分析に基づく CS 設計・開発，ぜい弱性やインシデントについて当局への報告，セキュリティ更新プログラムの提供やサポート期間の公表などがある．要件の詳細を，表 6，7 にまとめる．

表6　CRA 付属書 I　Part 1 製品 CS 要件

項目	内容
(1)	リスクを踏まえた CS 設計・開発・生産
(2)	(a) 既知の悪用可能なぜい弱性がない状態での上市 (b) CS が確保された状態での上市 (c) ぜい弱性対応としてのソフトウェア（SW）更新の保証 (d) 不正アクセスからの保護 (e) データに関する機密性の担保 (f) データに関する完全性の担保 (g) 取り扱うデータの最小化 (h) DoS 攻撃からの回復力，可用性の確保 (i) ほかの製品やネットワークへの悪影響の最小化 (j) 攻撃対象領域を制限する設計・開発・生産 (k) 損害の影響を低減する設計・開発・生産 (l) ログの記録と監視，オプトアウト機会の提供 (m) データや設定の削除機能の提供

表7　CRA 付属書 I　Part 2 ぜい弱性対応の要件

項目	内容
(1)	ソフトウェア部品表（SBOM）の整備，ぜい弱性及びソフトウェアコンポーネントの特定
(2)	迅速なセキュリティ更新プログラムの提供
(3)	定期的な CS テストとレビューの実施
(4)	修正されるぜい弱性に関する情報の開示
(5)	ぜい弱性情報の開示に関する方針の策定と実施
(6)	ぜい弱性の報告を受け付ける窓口の公表
(7)	CS を確保した更新と自動的な SW 更新の仕組み
(8)	更新プログラムの迅速かつ無料での提供

CRA では付属書 I　Part1 に，製品に対する CS 要件が列挙されている．CS リスク分析に基づく CS 設計・開発・生産を行うこととされており，分析の結果から項目（2）の（a）～（m）の要件に対応する必要がある．例えば，(e) 個人やそのほかのデータを保管，送信，処理において漏えいさせない機密性の担保，(f) 個人やそのほかのデータ，コマンド，プログラムを勝手な操作や変更から保護する完全性の担保，(l) データや処理をログとして記録して活動を監視する機能と，監視や他者へのログ提供をユーザが拒否するオプトアウトの機会の提供などを求めている．

CRA の付属書 I　Part2 には，ぜい弱性対応の要件として表 7 の項目（1）から（8）が列挙されている．例えば，(1) ではソフトウェア部品表（SBOM: Software Bill of Materials）を整備して，ぜい弱性が含まれる製品やコンポーネントの特定や依存関係を把握できるようにすること，項目（5）では「Vulnerability Disclosure

表８　積極的に悪用されているぜい弱性の報告

期限	提出物	内容
24 時間以内	早期警告通知	・製品が利用されている国
72 時間以内	ぜい弱性通知	・製品に関する一般的な情報 ・エクスプロイト及びぜい弱性の一般的な性質 ・講じられた対策やユーザが講じることのできる対策に関する一般的な情報
対策が利用可能になった後 14 日以内	最終報告書	・深刻度を含むぜい弱性の説明 ・可能な場合，悪用者に関する情報 ・セキュリティ更新プログラムまたはほかの是正措置の詳細

表９　CS に影響を及ぼす重大なインシデントの報告

期限	提出物	内容
24 時間以内	早期警告通知	・製品が利用されている国
72 時間以内	インシデント通知	・インシデントの一般的な情報，性質，初期評価の結果 ・講じられた対策やユーザが講じることのできる対策に関する一般的な情報
インシデント通知の提出後１カ月以内	最終報告書	・深刻度を含むインシデントの説明 ・脅威または根本原因 ・適用された／継続中の対策

Policy（VDP）」と呼ばれるぜい弱性情報の開示に関する方針を策定して運用することなどを求めている．

ここで，CRA の第 14 条には製造者の報告義務として，製造者は自ら認識した積極的に悪用されているぜい弱性，若しくは CS に影響を及ぼす重大なインシデントについて，迅速に当局へ報告しなければならないとされている．報告期限や内容について，表 8，9 にまとめる．

報告は，早期警告通知，ぜい弱性／インシデント通知，最終報告の３回となっている．早期警告通知は，自らぜい弱性を認識してから 24 時間以内とされており，製品を上市した国を事前に整理しておく必要がある．また，ぜい弱性／インシデント通知は，ぜい弱性／インシデントに関する性質やユーザが講じることのできる対策に関する情報を含む必要があり，72 時間以内に報告するには，ぜい弱性／インシデントの分析能力を適切に持つ必要がある．最終報告書には，根本原因や対策に関する情報が求められ，ここでも適切な分析能力を持つ必要がある．なお，報告に関する第 14 条は，2026 年 9 月 11 日から施行され，これまで欧州域へ上市して利用されている全てのディジタル製品が対象となる．

3　CSMS の構築の一例

UN-R155 では，組織として CS 対応を適切に推進するための CSMS の構築を求めている．これは，組織として体制とルールを整備し，誰が実施しても同様の結果を得られる業務フローを作成・運用し，監査によって継続的な改善を推進することである．

3．1　体制とルールの整備

図 1 に CSMS 構築のための体制とルール整備の一例を示す．人員，IT システム，文書などの情報セキュリティ管理は，Information Security Management System（ISMS）で推進し，これを組織運営の土台とする．この上に，製品の CS を実現するための CSMS を乗せるとよい．

製品・情報セキュリティに関する最高責任者は，製品・情報セキュリティ委員会を運営し，目的や方向性を示す最上位文書として基本方針を掲げ，これを実現するために何をどの程度達成するかを示す対策基準を策定

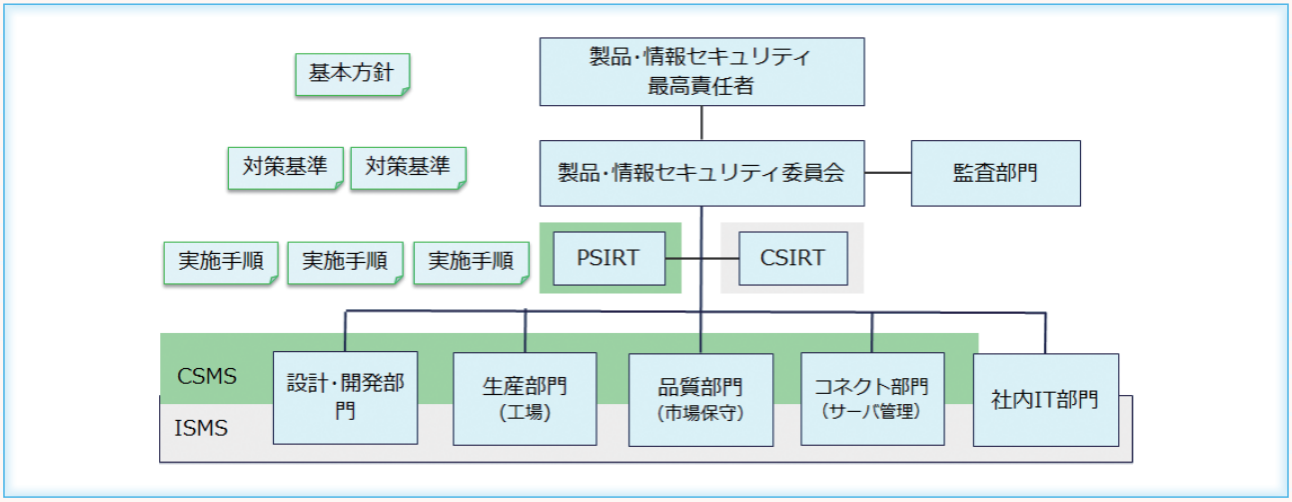


図 1　体制とルールの整備の一例

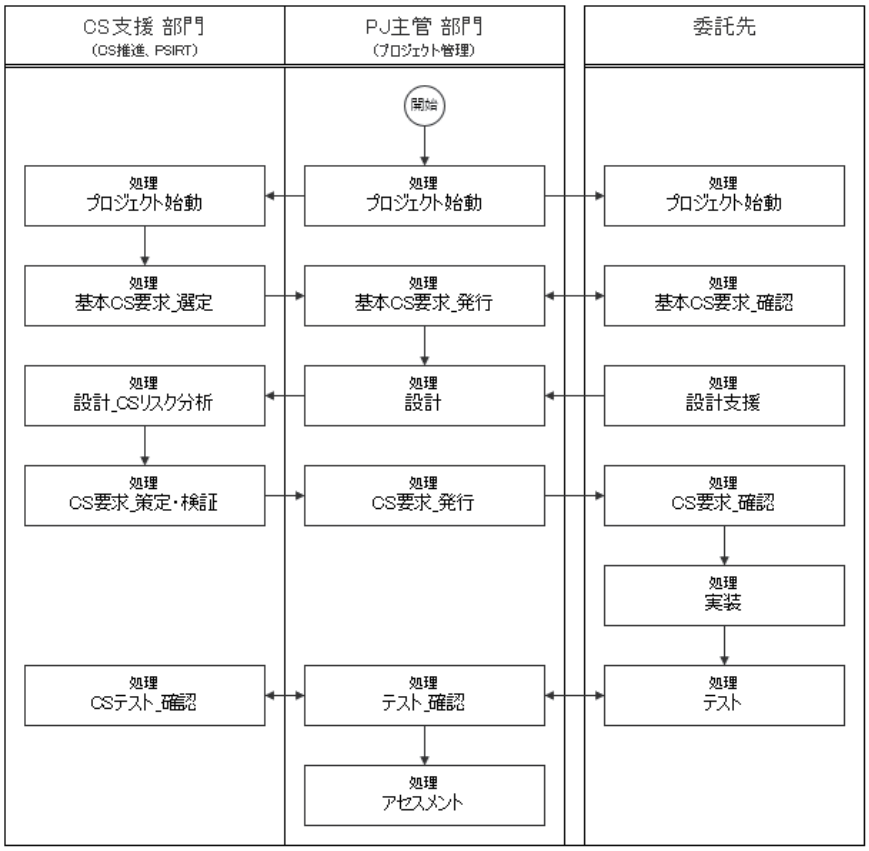


図２　業務フローの一例

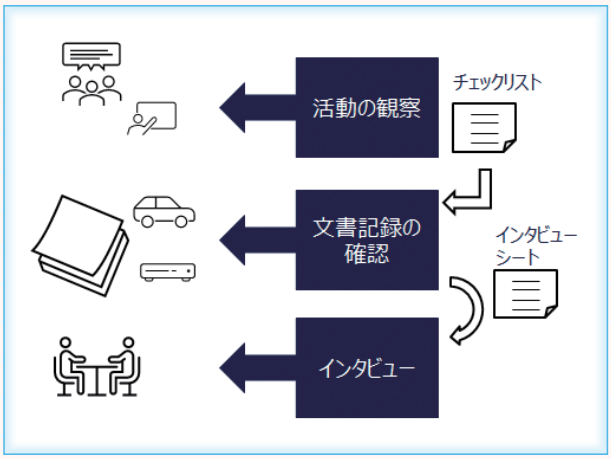


図３　監査活動の一例

し，周知する．CSMS の対象となる部門は，基本方針と対策基準を受けて，現場での対応を具体的な手順としてまとめた実施手順を策定し，その教育を推進する．また，製品に発覚するぜい弱性やインシデントへの対応として Product Security Incident Response Team（PSIRT）を設置して，組織間連携による迅速な対応を推進する．

3．2　業務フローの整備

図２に，実施手順の一つである製品の CS 開発を進めるための業務フローの一例を示す．製品 CS 開発のプロジェクトを主管する部門と，そこを支援する部門，委託先など連携先を明確にする．これにより，関係者の役割と業務の流れが分かるようになる．また，各業務の実施手順には，入力 →処理 →出力を定義するとよい．

3．3　監査による改善

図３に，CSMS に関わるルールの確認や運用を確認する監査活動の一例を示す．CSMS の機能の一つとしてルールが策定され適切に運用されていることを確認する監査部門を設けることもある．監査は，各部門の活動や，活動が記録された文書を確認する．また，各部門の担当者へのインタビューを通じて，CSMS が適切に運用されていることを確認する．結果は，適合，改善の機会，軽微な不適合，重大な不適合に分類して，改善活動を推進する．

4　技術的 CS 対策の一例

ここでは，製品 CS 開発において求められる CS リスク分析の流れ，CS 開発活動が適切であることを確認する検証・テスト・アセスメント，そして技術的 CS 対策の基本となる暗号技術の一例を紹介する．

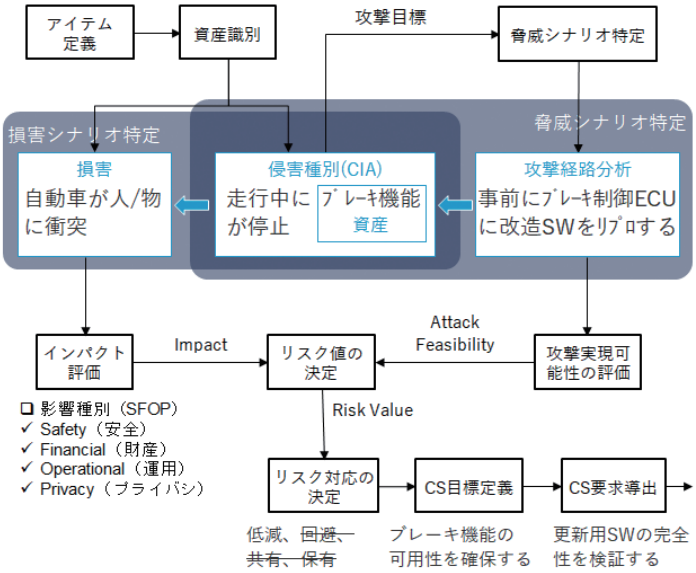


図４　CS リスク分析の流れ

４．１　CS リスク分析の手法

図４に，CS リスク分析の流れを示す．分析は，資産識別 →損害シナリオ特定 →インパクト評価 →脅威シナリオ特定 →攻撃経路分析 →攻撃実現可能性評価 →リスク値の決定 →リスク対応の決定へと進む．リスク値を参考に対応として，低減：技術的 CS 対策などを図る，回避：リスクを含む機能を取り除くなど，共有：開発プロジェクトを保険に加入させるなど他者と共有する，保有：リスクが小さい場合，今は対応しないと判断するなどがある．

対応として低減を選択した場合は，CS 目標定義 → CS 要求導出へと進む．図４では，ブレーキ機能の可用性を確保することを CS 目標として，これを達成するための CS 要求として，更新用 SW の完全性を検証することを導出している．その後も，システム設計を考慮した詳細な CS リスク分析を繰り返す段階的詳細化を進めることで，製品の全体的な技術的 CS 対策から，ハードウェア／ソフトウェアの細部まで技術的 CS 対策を適切に施すことができる．

脅威シナリオ特定には幾つかの手法があるが，Microsoft が提唱する脅威のモデル化 “STRIDE”[8] がよく知られている．資産識別，損害シナリオ特定，攻撃経路分析なども，様々な手法があるため，製品の特性に合わせた適切な手法を実施手順として整備しておくとよい．インパクト評価と攻撃実現可能性評価からリスク値を決定する指標には，CVSS v3[9] がある．こうした指標に基づき，リスク対応として，何らかの対策を図る “低減”，リスクを含む機能を排除するなどの “回避”，CS 開発プロジェクトを保険に加入させるなど他者との “共有”，リスク値が極めて低いと判断して “保有” するなど，決定していく．低減と決定されると，攻撃実現可能性を低減させる対策としての CS 要求を導出する．

４．２　検証・テスト・アセスメント

開発プロジェクトの計画や，CS リスク分析の流れが適切に行われたことをチェックリストなどを用いて検証することで，設計レベルでのぜい弱性の混入を防ぐことができる．検証で用いるチェックリストの例を図５に

【チェックリスト】

No.	対象	チェック項目	判定結果	判定理由
1	具体的な攻撃方法の列挙	脅威シナリオを実現する具体的な攻撃方法が攻撃経路として記されている。		
2		脅威シナリオごとに、考えられる攻撃経路が列挙されている。（一つの脅威シナリオに対して、複数の攻撃経路があり得る。）		
3		脅威シナリオと攻撃経路に矛盾はない。		

図５　検証で用いるチェックリストの一例

示しておく．製品の特性に合わせて，チェック項目を導出しておくとよい．設計後は，実装へと進み，CS 要求に基づく機能が仕様どおりに動作していることをテストを通じて確認する．最後に，一連の活動記録を第三者的な視点でレビューすることで，妥当性を判断するアセスメントを行うとよい．

4．3　暗号技術

情報セキュリティにおける機密性，完全性，可用性を確保するためには暗号技術の適用が一般的である．例えば，情報の機密性を担保するための暗号化や，完全性を確認するための署名・検証などがあり，これらを組み合わせてシステムの可用性を確保する．暗号技術には，十分な強度を有するアルゴリズムや「鍵長」を適用する必要がある．これら暗号アルゴリズムや鍵長については，CRYPTREC[6] や NIST SP-800[7] などで評価されており，参考にするとよい．

5　まとめ

本稿では，世界的に法規化が進む自動車やコネクティッド・ディジタル製品を対象とした CS 法規の要件を整理し，これを満たすための管理策や技術的 CS 対策の一例を概観した．自主的な CS 開発・生産・保守を進めていた時代は終わり，自動車やコネクティッド・ディジタル製品の製造者として CS 要件への対応が必須となってきている．CS 法規に違反すると製品の生産や販売の停止や，大きな制裁金を課されることになるため，各種 CS 法規要件の把握と，CSMS の構築，技術的 CS 対策の理解に努める必要がある．

■ 文献

(1) 国連欧州経済委員会自動車基準調和世界フォーラム（WP.29），“UN-R155.” https://unece.org/transport/documents/2021/03/standards/un-regulation-no-155-cyber-security-and-cyber-security
(2) 欧州委員会，“Cyber Resilience Act（CRA）.” https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R2847
(3) 英国，“Product Security and Telecommunications Infrastructure Act（PSTI）.” https://www.legislation.gov.uk/ukpga/2022/46/part/1/enacted
(4) 国土交通省，“国連の車両等の型式認定相互承認協定（1958 年協定）の概要.” https://www.mlit.go.jp/common/000055228.pdf
(5) 自動車技術総合機構，“CSMS 試験.” https://www.naltec.go.jp/publication/regulation/gtg5d20000000fht-att/ih3l1n00000008r8.pdf
(6) CRYPTREC，“電子政府推奨暗号の評価・監視.” https://www.cryptrec.go.jp/
(7) NIST Computer security resource center, “SP 800 guidance.” https://csrc.nist.gov/publications/sp800
(8) Microsoft，“脅威のモデル化 STRIDE.” https://learn.microsoft.com/ja- jp/azure/security/develop/threat-modeling-tool-threats
(9) 情報処理推進機構，“共通脆弱性評価システム CVSS v3概説.” https://www.ipa.go.jp/security/vuln/scap/cvssv3.html

竹森敬祐

DNV ビジネス・アシュアランス・ジャパン株式会社．これまでインターネットセキュリティ，スマートフォンのセキュリティ＆プライバシー，自動車セキュリティ，ディジタル製品セキュリティに関する研究・開発や制度設計などに関わる．

生体認証をめぐる脅威と対策技術の研究動向

高田直幸　Naoyuki Takada　セコム IS 研究所

1 はじめに

生体認証技術は「バイオメトリクス」とも呼ばれ，今日の私たちの生活に不可欠なものとなっている．PC やスマートフォンのロック解除やサインイン，公的サービスにおける本人確認，施設へのアクセスコントロールなど，様々な生活シーンで，顔認証や指紋認証，静脈認証といった生体認証技術が活用されている．利用者の「生体情報」を利用する生体認証技術は，従来のパスワードや暗証番号などの「記憶」に基づく認証や，IC カードなどの物理的なデバイスの「所持」に基づく認証に比べて，忘却や紛失のリスクが低く，利便性にも優れるとされ，広範な分野で実用化が進んでいる．

一方で，生体情報には変更が困難であるという本質的な性質があり，その漏えいや偽造，不正利用が深刻なセキュリティリスクを招く．このため，偽造生体情報への対策や，システム内に保持される生体情報の保護は，生体認証システムを安全かつ高い信頼性で運用するための最重要課題の一つとなる．更に近年，個人情報保護法や欧州一般データ保護規則（GDPR），欧州 AI 法などの法整備が進み，生体情報及び生体認証技術の利活用においては，倫理的な妥当性やプライバシーへの配慮も不可欠となっている．

本稿では，生体認証技術／システムの基礎を説明した上で，生体認証をめぐる脅威と対策技術の研究を紹介する．特に，偽造生体情報による攻撃と「プレゼンテーションアタック検知」と呼ばれる対策技術，セキュリティのみならずプライバシーへの配慮において重要な技術となるテンプレート保護技術に焦点を当てる．

2 生体認証技術の基礎

生体認証は，人体の生物学的特徴に基づき，個人を識別・認証する技術である．この生体認証に用いられる生物学的特徴を「モダリティ」と呼ぶ．

2．1　生体認証のモダリティ

表 1 に代表的なモダリティを示す．生体認証のモダリディは，身体的特徴と行動的特徴に大別される．身体的特徴は人体の構造的な特徴に基づく「静的な」モダリティであるのに対し，行動的特徴は，個人特有の行動パターンや振る舞いによる「動的な」モダリティである．行動的特徴には環境や心理状態，身体の状態による影響が表れやすく，一般には身体的特徴の方が安定的で，高い識別精度を示す．その一方で，継続的に変化する行動的特徴の方が，生体情報の偽造や「なりすまし」への耐性が高いとされる．

複数のモダリティを組み合わせる認証方式も存在し，「マルチモーダル認証」と呼ばれる．マルチモーダル認証には，異なるモダリティの生体認証を順次実行する方式と，指紋と指静脈，掌紋と手のひら静脈といった，同時に取得可能な複数のモダリティを組み合わせる方式がある．このほか，高い精度を示す身体的特徴と，高い「なりすまし耐性」を示す行動的特徴を補完的に組み合わせるという考え方もある．

2．2　生体認証システムの構成

一般的な生体認証システムの構成を図 1 に示す．生体認証システムの多くは，「生体センサ」「特徴抽出部」「テンプレートデータベース」「照合部」から構成される．システムの動作は，ある人物の生体情報をシステムに「登録」するフェーズと，ある人物が生体情報を登録済みの人物であることを「認証」するフェーズに分けられる．

生体センサは，人物から生体情報を取得する役割を果たし，モダリティに合わせたセンサが利用される．例えば，指紋認証であれば，光学式，静電容量式，超音波式等の各種の指紋センサ，顔認証であれば，可視光カメラ

表 1　代表的な生体認証モダリティ

分類	モダリティ	解説	精度＊
身体的特徴	指紋認証	手の指先の皮膚の凹凸パターン（隆線，汗腺など）を解析．手荒れや乾燥に影響を受けやすい	高
	顔認証	顔の画像特徴や目、鼻等の器官特徴を解析．照明や表情，年齢変化に影響を受けやすい	中〜高
	虹彩認証	眼球の虹彩パターンを解析	高
	静脈認証（手のひら・手指）	手のひらや手指の血管パターンを近赤外線で読み取って解析．体表に露出せず偽造耐性がある	高
	耳介形状認証	耳の形状を解析．実用化事例は少ない	中
	DNA 認証	遺伝子情報を基に識別．精度は極めて高いが一卵性双生児は区別できない．即時認証には不向き	高
行動的特徴	筆跡認証（署名認証）	書字動作の軌跡や筆圧，速度などを解析	中
	キーストローク認証	キーボード入力のタイミング・強さなどを解析	低〜中
	歩容認証（歩き方）	歩行時の動きや体重移動パターンを解析．遠隔からの認証用途に親和性が高い	中
	音声認証（行動的側面）	声のスペクトルや発話内容の癖（イントネーション，リズム）を解析	中
	タッチ動作認証	スマートフォンなどの操作パターン（スワイプ，タップ等）を解析	低〜中

＊特定の製品／技術の精度ではなく一般的な傾向として記載．

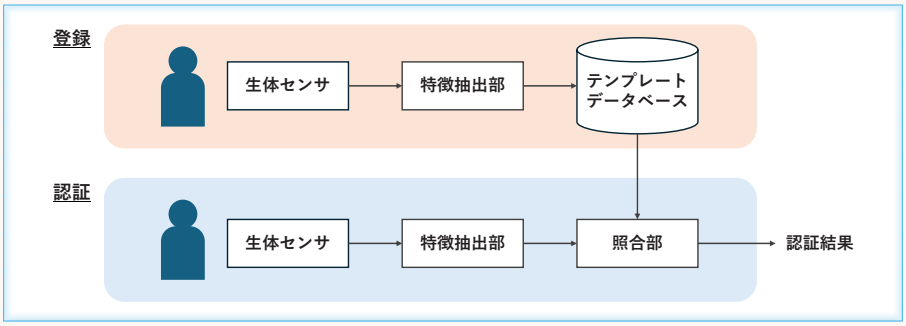


図 1　生体認証システムの構成

や，スマートフォンや PC にも搭載されている Depth カメラなど，多様なセンサが存在する．従来，生体情報の取得には，センサと生体を接触または近接させる必要があったが，近年のセンシング技術や，取得した生体情報からの特徴抽出処理の進化により，非接触や遠隔から取得できるモダリティが増えている．

特徴抽出部では，生体センサで取得した生体情報から個人の識別に用いる特徴が抽出され，システムが扱いやすい「特徴量」に変換される．登録フェーズで抽出された特徴量は，事後の認証フェーズで参照できるようデータベースに格納され，「テンプレート」と呼ばれる．特徴量の形式は，一般に，モダリティや認証アルゴリズムによって異なるが，指紋認証におけるマニューシャ特徴や，John G. Daugman 教授が特許化し，虹彩認証で利用されるアイリスコード等，広く認知されている形式も存在する．このほか，深層学習をはじめとする機械学習を活用して特徴量を算出する方式も，数多く研究されている．

照合のプロセスでは，利用者の生体情報から算出した特徴量と，データベースに格納された登録人物のテンプレートとの類似度が算出される．類似度が所定の照合しきい値を超えた場合，認証対象者は登録人物と同一人物であると判定される．特徴量間の類似度の算出には，認証アルゴリズム独自の手法やコサイン類似度などが用いられる．

3　生体認証システムに対する脅威と攻撃手法

生体認証技術は，認証に個人特有の生体情報を用いるため，安全で便利であると考えられがちであるが，セキュリティの観点からは，生体認証特有の二つの大きな「ぜい弱性」がある．第 1 に，生体情報の一部は体表に露見しており，容易に詐取される．顔画像は誰でもカメラで撮影することができる．人が素手で物に触れれば，

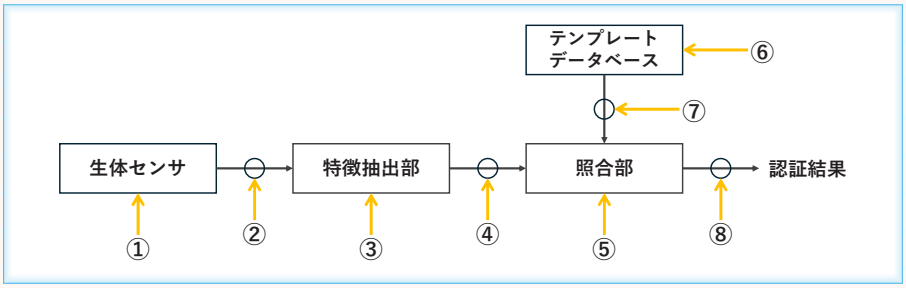


図２　生体認証システムの構成と攻撃のポイント

そこに指紋が残る．更に，センシング技術の発展により，非接触で取得できる生体情報も増えている．第２に，パスワードやICカード等の所持や記憶による認証手段とは異なり，生体情報は本質的に変更が困難である．例えば，個人の指紋画像が漏えいしたとしても，指紋の変更は容易ではない．これらのぜい弱性は，プレゼンテーションアタック（Presentation Attack）と呼ばれる偽造生体情報による攻撃につながる．

プレゼンテーションアタック以外にも，生体認証システムのぜい弱性を突いた攻撃が存在する．Rathaらは，生体認証システムにおける八つの攻撃ポイントを示した[(1)]（図 2）．これらのポイントへの攻撃は，次の四つに分類される[(2)]．

(A) ユーザインタフェースへの攻撃

(B) 特徴抽出部，照合部などのモジュールへの攻撃

(C) モジュール間の通信路への攻撃

(D) テンプレートデータベースへの攻撃

代表的な攻撃と攻撃のポイントを上記の分類で整理したものを表２に示す[(3)]．本稿では，生体認証システム特有の攻撃として，ユーザインタフェースへの攻撃であるプレゼンテーションアタックと，テンプレートデータベースへの攻撃を取り上げ，対策を交えて詳説する．モジュールや通信路への攻撃は，一般的なITシステムと同様に，セキュアなソフトウェア実行環境（TEE: Trusted Execution Environment）やコードサイニングの導入，通信路の暗号化やタイムスタンプ，チャレンジレスポンス方式の導入といった対策が有効に作用するため，本稿での対策の詳説は割愛する．

表２　生体認証システムへの代表的な攻撃（文献 (3) の表を一部改変）

番号	攻撃の種別	攻撃ポイント
1	プレゼンテーションアタック：偽造生体情報の提示（スプーフィング）	①
2	一卵性双生児や血縁者等の類似性の利用（顔など）	①
3	正規ユーザから攻撃者への生体情報サンプルの提供（正規ユーザと攻撃者との共謀）	①
4	ゼロ・エフォート試行：攻撃者が正規ユーザになりすまして自身の生体情報を提示，登録詐称	①
5	センサ上に残された残留生体情報の不正取得	①
6	センサや操作パネルの不正操作，物理的な破壊	①
7	通信路を流れる正規ユーザの生体情報や特徴情報の傍受，傍受した情報によるリプレイ攻撃	②，④
8	生体情報／特徴量／テンプレートの継続的な改ざんによるヒル・クライミングアタック	②，④，⑦
9	マルウェアによる攻撃，トロイの木馬の注入	③，⑤
10	多くのテンプレートと一致する偽造生体情報の推定，提示（ウルフ攻撃）	③，⑤（①，②）
11	テンプレート情報の不正入手	⑥
12	テンプレートの改ざん，攻撃者のテンプレートへの置換え	⑥
13	通信路からのテンプレート傍受，傍受したテンプレートによるリプレイ攻撃	⑦
14	攻撃者のテンプレート注入	⑦
15	通信される認証結果の改ざん：正規ユーザのアクセス拒否，攻撃者のアクセス許可	⑧
16	サービス拒否攻撃：通信路の切断／情報の改ざん／継続的なサンプル注入	②，④，⑦，⑧

表3　主なプレゼンテーションアタック手法とその対策技術（PAD）[5]~[10]

モダリティ	攻撃手法
指紋	・ゼラチン，シリコン，ラテックス，木工用ボンド，粘土等による人工指紋の作成 ・3D プリンタでの指紋型の作成 ・他人の切断指や死者の指の利用
顔	・写真や動画像の提示 ・顔写真を部分的にくり抜いて，お面として被って提示 ・シリコンや樹脂等による 3D マスクの作成，提示 ・ディープフェイク技術による顔画像や動画像の作成，提示 ・一卵性双生児等によるなりすまし
虹彩	・虹彩画像の提示（赤外撮影に対応したプリンタでの印刷，モニタへの表示） ・虹彩画像の上にコンタクトレンズを載せて提示（眼球の湾曲再現） ・虹彩模様を印刷したカラーコンタクトレンズを装着して提示
静脈	・蝋（ろう）やシリコン樹脂，大根などの野菜による人工物の作成 ・赤外インクで静脈パターンを紙に印刷して提示
音声	・声真似：声色や話し方の真似 ・録音音声の再生 ・音声変換：正規ユーザの音声特徴を学習し，攻撃者の声質を変換して再生 ・音声合成：正規ユーザの音声特徴を学習し、任意テキストを合成して読み上げ

3.1　プレゼンテーションアタック

プレゼンテーションアタックは「なりすまし攻撃」とも呼ばれ，生体センサに対し偽造した人工物や映像等の情報を提示して，生体認証システムの誤動作を誘発する攻撃である．古くは，顔認証システムへの顔写真の提示など，単純なアプローチが主流であったが，松本らが，食品用のゼラチンを使用して偽造指紋を作成して攻撃を成功させて以降，プレゼンテーションアタックへの注目が高まった[4]．

2016 年には，ISO/IEC30107 シリーズとして，プレゼンテーションアタック検知（PAD: Presentation Attack Detection）の基本フレームワークや評価方法が国際標準化されている．近年では，生体認証技術やデバイスの PAD 性能の評価を，第三者となる品質保証機関に依頼し，国際標準で規定されたテスト手法・報告方法に従った評価として公表するケースもある．

表 3 に，指紋，顔，虹彩，静脈，音声のモダリティ別に，主なプレゼンテーションアタック手法とその対策技術（PAD）を示す[5] ~ [10]．ここに示した攻撃手法は，シリコン樹脂や印刷物などを使用して物理的な偽造生体を作成し，センサに提示するアプローチと，写真や映像などの偽造情報を生成して提示するアプローチに大別される．前者は 3D プリンティング等の成形技術の進化によって，後者はディープフェイク[7] をはじめとする生成 AI 研究の進化によって高度化しており，偽造生体情報のクオリティは日増しに高まっている．

対策技術となる PAD 手法については，主にハードウェアによるものと，ソフトウェアによるものに大別される（表 4）．ハードウェアによる手法の多くは，生体の物理的な形質を直接計測することで生体検知を実現する．生成 AI による偽造情報生成攻撃に対しても効果的に作用すると期待できるが，専用のハードウェアを必要とし，適用できる場面が限定される．このため汎用的な技術として，ソフトウェアベースの PAD 研究も盛んとなっている．

3.2　テンプレートデータベースへの攻撃

本攻撃は，大きくは攻撃者テンプレートの不正注入と，テンプレートの不正取得に分けられる．前者は攻撃者の生体情報による不正アクセスを企図したものである．後者の攻撃でテンプレートが不正取得されると，次の脅威につながり，より大きな影響が想定される．

・リプレイ攻撃での利用
・テンプレートの基になった生体情報の復元
・本来用途以外への不正転用

リプレイ攻撃とは，傍受した信号を再度通信路に流す攻撃である．通信路にぜい弱性がある場合，不正取得されたテンプレートによる攻撃は，深刻な脅威となる．

テンプレートから元の生体情報を復元する試みは，テンプレート復元攻撃（template inversion attack）と呼ばれ，顔や指紋などのモダリティで多くの研究者が取り組んでいる．近年では，GAN などの生成 AI を活用した研究が進み，クオリティの高い復元が実現されている[11]．テンプレートからの顔画像の復元は，生体認証システムの利用者を特定されるプライバシーリスクにもつながる脅威となる．

本来用途以外への不正転用というのは，異なるシステム間のクロスマッチングを意味する．例えば，企業のア

表４　プレゼンテーションアタックの主な対策技術（PAD）[5]～[10]

モダリティ	ハードウェアによる対策	ソフトウェアによる対策
指紋	・静電容量式センサの利用（指表面の電位差の検知） ・型取り／印刷の解像度を超える高解像度での指紋センシング，微細特徴の利用	・発汗により指紋のコントラストが変化する現象の検知
顔	・Depth センサによる顔の 3D 形状の取得 ・マルチスペクトル／赤外線反射特性の観測による印刷物判別 ・カメラのピントを意図的にずらした画像からの立体判別 ・ランダムに表情や動作を指定するチャレンジ応答方式 ・光に対する縮瞳や視線の動きなどの眼球反応の観測	・瞬き検出 ・顔の血流による皮膚の微小な色変化からの心拍の検知
虹彩	・可視光照明制御による瞳孔の収縮反応の誘発，観測	・虹彩に映り込む環境光の反射，瞳孔と虹彩の境界線の解析 ・瞬き検出 ・光の変化や視線移動に対する瞳孔の動き，形状の解析
静脈	・マルチスペクトル撮像による差異の解析	・複数枚の静脈画像からの脈動，血流の観測
音声	・認証時にチャレンジ音／超音波を混入（再生攻撃ではチャレンジ音が欠落／相違） ・ランダムに文章や数字列を提示するチャレンジ応答方式	・録音・再生機器固有の雑音，微小なエコー，量子化雑音の検知 ・音声変換や音声合成モデル特有のスペクトル的特徴の検知 ・話速やピッチの時間変動，間の取り方などのダイナミクスの精査 ・RNN，Transformer を応用した判別
共通	―――	・画像特徴の解析 　- 雑音やコントラスト等の画像品質指標からの判別 　- 印刷画像のモアレ，ドットパターン等の検出 　- ディスプレイ映像に生じる輝度むらや反射光，モアレ，ドットパターン等の検出 　- 微細特徴の欠落の検知 　- CNN や Vision Transformer を応用した判別

クセスコントロール用のデータベースから不正取得した顔テンプレートを使って，街頭の防犯カメラ映像から当該人物を追跡したり，別のシステムに保管された個人の健康情報を検索したりといった，セキュリティやプライバシーの深刻な脅威につながる．

これらの脅威を回避する手段として注目されているのが次章で解説するテンプレート保護技術である．

4　テンプレート保護技術

これまで解説したとおり，生体認証システムにおけるテンプレートの流出は，リプレイ攻撃や，元の生体情報の復元，ほかのシステムデータとのクロスマッチングなど，深刻な脅威につながる．一般に，パスワードや暗証番号，IC カードなどの記憶や所持に基づく認証手段で漏えいや紛失が発生した場合には，危たい化した認証情報（クレデンシャル）の無効化や更新という措置が取られるが，生体情報には変更が難しいという本質的な性質がある．テンプレート保護技術は，こうした問題を解決する技術として期待されている．

テンプレート保護技術の基本的なフレームワークは，ISO/IEC 24745:2022 で国際標準化されている．その中で，テンプレート保護技術に求められる要件は，次の三つとされている．

(1) 不可逆性（Irreversibility）：テンプレートから元の生体情報の類推や復元ができないこと．

(2) 更新可能性（Renewablity）：同じ生体情報から異なるテンプレートデータを生成できること．（危たい化したテンプレートを無効化し新たなテンプレートに更新できること．）

(3) リンク不可能性（Unlinkability）：異なる生体認証システム／データベースとのクロスマッチングに利用できないこと．

実運用の観点からは，上記三つの要件に加えて，テンプレート保護により認証精度が著しく劣化しないとする「精度の維持」を要件とする．本稿では，代表的なテンプレート保護技術として，特徴変換，バイオメトリック暗号，暗号化照合の三つのアプローチを解説する．

4．1　特徴変換（Feature Transformation）

特徴変換アプローチでは，生体特徴に不可逆な変換を適用し，変換後の特徴をテンプレートとして照合に利用する．2001 年に Ratha らによって，キャンセラブルバイオメトリクス (Cancelable Biometrics) の名で，概念

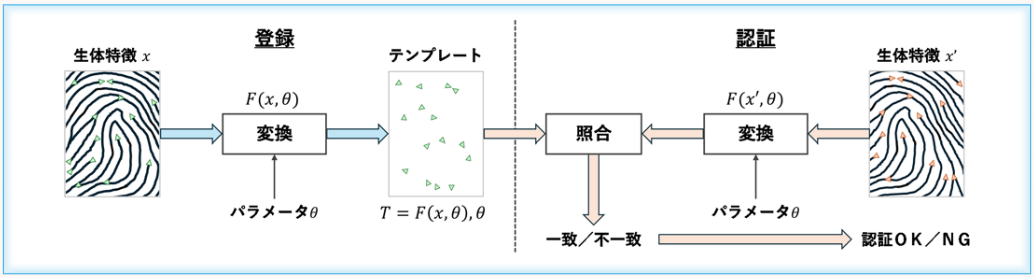


図３　特徴変換アプローチにおける登録と認証のフロー

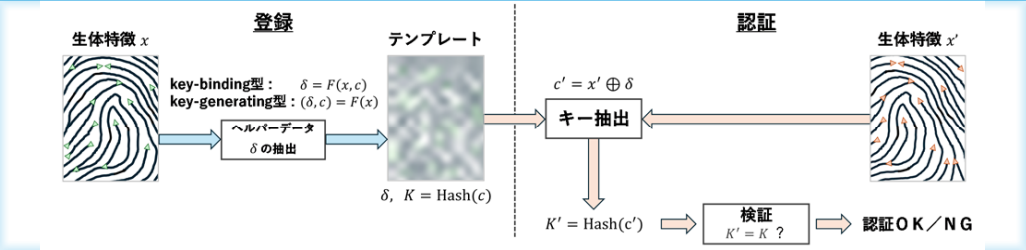


図４　バイオメトリック暗号を活用した生体認証システムの登録と認証のフロー

と基本アプローチが提案された[1]．

図３に本アプローチで構成する生体認証システムの登録と認証のフローを示す．生体特徴xを入力，θをパラメータとした変換関数$F(x,\theta)$によって生体特徴を変換し，変換した状態で照合を行う．θの変更によって生体特徴の変換結果を変更し，更新可能性とリンク不可能性を担保する．不可逆性は，変換関数の種別によって担保され，幾何的な変換やランダムプロジェクションなどの適用が研究されている[12], [13]．

幾何的な変換の場合，変換後の特徴量同士の照合に，一般的な特徴量照合アルゴリズムも適用できる場合があり，システムへの導入負荷が低いメリットがある．一方で，変換アルゴリズムが単純であったり，パラメータθの漏えいがあったりすると，元の生体特徴や，元の生体特徴と同じ変換特徴が得られる生体特徴パターンを推定される危険性がある．θは登録時だけでなく，認証時にも必要となることから，厳格な管理が必要となる．また，変換後の特徴量類似度に着目し，遺伝的アルゴリズムで元の生体情報を探索する攻撃[14]や，異なる変換パラメータで生成された複数のテンプレートから元の生体情報を再構成するARM（Attack via Record Multiplicity）攻撃[15]などが知られている．

4．2 バイオメトリック暗号（Biometric Cryptosystems）

生体特徴は，個人内でも揺らぎがあるため，生体特徴そのものを，例えば暗号鍵のように安定的なディジタル情報として扱うことは困難である．そこで，特徴量を量子化して，誤り訂正符号を用いることで生体特徴の揺らぎを吸収し，更に暗号技術の知見を生かして不可逆性，更新可能性，リンク不可能性を実現する手法として，バイオメトリック暗号が研究されている．

図４に本アプローチで構成する生体認証システムの登録と認証のフローを示す．登録時は，生体特徴を使って，鍵Kを安全に封入したヘルパーデータを作成する．認証時は，ヘルパーデータと生体特徴から鍵を抽出する．認証は鍵の検証によって行う．

バイオメトリック暗号には，生体特徴とは独立に鍵Kを定める鍵結合型（key-binding型）と，生体特徴から鍵Kを生成する鍵生成型（key-generating）がある．本稿では，鍵結合型と鍵生成型の代表的な手法として，Fuzzy Commitment[16]とFuzzy Extractor[17]について解説する．鍵結合型でよく知られている手法としてFuzzy Vault[18]もあるが，本稿では割愛する．

4．2．1　Fuzzy Commitment

Fuzzy Commitmentは，生体特徴と誤り訂正符号を利用して作成したコミットメントから，鍵情報を安定的に抽出する方法である．

基本的なアイデアを以下に示す．ビット量子化された生体特徴xと，ランダムに選択した誤り訂正可能な符号cから，次式でヘルパーデータδを算出する．

$$\delta = x \oplus c$$

$\oplus$はビットごとの排他的論理和演算を表す．今，個人内の揺らぎを含むビット量子化された生体特徴x'が得られたとし，x'とδから次式で符号c'を計算する．

$$c' = x' \oplus \delta$$

c'は，x'とxの差異をcに反映したものとなるため，xとx'の差異が許容範囲内にあれば，誤り訂正により符号c'から符号cが取得できる．しかるに，登録フェーズでは，生体特徴xと符号cから関数$F(x,c)$でヘルパー

データδと$K = \mathrm{Hash}(c)$を生成し，認証フェーズでは生体特徴x'とヘルパーデータδから$K' = \mathrm{Hash}(c')$を生成して，K'とKの一致判定で認証する仕組みが構築できる（図4）．

生体特徴の揺らぎが大きい場合，安定的に鍵Kを抽出するためには，誤り訂正符号による訂正能力を高める必要があるが，誤り訂正能力を過度に高めると，他人受入れの発生による精度低下や，計算量の増大といった問題が生じる．このため，できるだけ安定的な生体特徴取得の実現が，Fuzzy Commitment導入の課題となる．

4.2.2　Fuzzy Extractor

「鍵生成型」の本手法は，生体特徴xから鍵Kとヘルパーデータδを生成する鍵生成関数$\mathrm{Gen}(x)$と，揺らぎを含む生体特徴x'とヘルパーデータδから，鍵K'を生成する鍵再現関数$\mathrm{Rep}(x', \delta)$の二つの関数で構成される．

典型的な実装では，$\mathrm{Gen}(x)$は，誤り訂正符号Cの中から，xに最近傍の符号語$c \in C$を選択し，次式で鍵Kと，ヘルパーデータδを生成する．

$$K = \mathrm{Hash}(c),\quad \delta = x \oplus c$$

$\mathrm{Rep}(x', \delta)$は，揺らぎを含む生体特徴x'とヘルパーデータδから$c' = x' \oplus \delta$を計算し，c'を誤り訂正した上で，$K' = \mathrm{Hash}(c')$を計算する．

両関数を活用することで，図4に示した認証の仕組みが構成できる．また，生体情報xに由来して鍵生成ができることから，公開鍵暗号基盤（PKI）やブロックチェーンの署名鍵生成への応用も検討されている．署名鍵生成に関しては，Fuzzy Signatureなど，より発展的なアプローチも研究されている[19), (20)]．

Fuzzy Commitment同様に，特徴抽出や量子化に精度と工夫が求められるとともに，ヘルパーデータδから，生体特徴や鍵に関する情報が漏れないように設計することが重要となる．また，符号語cが漏えいすると，δとcから元の生体特徴xが復元できてしまうため，厳重な管理が必要となる．

4.3　暗号化照合（Biometrics in encrypted domains）

「暗号化照合」は，生体特徴を暗号化したまま安全に照合するアプローチで，準同形暗号（Homomorphic Encryption）を用いる方法や，セキュアマルチパーティ計算を用いる方法が広く研究されている．

準同形暗号とは，二つの暗号文$\mathrm{Enc}(m_1)$と$\mathrm{Enc}(m_2)$を復号することなく，$\mathrm{Enc}(m_1 + m_2)$や$\mathrm{Enc}(m_1 \times m_2)$などの演算ができる準同形性を有する暗号で，加算及び乗算の双方，及び組合せに準同形性を持つ完全準同型暗号（FHE: Fully Homomorphic Encryption）と，加算または乗算のいずれか一方にのみ準同形性を持つ部分準同形暗号（PHE: Partially Homomorphic Encryption）がある．生体認証技術においては，生体特徴の類似度計算に積和演算が用いられることが多いことからFHEの活用が期待されたが，FHEには計算コストが極めて高いという課題がある．現在は，FHEの計算効率改善研究と，より軽量で実装が容易なPHEや，加算は無制限であるが乗算には回数制限があるsomewhat準同形暗号（SHE: Somewhat Homomorphic Encryption）の活用研究が並行して進められている[21)]．

準同形暗号で暗号化した生体特徴に対し，暗号化したまま照合演算を適用した場合，照合結果もまた暗号化された状態となる．照合結果を取り出すためには，これを復号する必要があるが，この復号処理をマルチパーティ計算で行うのが，セキュアマルチパーティ計算による暗号化照合アプローチである．代表的な手法としては，Shamirの秘密分散法を用いて準同形暗号の復号鍵を複数のパーティに分散し，各パーティで部分復号する方式がある．各パーティの部分復号結果を集約してLagrange補間等の手法で再構成することで，照合結果を復元する．この方式では，照合に関与するいずれのパーティも部分的な復号鍵しか保持しないため，完全な復号鍵が漏えいし，結果として生体特徴や照合結果が漏えいしてしまうリスクを低く抑えることができる．

5　今後の展望

本稿では，生体認証技術／システムに対する攻撃と，その対策技術としてPADと，テンプレート保護技術に焦点を当てて詳説した．

昨今，特殊詐欺やソーシャルエンジニアリングの温床として，映像や音声を介したリモートコミュニケーションにおける「なりすまし」が社会的な問題となりつつあるが，本稿で解説したPAD技術は，こうした「なりすまし」検知への応用も期待される．また，日本企業をはじめとした事業者を中心に，本稿で解説したテンプレート保護技術や，より発展させた技術を応用し，セキュリティやプライバシーに配慮した認証基盤の構築・実用化を進める動きもある．生体認証の持続的な発展には、本稿で紹介した技術を活用しつつ，プライバシー保護設計（Privacy by Design）といった観点からの包括的な対応も必要となろう．

このほか，本稿では解説を割愛したが，FIDOやFIDO2など，PCやスマートフォンに搭載された生体認証技術を活用した安全なオンライン認証基盤の標準化も進んでいる[22)]．スマートフォンの普及に伴って，オンラインの公的サービスや金融サービス等における生体認

証活用にも，一層の拍車がかかっている．

このように，今後も生体認証技術の活用は拡大していくことが予想される．生体認証技術の安全性を理解する上で，本稿の内容が一助となれば幸いである．

■ 文献

(1) N. K. Ratha, J. H. Connell, and R. M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems, ” IBM Syst. J., vol. 40, no. 3, pp. 614-634, Jan. 2001.

(2) A. K. Jain, K. Nandakumar, and A. Nagar, “Biometric template security,” EURASIP J. Adv. Signal Process, Article 113, March 2008.

(3) W. Yang, S. Wang, J. Hu, G. Zheng, and C. Valli, “Security and accuracy of fingerprint-based biometrics: A review,” Symmetry 11, no. 2: 141, Jan. 2019.

(4) T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, “Impact of artificial "Gummy" fingers on fingerprint systems,” Proc. SPIE, vol. 4677, April 2002.

(5) K. Karampidis, M. Rousouliotis, E. Linardos, and E. Kavallieratou, “A comprehensive survey of fingerprint presentation attack detection,” J Surveill Secur Saf 2021;2:117-61, Oct. 2021.

(6) L. Debiasi, C. Kauba, H. Hofbauer, B. Prommegger, and A. Uhl, “Presentation attacks and detection in finger- and hand-vein recognition,” Proc. the Joint Austrian Computer Vision and Robotics Workshop 2020, pp.65–70, 2020.

(7) T. T. Nguyen, Q. H. Nguyen, D. T. Nguyen, D. T. Nguyen, T. Huynh-The, S. Nahavandi, T. T. Nguyen, Q. Pham, and C. M. Nguyen, “Deep learning for deepfakes creation and detection: A survey,” Computer Vision and Image Understanding, vol. 223, Aug. 2022.

(8) D. Sharma and A. Selwal, “A survey on face presentation attack detection mechanisms: hitherto and future perspectives,” Multimed Syst. Vol. 29, no. 3, pp. 1527-1577, March 2023.

(9) C. B. Tan, M. H. A. Hijazi, N. Khamis, P. N. E. binti Nohuddin, Z. Zainol, F. Coenen, and A. Gani, “A survey on presentation attack detection for automatic speaker verification systems: state-of-the-art, taxonomy, issues and future direction,” Multimedia Tools and Appl., vol. 80, pp. 32725–32762, Sept. 2021.

(10) S. Marcel, J. Fierrez, and N. Evans, Handbook of biometric anti-spoofing: presentation attack detection and vulnerability assessment, Springer, 2023.

(11) M. Gomez-Barrero and J. Galbally, “Reversing the irreversible: A survey on inverse biometrics,” Comput. & Security, vol. 90, no. 2, Dec. 2019.

(12) N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle, “Generating cancelable fingerprint templates,” IEEE Trans. Pattern Analysis and Machine Intell., vol. 29, no. 4, pp. 561–572, April 2007.

(13) A. B. J. Teoh, A. Goh, and D. C. L. Ngo, “Random multispace quantization as an analytic mechanism for BioHashing of biometric and random identity inputs,” IEEE Trans. Pattern Analysis and Machine Intell., vol. 28, no. 12, pp. 1892-1901, Dec. 2006.

(14) X. Dong, Z. Jin, and A. T. B. Jin, “A genetic algorithm enabled similarity-based attack on cancellable biometrics,” IEEE 10th International Conf. Biometrics Theory, Appl. and Syst.(BTAS), Sept. 2019.

(15) C. Li and J. Hu, “Attacks via record multiplicity on cancelable biometrics templates,” Special Issue: Security of new generation computing systems(NSS 2012), vol. 26, no. 8, pp. 1593-1605, June 2014.

(16) A. Juels and M. Wattenberg, “A fuzzy commitment scheme,” Proc. 6th ACM conf. Comput. and commun. security(CCS '99), Association for Comput, Mach., pp. 28–36, Nov. 1999.

(17) Y. Dodis, L. Reyzin, and A. Smith, “Fuzzy extractors: how to generate strong keys from biometrics and other noisy data,” EUROCRYPT 2004, Springer Lecture Notes in Comput. Sci., vol.3027, pp.523-540, 2004.

(18) A. Juels and M. Sudan, “A fuzzy vault scheme,” Designs, Codes and Cryptography, vol. 38, no. 2, pp. 237–257, Feb. 2006.

(19) K. Takahashi, T. Matsuda, T. Murakami, G. Hanaoka, M. Nishigaki, “Signature schemes with a fuzzy private key,” International J. Inf. Security, vol. 18, pp. 581–617, Feb. 2019.

(20) H. Higo, T. Isshiki, S. Otsuki, and K. Yasunaga, “Fuzzy signature with biometric-independent verification,” International Conf. Biometrics Special Interest Group (BIOSIG), pp.1-6, Sept. 2023.

(21) M. Yasuda, T. Shimoyama, J. Kogure, K. Yokoyama, and T. Koshiba, “Practical packing method in somewhat homomorphic encryption,” International Workshop Data Privacy Management 2013, Springer Lecture Notes Comput. Sci., vol. 8247, pp. 34-50, Jan. 2014.

(22) FIDO Alliance. https://fidoalliance.org/

高田直幸

1998 東北大・工・電気卒．2000 同大学院修士課程了．同年セコム株式会社入社．以来，生体認証，画像解析技術の研究に従事．2021～2023 バイオメトリクス研究専門委員会副委員長．現在，セコム株式会社 IS 研究所研究戦略部長．東京科学大技術経営専門職学位課程在学中．

サイバー攻撃の要因である「ぜい弱性」とは？
――ぜい弱性がもたらすリスクとぜい弱性管理について――

西村隆廣　Takahiro Nishimura　NRI セキュアテクノロジーズ株式会社

1 はじめに

インターネットを利用するサービスが停止すれば，私たちは友人への連絡もままならず，「口コミ」評価の良い飲食店を見つけることも，その店を予約することも，これまで行ったことのない新しい店までたどり着くこともできなくなってしまうだろう．もちろんこれらに限らず企業や組織の業務においても，現代の我々の生活のあらゆる部分において，インターネットは必要不可欠なものとなっている．

サービス停止のみならず，企業の顧客データや個人情報の漏えい，データの暗号化など，企業や個人に限らず社会にも大きな影響を及ぼすサイバー攻撃は，ニュースなどでも多く取り上げられている．多くの人が「ソフトウェアのぜい弱性を狙った攻撃」などという言葉を耳にした経験があるのではないだろうか．漠然と「ぜい弱性」という言葉を聞いたことがある方は多く，何となく理解しているかもしれないが，詳しく理解している方は少ないと思われる．サイバー攻撃発生の要因の一つとなるぜい弱性について，発生する原因や攻撃例，企業や組織における対策方法を中心に解説していく．

2 ソフトウェアのぜい弱性とは

ぜい弱性とは，システムやソフトウェアなどのセキュリティ上の欠陥のことを指す．その欠陥は不正アクセスやコンピュータウイルスなどのサイバー攻撃により，甚大な被害をもたらす原因の一つとされている．ぜい弱性は英語では Vulnerability と言い，攻撃の受けやすさ，被害の遭いやすさなどを意味する．コンピュータ用語ではなく，人の状態などに対しても使われる言葉である．

似たような言葉として「セキュリティホール」もあるが，こちらはシステムやソフトウェアの開発段階における設計ミスや，プログラムコードの書き間違いなどによって発生するセキュリティ上の欠陥を指す．対してぜい弱性は，管理態勢や人的ミスなども含めたシステム環境全体における欠陥を指す．つまり，セキュリティホールはぜい弱性の一種であり，不具合や設計ミスが原因で発生する欠陥のみを指している．

システム上の欠陥については「バグ」という言葉も一般的ではないだろうか．こちらはセキュリティにかかわらず，システムやソフトウェアが正常に動作しない場合に使用される言葉である．このバグも，ただ単に機能が想定どおりの動きをしない場合だけでなく，セキュリティに影響がある場合にはぜい弱性と呼ばれる．バグは想定どおりの動きをしているかどうかテスト段階で見つけることは難しくないのだが，ぜい弱性やセキュリティホールは通常の使い方をする分には問題が発生しないため，開発側が気付きにくいと言われている（図 1）．

3 ソフトウェアぜい弱性が発生する原因

ぜい弱性が発生する原因とは何か，その発生原因を対策すれば，ぜい弱性によって引き起こされるサイバー攻撃を防ぐことができるのではないだろうか．ぜい弱性の主な発生原因は以下のようなものが挙げられる．

- 開発段階における設計ミス
- プログラムコードの書き間違い
- OS や基幹システムの更新に伴い発生する欠陥
- 使用しているソフトウェアの設計，または更新に伴い発生する欠陥
- OS やソフトウェアのアップデートや修正パッチの未適用
- 運用や設定のミス
- 新たなサイバー攻撃手法の登場により，従来正常であった状態がぜい弱性となるケース

これら全てを含め漏れなく対応ができていれば，ぜい弱性の発生を防ぐことができるかもしれない．しかし，

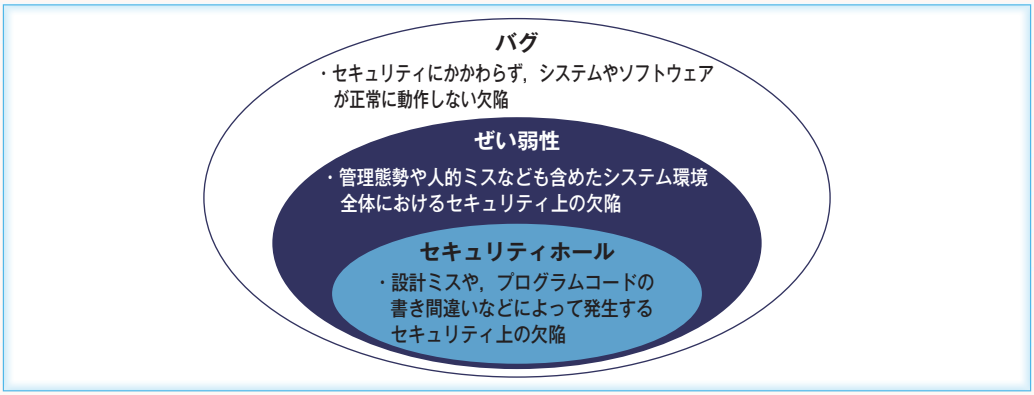


図 1　ぜい弱性，セキュリティホール，バグの関係図

サイバー攻撃の進化によって日々新たな攻撃手法が生み出され続けており，想定し得ないことがぜい弱性として見つかるなど事前の対策は非常に難しい．

既に見つかったぜい弱性については，OS やソフトウェアのアップデート，修正パッチを早期に適用することにより，サイバー攻撃の被害に遭う可能性を低くすることが可能である．個人であればアップデートや修正パッチを即適用することはそう負担が掛からないことかもしれない．しかし，企業や組織などで様々なシステムを利用していればいるほど，アップデートや修正パッチによりほかのシステムに影響がないか入念な確認が必要なため，大きな負担となる．そのためぜい弱性が放置されてしまい，放置されたぜい弱性を突かれてサイバー攻撃の被害に遭う企業や組織も少なくない．

4　ソフトウェアぜい弱性がもたらす影響と攻撃例

サイバー攻撃の多くはぜい弱性を悪用して行われており，ぜい弱性を放置することでサイバー攻撃の被害に遭う可能性は高まり，事業活動に甚大な影響を及ぼすこととなる．本章では，具体的な攻撃例について解説する．

4．1　マルウェア感染

マルウェアとは，不正かつ有害な動作を行う意図で作成されたコンピュータウイルス，トロイの木馬，ワーム，スパイウェアなどを含む，悪意のあるソフトウェアや悪質なコードの総称である．

ぜい弱性を放置することで，マルウェアに感染する可能性が高まることは確実である．PC がマルウェアに感染すると，勝手に外部と接続して操作されたり，PC 内のファイルが破壊や暗号化されたり，外部へ持ち出されるなどの被害に遭う．企業や組織にとっては業務に支障をきたすばかりでなく，存続も危ぶまれるような大きな損害額となる可能性もある．

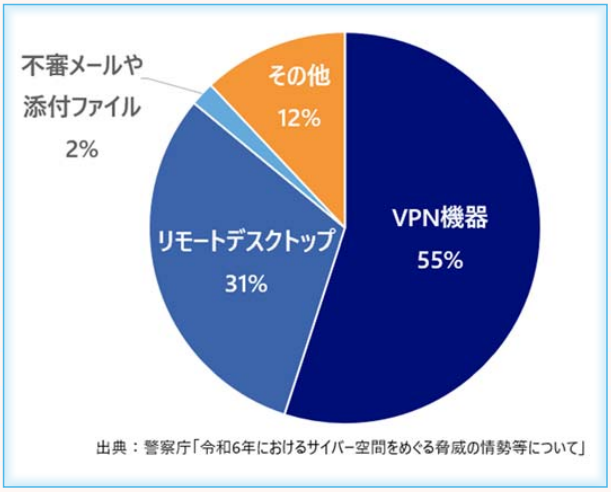


図 2　ランサムウェア感染経路

昨今のマルウェアの多くは OS やソフトウェアのぜい弱性を利用して感染するため，ぜい弱性を解消するためのアップデートやパッチの適用は極めて重要である．

4．2　不正侵入

ここでいう不正侵入とは，攻撃者による直接的な端末やサーバへの不正侵入を伴う侵害を指す．侵害の影響度が高いことから，ランサムウェアを用いた侵害が国内外で非常に注目されており，警察庁は「ランサムウェア被害において，VPN やリモートデスクトップ用の機器からの侵入が，全体の感染経路の 8 割以上を占める状況である」と報告[*1] している（図 2）．

ランサムウェアの攻撃者は，VPN などのぜい弱性のある機器を攻撃の足掛かりとして企業や組織のネットワークに侵入する．そしてドキュメントファイルなどを「人質」として暗号化し，復号のために被害者に金銭を要求する．攻撃を受けた企業や組織は，被害拡大を防止

＊1　警察庁「令和6年におけるサイバー空間をめぐる脅威の情勢等について」

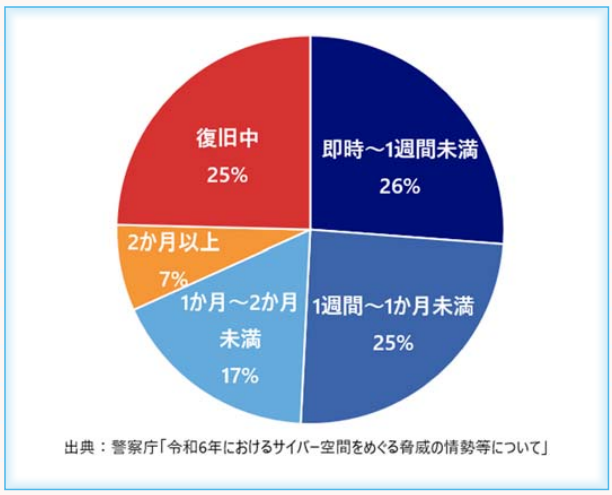


図３　ランサムウェア被害の復旧等に要した期間

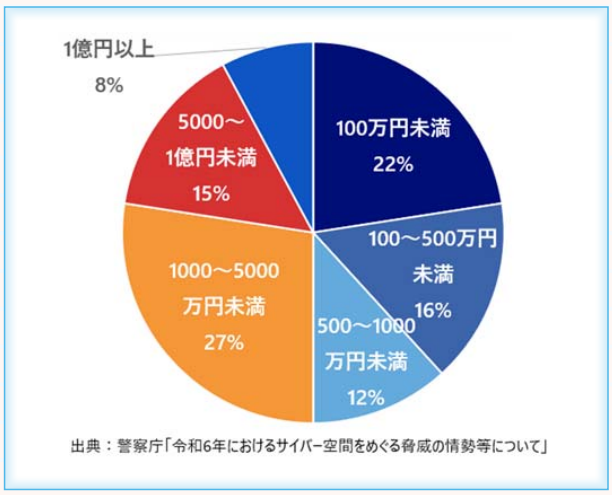


図４　ランサムウェア被害の復旧等に要した費用

するためにシステムを停止せざるを得なくなる場合や，業務の継続に影響が及ぶ場合がある．そのため企業や組織は自分たちが使用している機器を把握し，ぜい弱性やアップデートに関する情報を確認し，ぜい弱性を放置することなくアップデートやパッチの適用を確実に実施することが重要である．システム保守を外部委託している場合は，ぜい弱性の対処が保守契約に含まれているかを確認し，その対処が適切に実施されていることを確認することも重要である．

4．3　データの盗聴

ここでいうデータの盗聴とは，企業が公開しているWebサイトなどにおいてデータが盗まれることを指す．Webサイトにおいて後述の**4．5．1**のSQLインジェクションのぜい弱性がある場合，Webサイト上の入力フォームに不正な文字列を挿入することで，不正にデータベースにアクセスすることができてしまう．そのためデータベース内の個人情報などが盗まれ，情報漏えいの被害につながる．

Webアプリケーションのぜい弱性診断を行い，見つかったぜい弱性に対策を施すことでこの被害を未然に防ぐことが可能である．

4．4　システムの停止

ぜい弱性を突かれてシステムに不正侵入された場合，調査や復旧までの間，システムやサービスを停止せざるを得ない状況に追い込まれる場合も少なくない．ランサムウェアによる企業のシステム停止のニュースは記憶に新しいのではないだろうか．その企業が提供する動画像配信サービスは，復旧までに実に２か月もの時間を要した．同社は被害拡大防止のためにサーバをシャットダウンしたが，攻撃者により遠隔からシャットダウンしたサーバを起動されるという執ような攻撃を受けた．そのためサーバの電源ケーブルや通信ケーブルを物理的に抜線したことで，データセンタ内のサーバ全てが使用不可となったことを公表している．犯行グループが発表した声明をみると，ここまで被害が大きくなった原因は仮想化基盤のぜい弱性であったと言われている．

企業や組織において，システム停止が発生すれば信頼の失墜のみならず，機会損失などの被害も甚大となる（図3，4）．

4．5　Webサイトの改ざん

Webサイトの改ざんとは，文字どおりWebサイトが悪意のある何者かによって管理者の意図しない何らかの内容に書き換えられてしまうことや，Webサイト上に悪意のあるファイルなどを埋め込まれてしまうことを指す．最悪の場合，改ざんされたサイトを閲覧した人のPCにマルウェアを感染させてしまうなど，自らが加害者となってしまう可能性もある．Webサイトの改ざんの原因となるぜい弱性を利用した攻撃として，主に以下の三つがある（表1）．

4．5．1　SQLインジェクション

Webアプリケーションのぜい弱性を利用し，Webサイト上の入力フォームにSQL文を注入（インジェクション）することで，データベースに保存されたデータを不正に読み取ったり，改ざんしたりする攻撃である．決済を伴うECサイトや会員登録を行うポータルサイトなど，SQLサーバを利用するサイトが標的となる．

この攻撃の対策としては，プレースホルダの利用，エスケープ処理，Webアプリケーションのバージョンを最新に保つこと，WAF（Web Application Firewall）の導入などがある[*2]．

表 1　Web サイトのぜい弱性を突いたサイバー攻撃

名称	手法	対象
SQL インジェクション	Web サイト上の入力フォームに SQL 文を注入（インジェクション）し，データベースに保存されたデータへの不正アクセスや改ざんなどを行う	決済を伴う EC サイトや，会員登録を行うポータルサイトなど
クロスサイト・スクリプティング	不正な JavaScript や HTML などを含んだ URL をクリックさせ，偽ページなどを表示させる	アンケートサイトやサイト内検索，ブログ，掲示板など
クロスサイト・リクエストフォージェリ	悪意のある URL をユーザにクリックさせ，第三者がユーザになりすまし不正な操作を行う	ログインを必要とする，オンラインバンキング，SNS，EC サイトなど

4.5.2　クロスサイト・スクリプティング

Web アプリケーションのぜい弱性を利用し，不正な JavaScript や HTML などを含んだ URL をクリックさせることで，ユーザの Web ブラウザに偽ページなどを表示させる攻撃である．アンケートサイトやサイト内検索，ブログ，掲示板などといったユーザからの入力内容を基に Web ページを生成するサイトなどが標的となる．

この攻撃の対策としては，サニタイジング（スクリプトの無害化），入力値の制限，WAF の導入などがある[*2]．

4.5.3　クロスサイト・リクエストフォージェリ

Web アプリケーションのセッション管理のぜい弱性を利用し，悪意のある URL をユーザにクリックさせることで，第三者がユーザになりすまして不正な操作を行う攻撃である．ログインを必要とするオンラインバンキング，SNS，EC サイトなどが標的となる．

この攻撃の対策としては，トークンの利用，リクエストヘッダの確認などがある[*2]．

これら Web サイトの改ざんに関わる攻撃は，全て **4.3** のデータの盗聴の対策と同様である．すなわち Web アプリケーションのぜい弱性診断を行い，見つかったぜい弱性に対策を施すことでサイバー攻撃の被害を未然に防ぐことが可能である．

5　ぜい弱性情報の収集・活用，確認・対策方法

私たちが日々の業務や生活の中で使用する様々なシステムは，一つのシステムの中で多数のソフトウェアが稼動している．ぜい弱性がもたらすサイバー攻撃の被害をできるだけ防ぐためには，自組織が利用するシステムに内在するソフトウェアを全て把握し，それらのぜい弱性情報を入手し，その情報を評価した上で，ぜい弱性への対応を行う必要がある．その一連のプロセスを「ぜい弱性管理」と呼ぶ．ぜい弱性管理を行うことで自組織のセキュリティが向上するばかりでなく，セキュリティ対策状況が可視化され，万が一サイバー攻撃被害に遭ったとしても，その対処を効率的に行うことができるようになる．その結果，被害を最小限に抑えることが期待できる．

5.1　ぜい弱性情報の収集

ぜい弱性は主に，ソフトウェアの開発ベンダなどのホワイトハッカーが探したり，攻撃者が実際に悪用したりすることで見つかる．こうして見つかったぜい弱性の情報は，公開データベースに登録されることが多い．主要なデータソースは以下のとおりである（表 2）．

5.1.1　JPCERT/CC

JPCERT/CC（一般社団法人 JPCERT コーディネーションセンター）[*3] は日本におけるコンピュータセキュリティインシデントへの対応を行う組織である．国内外のセキュリティ関連情報を収集し，セキュリティアドバイザリーやぜい弱性情報，注意喚起を公開している．

5.1.2　IPA

IPA（独立行政法人情報処理推進機構）[*4] は日本の IT 国家戦略を技術面・人材面から支えるために設立された独立行政法人である．本機構の所管官庁は経済産業省であり，セキュリティインシデントへの対応支援，ぜい弱性情報の収集と公開，セキュリティ意識の向上を図るための啓発活動などを実施している．JPCERT/CC とともに日本国内における情報セキュリティの中心的な役割として機能している．

5.1.3　JVN

JVN（JP Vendor Status Notes）[*5] は経済産業省告示「ソフトウェア等脆弱性関連情報取扱基準」を受けて，日本国内の製品開発者のぜい弱性対応状況を公開するサイトである．JPCERT/CC と IPA により共同で運営さ

＊2　これらは具体的な実装方法の説明となるため，本解説では割愛する．
＊3　JPCERT/CC（https://www.jpcert.or.jp/）
＊4　独立行政法人情報処理推進機構：IPA（https://www.ipa.go.jp/）
＊5　Japan Vulnerability Notes（https://jvn.jp/）

表２　ぜい弱性情報の主要なデータソース

名称	概要
JPCERT/CC （一般社団法人 JPCERT コーディネーションセンター）	日本国内の情報セキュリティインシデントに対応するための民間のコンピュータセキュリティインシデント対応チーム（CSIRT）
IPA （Information-technology Promotion Agency, Japan：独立行政法人 情報処理推進機構）	日本政府（所轄官庁は経済産業省）が設立した情報セキュリティや IT 人材育成の推進を担う独立行政法人
JVN （Japan Vulnerability Notes）	JPCERT/CC と IPA が共同運営する日本で使用されているソフトウェアなどのぜい弱性関連情報とその対策情報を提供するポータルサイト
JVN iPedia	IPA が運営する国内外問わず日々公開されるぜい弱性対策情報を，JVN，JPCERT/CC，NVD などと情報と連携して，収集・蓄積することを目的としたぜい弱性情報データベース
NVD （National Vulnerability Database）	NIST（米国国立標準技術研究所）が運営するぜい弱性情報データベース
Exploit Database	OffSec 社が運営する実際に確認されたセキュリティぜい弱性を利用した攻撃コード（Exploit）や対策方法を収集・公開しているオープンな情報共有データベース
OWASP （Open Web Application Security Project）	Web アプリケーションやソフトウェアのセキュリティを向上させることを目的とした米国の非営利団体で，Web アプリケーションのセキュリティに関するガイドラインやツール，フレームワークを無償で提供

れている．

5．1．4　JVN iPedia

JVN iPedia[*6] はぜい弱性情報データベースである．JVN が新たなセキュリティ情報をいち早く提供するプラットホームであるのに対し，JVN iPedia は日々発見されるぜい弱性対策情報を蓄積し，アーカイブ機能や検索機能を実装して幅広く利用することを目的としている．

5．1．5　NVD

NVD（National Vulnerability Database）[*7] は NIST（米国国立標準技術研究所）によって管理されているぜい弱性情報データベースである．世界で最も広く利用されているぜい弱性情報のソースである．

5．1．6　Exploit Database

Exploit Database[*8] はセキュリティに関する教育やテストなどを提供する OffSec 社が，実際に使用可能なぜい弱性のエクスプロイトコードや攻撃手法を収集したオンラインデータベースである．セキュリティ研究者が新たに発見されたぜい弱性の悪用方法を共有するためのプラットホームで，システムのセキュリティ検証を行う際の実践的なリソースとして利用されている．

5．1．7　OWASP

OWASP（Open Web Application Security Project）[*9] は Web アプリケーションのセキュリティに関する包括的な情報，ガイドライン，ツール，リソースを提供しており，セキュリティ対策の基準として開発者やセキュリティ専門家に広く活用されている．

これら信頼性のある情報源から収集した，ぜい弱性情報について，主に以下が確認ポイントとなる．

- CVE 識別子（ぜい弱性に割り当てられる一意の ID）
- ぜい弱性の概要
- 影響を受ける製品／バージョン
- CVSS スコア（ぜい弱性の深刻度）
- 修正情報（アップデートやパッチなどの対応策）

5．2　ぜい弱性情報の活用

収集したぜい弱性情報から，自組織内で影響を受ける可能性のあるシステムやアプリケーションを特定する必要がある．次にそのぜい弱性が自組織に与えるリスクを評価し，優先して対応すべきか否かといった対応方針を決定する．

ぜい弱性情報は毎日多数公開されており，その全ての情報を収集してリスクを評価することは大変労力のいる仕事である．リスクを評価して対応優先度を決定するに

＊６　ぜい弱性対策情報データベース：JVN iPedia（https://jvndb.jvn.jp/）
＊７　NIST National Vulnerability Database（https://nvd.nist.gov/vuln）
＊８　Exploit Database（https://www.exploit-db.com/）
＊９　Open Web Application Security Project: OWASP（https://owasp.org/）

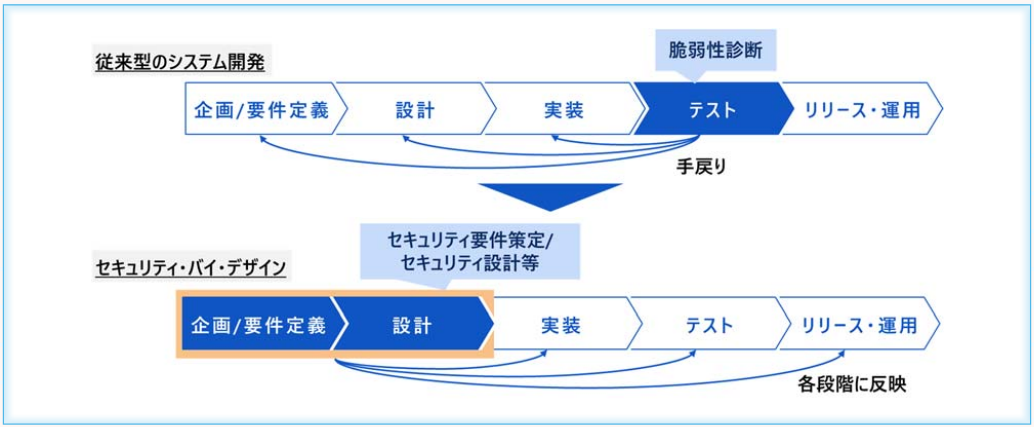


図５　セキュリティ・バイ・デザイン推進による安全性向上のイメージ

は高度な専門的知識も必要となるため，ぜい弱性管理においては外部の専門家を利用する企業や組織も少なくない．

5.3　ソフトウェアのぜい弱性の確認

日々新たな攻撃が生まれているため，自組織が利用するソフトウェアにぜい弱性が潜んでいないか，定期的な確認も必要である．

ぜい弱性を確認する方法には，ソースコード解析，ぜい弱性スキャン，ペネトレーションテストなどがある．しかし，専門知識を要するため自組織のみで実施するハードルは高く，こちらも外部の専門家による実施が一般的である．

5.4　ソフトウェアのぜい弱性の対策

常に新しいぜい弱性情報を収集し，自組織に影響のあるぜい弱性を見つけ，それを評価し，対応方針を決め，対応を実施する．また，自組織が利用するシステムに内在するソフトウェアにぜい弱性が潜んでいないか，定期的に確認して対応を実施する．ソフトウェアのぜい弱性の対策としてできることは，ぜい弱性管理プロセスの整備と，そのプロセスにのっとった継続的な対応を地道に行うことである．言葉にするのは簡単だが，多くの企業や組織にとって，これらに漏れなく対応するスキルを持ち合わせた人材，予算の確保などは非常に困難であろうと思われる．

ぜい弱性を完全になくすことはできないが，ぜい弱性を減らす方法として「セキュリティ・バイ・デザイン」という考え方がある．次章でその考え方について解説する．

6　セキュリティ・バイ・デザイン

システムやソフトウェア，製品の開発初期段階からセキュリティを考慮し，設計プロセスに組み込む考え方，

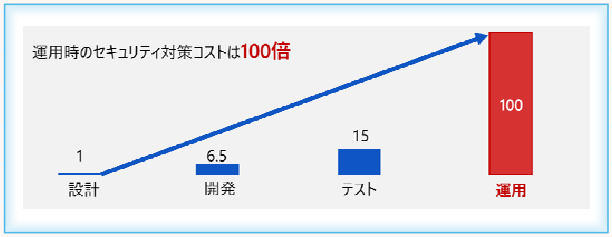


図６　ぜい弱性修正コストのイメージ

セキュリティ・バイ・デザインについて解説する．

セキュリティ・バイ・デザインという概念は新しいものではなく，2000 年頃には登場していたと言われている．従来はリリース前のぜい弱性診断を軸としてセキュリティ対策を行うことが主流だった．しかし，リリース直前や運用中に見つかったぜい弱性を修正するには，システム全体を再設計することが必要な場合もあるため，相当なコストや時間を要することもある．そのため，ぜい弱性診断によってぜい弱性が発見されても，緊急度の高いぜい弱性は修正するが，コストやリリースまでの時間の制約から，低リスクのぜい弱性は修正しないケースが多いのが実態である．

こうした状況を防ぐため，セキュリティ・バイ・デザインを採用することで，企画・設計段階で検出可能なぜい弱性が存在しない状態にすることが可能となる（図 5）．大幅なコスト削減効果を望むことができ（図 6），従来見過ごされていた低リスクのぜい弱性や，リスクを受容していた比較的リスクの高いぜい弱性にまで対策が施され，セキュリティ品質が向上する．

近年の DX（ディジタルトランスフォーメーション）やディジタルビジネスの進展に伴い，セキュリティ・バイ・デザインの重要性が改めて認識されており，デジタル庁が 2022 年に「政府情報システムにおけるセキュリティ・バイ・デザインガイドライン[*10]」を発行している．

2024 年には，金融庁が公表した「金融分野における

2.3.4.3. セキュリティ・バイ・デザイン
基本的な対応事項
①金融商品・サービスの企画・設計段階から，セキュリティ要件を組み込む「セキュリティ・バイ・デザイン」を実践すること．サービス全体の流れの中で，重要なサードパーティも含めてリスクを検証し対策を講ずること． ②自組織にシステムを提供する重要なサードパーティにおいて，セキュリティ・バイ・デザインを実施できる体制となっているかを確認すること．
対応が望ましい事項
a．セキュリティ・バイ・デザインにかかる管理プロセスを，以下の点も考慮のうえ，整備し，運用すること． ・セキュアコーディングに係る基準を策定し，実施すること． ・データの機密性，アクセス管理，イベントログ取得等のセキュリティ要求事項を明確化していること． ・セキュリティ技術・アーキテクチャーに係る設計標準を策定すること． ・アプリケーションソフトウェアのリリース前及びリリース後定期的に，アプリケーションのぜい弱性診断を実施すること． ・システムへのぜい弱性の混入及び作込みを未然に防止し，早期に検知するツール（ソースコード解析ツール等）を活用すること．

図7　金融庁が公表した「金融分野におけるサイバーセキュリティに関するガイドライン」の要件

サイバーセキュリティに関するガイドライン[*11]」においても，セキュリティ・バイ・デザインの対応が要件の一つとして盛り込まれている（図7）．本要件では企画・設計段階だけではなく，開発・運用工程における事項も挙げられている．本ガイドの要件に準拠するためには，システム開発のライフサイクル全体でセキュリティ対策を考慮していく必要があると考えられる．

セキュリティ・バイ・デザインは概念であり，決まった型はなく，企業や組織によって対応策は様々で大きく異なることもある．まずは小さな取組みからでも，セキュリティ・バイ・デザインを実施して頂きたい．

7　おわりに

残念ながら，ぜい弱性を完全になくすことはできない．なぜなら技術の進歩，攻撃手法の進化，意図しない操作などとともに，新たなぜい弱性が常に発見されるためである．

しかし，ぜい弱性を放置すればその分セキュリティリスクは高まり，事業継続の危機はもちろん，社会的混乱を及ぼす可能性もある．ぜい弱性対策の重要性を理解し，万が一サイバー攻撃の標的となったとしても，被害を最小限に食い止めることができるようにぜい弱性に挑んで頂きたい．

※10　デジタル庁（https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/e2a06143-ed29-4f1d-9c31-0f06fca67afc/7e3e30b9/20240131_resources_standard_guidelines_guidelines_01.pdf）
※11　金　融　庁（https://www.fsa.go.jp/news/r6/sonota/20241004/18.pdf）

西村隆廣

2002 野村総合研究所に入社．2009 NRI セキュアテクノロジーズに出向後，セキュリティコンサルティング業務に従事．セキュリティ戦略の企画・評価・マネジメント，サイバーセキュリティ態勢整備，サイバーセキュリティ分野の先進的な事業開発等を担当．

ディープフェイク音声の脅威と対策について

鵜木祐史　Masashi Unoki　北陸先端科学技術大学院大学

1 はじめに

情報通信技術や AI（人工知能）技術の急激な発展により，我々を取り巻くマルチメディア情報通信の利用形態が劇的に変化し，それと同時に社会的脅威も急激に高まっている．これらの技術は我々の生活に様々な面でメリットをもたらし生活を豊かにする一方で，個人情報の漏えいによる悪用といったリスクにさらされている．

例えば，音声メディアに焦点を当ててみると，10 数年前までは，違法コピー・違法配信が主たる社会問題であり，近い将来，ディジタルフォレンジクス（ディジタル機器に残された情報を収集・解析し，証拠として活用する技術）における音声改ざん等が問題になると予想されていた[(1)]．しかし，近年では，図 1 に示すように，深層学習ベースの音声合成技術で制作された偽物の合成音声（以後，「ディープフェイク音声」と呼ぶ）による「なりすまし」といった社会的脅威が急激に高まっている[(2)]．そのため，音声メディアを安心安全に利用するための仕組みや，音声に含まれるプライバシー情報をディープフェイクのような攻撃から守る革新的技術の確立が喫緊の課題になっている．

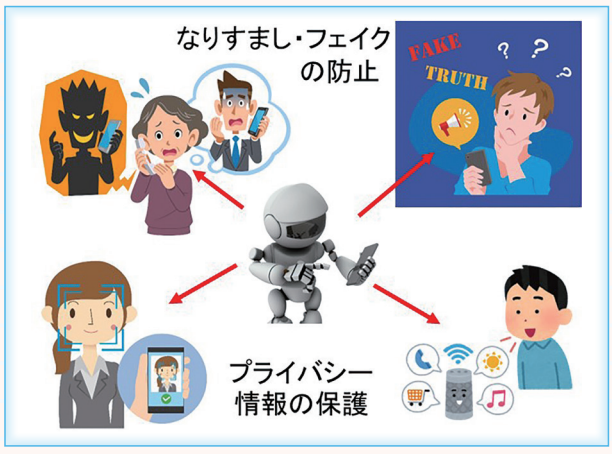


図 1　音声セキュリティの課題

本稿では，ディープフェイク音声による脅威を紹介し，その対策と課題を概説する．また，最新の研究成果を紹介し，現在の動向と残された課題を検討する．

2 音声合成技術の歴史の概観と現在の展開

音声合成技術は，元々フェイク音声を作成するために開発されたわけではない．音声合成技術が目指すゴールは，「福祉応用」など豊かな人間社会形成の一助になることである．例えば，将来何かの理由（病気等による発話器官の障がいなど）でしゃべれなくなったとしても，発話内容をテキストで打ち込むことで，自分の声でその内容を機械にしゃべらせることができるようになることである．その結果，QoL（Quality of Life）の向上にもつながり，皆ハッピーになるというわけである．ところが最近の展開をみるとそうでもないような気がしている．

音声合成技術の研究は，コンピュータの発展とともに進歩してきた．詳細な歴史の説明は専門書に譲るが（全体を俯瞰し，AI 音声合成技術までを含む教科書が見当たらないため，ウィキペディアの音声合成[(3)] を参照して頂きたい．），れい明期の頃は機械方式で合成音声の実現を試み，人間の発話生成機構や機能をまねる形で検討されてきた．その後，ボコーダが発明され，波形接続・音声素片合成形式，統計的手法（HMM: Hidden Markov Model 音声合成），そして深層学習ベース音声合成技術へと発展した．深層学習ベースと一くくりにしているが，波形生成モデル WaveNet やエンドツーエンド方式の音声合成器，ニューラルボコーダの登場など，その発展スピードは急激である．現在は，大規模音声発話データを活用して学習することで，自然で高品質の合成音声が提供できるようになり，図 2 に示すような応用展開に至っている．

最近の深層学習ベースの音声合成技術では，本人の大規模音声発話データがあれば，本人そっくりの自然な音

図２　深層学習ベース音声合成技術の応用先

声を合成できるところまできた．現在はそれを，本人の少数音声発話データから実現できるところまで競争が激化しているところである．その結果，本人のディジタル分身を作る技術（例えば，小池百合子都知事の広報キャラクター，AI ゆりこ 2.0[4]）やサイバー空間でのアバタ音声合成，更には故人の声の復活（AI 故人[5]）という話にまで展開されている．

このような展開が今後どこまで広がるのか予想もつかないが，一つだけ言えることは，これらの技術が悪用されることにより，なりすましといった社会問題を増長しかねないことである．今後は，倫理観をもってこれらの技術を使わなければならないのかもしれない．

3　なりすまし音声の脅威

2015 年 9 月，音声科学のトップカンファレンスの一つである Interspeech 2015 のスペシャルセッションにおいて，ASVspoof チャレンジ[6] が初めて開催された．このチャレンジでは，図 3 のように，音声合成技術や音声変換技術による音声なりすまし，あるいは収録音声（本人に同意なく盗聴された音声）と本物の声を判別するアルゴリズムの開発を通じて，なりすまし攻撃とその対策を独立に評価することを狙いとした．これまでに 5 回（ASVspoof2015，ASVspoof2017，ASVspoof2019，ASVspoof2021，ASVspoof2024）開催され，なりすまし音声対応やシステムの攻撃耐性の強化が検討されてきた．

ASVspoof2015 で最も驚いたのは，なりすまし音声の検出性能にうんぬんする前に，話者認証システムが当時の合成音声によって突破されたことであった．これは過度に評価されたこともあったかもしれないが，音声合成によるなりすまし攻撃が非常に脅威と感じた最初のきっかけでもあった．当然ながら回を増すごとに，合成音声の品質も向上し，なりすまし音声も巧妙化されている印象を持った．

このようなチャレンジにより，学術的にはなりすまし攻撃やその脅威への対策は練られているものの，現実的には下記のようななりすまし音声による詐欺事件やオンライン選挙活動被害の実例も起こっている．

・AI を悪用した音声詐欺が世界で増加[7]

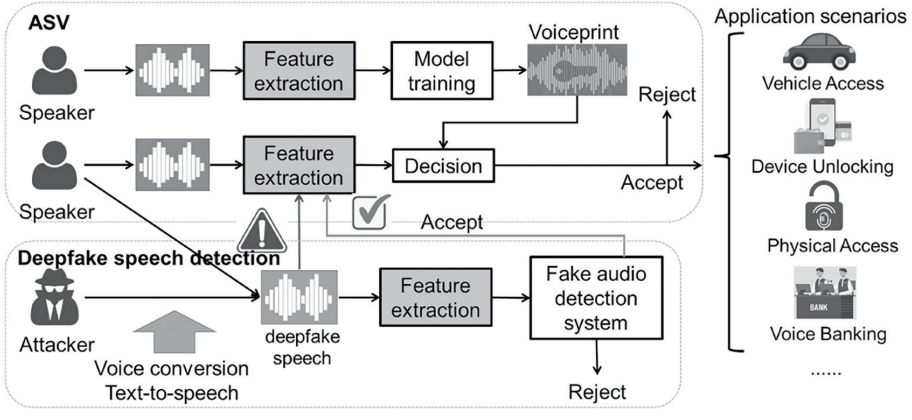


図３　なりすまし音声の対策

・「AI 音声を詐欺に悪用　3 秒の素材で親族らになりすまし」[8]
・前澤友作氏や堀江貴文氏に（声まで）なりすました詐欺広告[9]
・ディープフェイクボイスによるなりすまし（AI 音声によるなりすまし詐欺）[10]
・「AI で音声など無断生成　声優らがルール作りを訴え啓発動画」[11]
・「生成 AI が『オレオレ詐欺』を助長？　技術革新が犯罪を高度に」[12]
・「バイデン大統領の AI『なりすまし音声』，男を起訴　罰金 9 億円も」[13]

これらは氷山の一角に過ぎず，現実的にはもっと多くの事件が起こっている可能性がある．また，今後はもっと複雑で大きな社会問題を引き起こす可能性もある．

4　ディープフェイク音声への対応

4．1　チャレンジ

ディープフェイク音声検出に限定した最初のチャレンジは ADD 2022（Audio Deepfake Synthesis Detection 2022）である[14]．これは，IEEE ICASSP 2022 信号処理グランドチャレンジの一環として開催された．

ADD 2022 では，人間の音声パターンを正確に模倣できるディープフェイク生成技術の高度化に焦点を当てている．これにより，検出システムにとって本物の音声とディープフェイク音声の区別が非常に困難なものになっている．第 2 回チャレンジ（ADD 2023）[15] も開催されており，世界中の研究者がディープフェイク音声検出のための革新的技術の開発を目指している*．

ディープフェイク音声検出に関するほかの調査報告として，Chen ら[16] や Wang ら[17] の報告がある．Chen らは，新しい音声合成方法が絶えず登場しているため，検出モデルが様々なディープフェイク音声に共通するパターンを特定しながら，新しい音声合成技術によってもたらされる変動に，耐性を持つことが難しくなっていることを指摘している．Wang らは，ディープフェイクの制作者がその検出を回避するために敵対的攻撃を使用する可能性があり，ディープフェイク音声検出が雑音に対して頑健でなければならないと指摘している．

Yi らは，ADD 2023 の調査を通じて，ディープフェイクスピーチ検出のためのラベル付きデータセットの入手が困難であり，真正サンプルとディープフェイクサンプルの分布が不均衡となり，ディープフェイク音声の検出性能が最適化されない可能性があることを示した[18]．また，ディープフェイク合成音声検出モデルでは，短い発話箇所でも正確な予測を行うため，より多くのコンテキスト情報を必要とする．そのため，これらの知見はディープフェイク音声の制作者による悪用につながる可能性がある．このような課題に対処するためには，ディープフェイク合成音声検出の継続的な研究開発が必要である．

4．2　データベース

ディープフェイク音声検出のための主要なデータベースは，図 4 に示すように，ASVspoof 2019/2021 と ADD 2022/2023 である．

ASVspoof 2019 は，なりすまし対策の比較評価プ

＊　ASVspoof は様々な種類のなりすましを対象にしているが，ADD は音声合成によるなりすましに限定している．

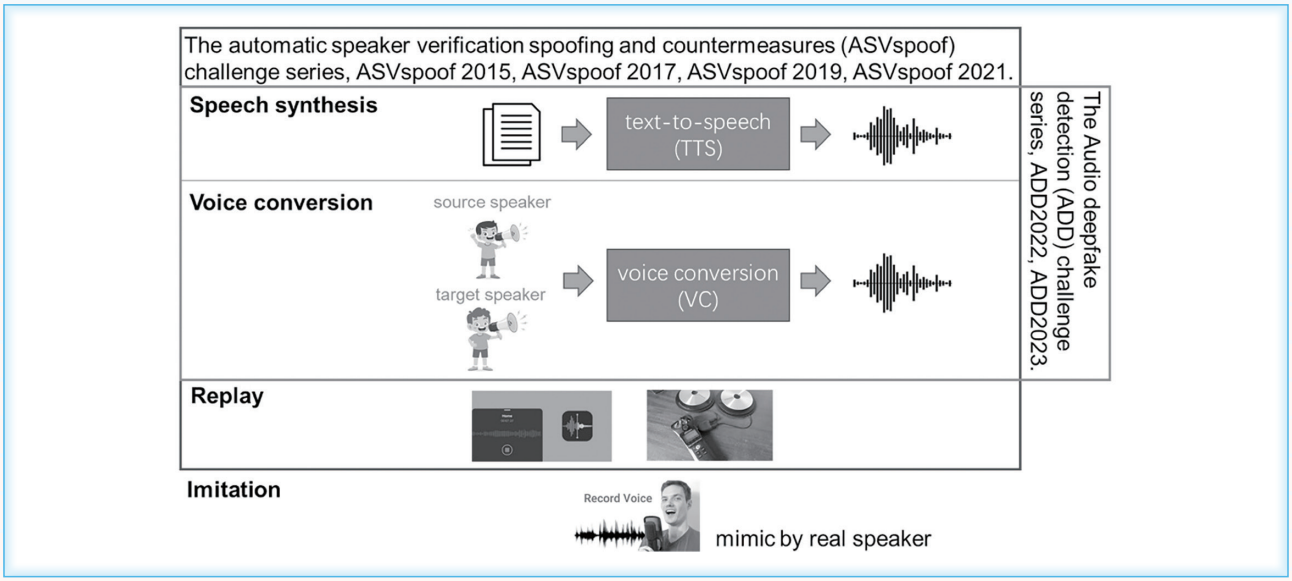
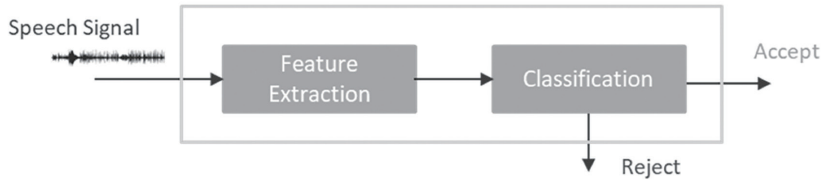


図 4　ディープフェイク音声データベース



Feature extraction	Classification
1. Linear frequency cepstral coefficients (LFCC)	Gaussian mixture model (GMM) [Baseline]
2. Constant-Q CC (CQCC)	GMM [Baseline]
3. CQCC and Spectrogram	CNN
4. Constant Q Transform (CQT)	ResNet-based
5. LFCC	LCNN

図５　代表的なディープフェイク音声検出の概要

ラットホームを作成するために使用されたものであり，正規の音声信号と偽装された音声信号の標準化されたデータセットが含まれている[19]．ASVspoof 2019 は，なりすましの種類を論理アクセス（音声変換及びテキスト音声合成）と物理アクセス（様々な環境におけるリプレイ攻撃）の二つのシナリオに再概念化されている[19), (20)]．

ASVspoof 2021 には，オンライン音声データに対するなりすまし音声検出の頑健性を評価するために，ディープフェイク音声という別のシナリオが追加されている[3]．

ADD データセットは，ディープフェイク音声の検出と分析に使用されている．ADD 2022 には，低品質フェイク音声検出（LF），部分的フェイク音声検出（PF），音声フェイクゲーム（FG）の三つのトラックがある．LF 音声検出は，実世界の雑音の中で完全に偽の発話を識別することに重点を置いている．PF 音声検出は，部分的に偽の音声を識別することを対象とする．FG 音声検出は，生成と検出の二つのタスクから成る敵対的ゲームである[14]．

ADD 2023 では，三つのサブ課題（FG，操作領域の位置，ディープフェイクアルゴリズムの認識）を通じて，部分的な偽音声の区間と偽音声の発生源を特定することを課題としている[15]．

４．３　最新の方法

ディープフェイク音声検出で利用される特徴量は，短時間・長時間特徴量，プロソディの特徴量，及び最先端の検出技術で利用されるディープ特徴量に分類される．

音声信号の短時間フーリエ変換（STFT）から得られる短時間スペクトル特徴には，対数振幅スペクトルや線形予測符号化など，大きさや位相に基づく特徴がある．定 Q 変換（CQT），ヒルベルト変換，ウェーブレット変換などの長時間スペクトル特徴は，時間周波数パターンを分析するために使用される[18]．特に，両スペクトル特徴において，ケプストラムベースの特徴として，固有周波数ケプストラム係数（LFCC），メル周波数ケプストラム係数（MFCC），CQT ベースのケプストラム係数（CQCC）が使用されている．

韻律的特徴には，基本周波数（F0），継続時間，パワー変化，話速など，より長い音声区間が含まれる．このような手作業で作成された特徴量はロバストであるが，バイアスに悩まされることもある．

ディープニューラルネットワーク（DNN）を介して学習されるディープ特徴は，学習可能なスペクトル，教師あり埋込み，及び自己教師あり埋込み特徴に分類される[18]．広く使用されている DNN には，Wav2vec や HuBERT がある．Zaman ら[21] はこれらの要素を組み合わせて，オーディオ分類で一般的に使用される深層学習アーキテクチャによって既存のモデルを分類している．更に，スタンドアロンの深層学習モデルは，サポートベクターマシン（SVM）や K 近傍法（KNN）といった従来の分類器と組み合わせることも可能である．

４．４　結果と限界

上述した代表的なディープフェイク音声検出の概要を図５に示す．ADD 2022/2023 チャレンジで提案された数多くの検出器の性能評価については，トラックごとに評価されランク付けされている．評価尺度は原則，等価エラー率（EER）が利用され，この値が限りなく 0 に近づくことを狙いとしている．いずれも高い検出性能を達成しているところがあるが，十分に満足のいくものではなく，それぞれのトラックで個別に今後の課題が述べられている状況である[14), (15)]．

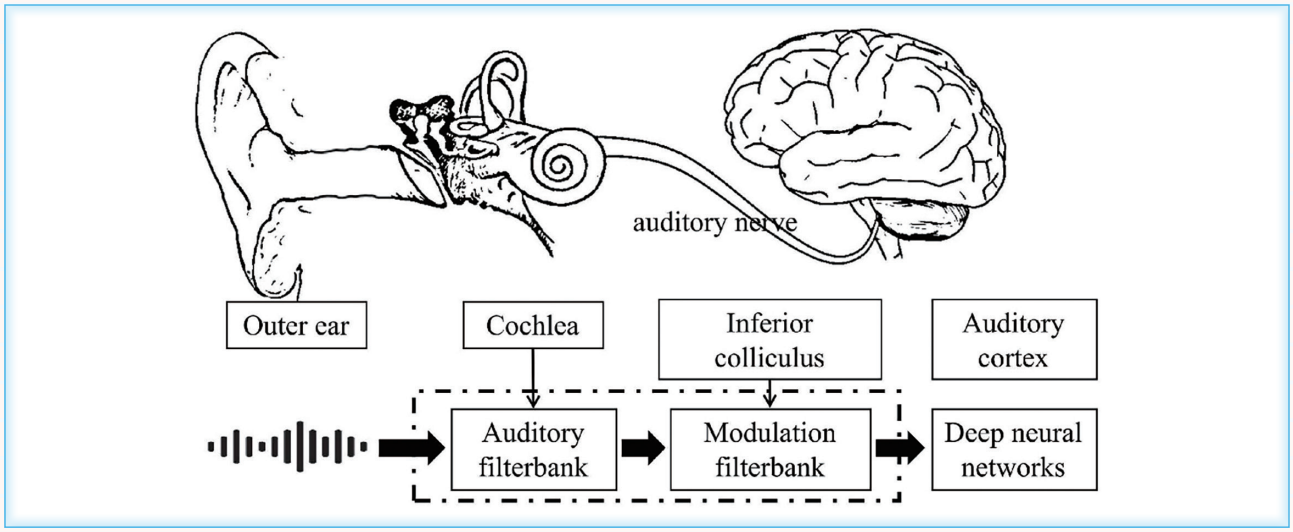


図６　聴覚のスペクトル変調・時間変調分析の概要

ディープフェイク音声検出において多くの特徴抽出技術法が実装されているが，まだ幾つかの欠点が残されている．従来の音響特徴のほとんどは，話者の個人性や言語情報などの一般的な情報を抽出するように設計されているため，真正音声と音声合成で作られた偽音声を区別する識別力が限られている．更に，フロントエンドの特徴抽出手法法は，入力音声の雑音（ノイズ）やばらつきの影響を受けやすく，特に録音条件や背景雑音が変化する実世界のデータを扱う場合，性能が低下する可能性がある．

これらの特徴のほとんどは，本物と偽物の音声の違いを視覚化したり，なぜそのような違いが生じるのかを説明したりするために使用できるものにはなっていない．また，ADD では合成音声＝ディープフェイク音声と限定しているが，ASVspoof では合成音声以外のフェイクを対象にしていることにも注意しておきたい．

5　人間の聴知覚メカニズムに基づく方法

筆者らのグループでは，上述の問題から，本物と偽物を判別する仕組みまで言及できるように，図 6 に示すメカニズムから説明可能な人間の聴知覚メカニズムに基づくディープフェイク検出法を検討した．

- 聴覚フィルタバンクを利用した音の時間周波数分析とそのスペクトル変調・時間変調（STM）表現を利用したディープフェイク音声検出法[22]
- 音声のプロソディに関わる基本周波数の時間変動でみられるジッター・シマーを音響特徴としたディープフェイク音声検出法[23]
- 聴知覚の主な違いとして音色評価に関わる特徴（シャープネス，ラフネス，ブライトネス，ハードネスなど 8 種類）を利用したディープフェイク音声検出法[24]
- フェイク音声を病理音声の一つと解釈した場合の知覚評価に関わる特徴（ジッター・シマー，声帯音源対雑音比など）を利用したディープフェイク音声検出法[25]

いずれの方法も ADD 2022/2023 や ASVspoof 2019/2021 のデータベースを対象に EER で評価した．検出性能だけみると，いずれの方法もこれらのチャレンジで提案された最新の方法の検出精度には負けてしまうが，聴知覚に関する音響特徴の段階で，本物と偽物の違いをおおむね説明できるという大きな利点を有する．これらの音響特徴の次元は深層学習をベースとする検出法の特徴の次元に比べて大幅に低くなることから深層学習ベースの識別器に整合し難いところも問題である．

一方で，ADD や ASVspoof では，ディープフェイク音声，なりすまし音声を対象に議論していたが，両方で考慮されていない，なりすまし音声もある．それは，人間による「ものまね音声」（Imitated speech）である．我々は，ものまね音声をどれだけ人間が正確に弁別できるか，22 人の大学院生を対象に 20 人の話者を対象とするものまね音声検出実験を行った．ここでは，実験参加者は 20 人の話者を学習し，その話者識別率が 90 % を超えるまで学習を行わせた．2 人がこの作業でギブアップし，話者識別率 90 % を超えた 20 人の学生がものまね音声検出実験に進んだ．その結果，検出精度は 70.1 % であった[26]．

この結果に対し，典型的なディープフェイク検出器（特徴としてメルスペクトルを利用）では，検出精度が 51 % であったのに対し，聴知覚に関係する音色指標を利用したところ検出精度が 62 % であった．人間の検出能力にはまだ追い付いていないが，前述したように聴知覚メカニズムを考慮する検出方法の優位性を確認できる．

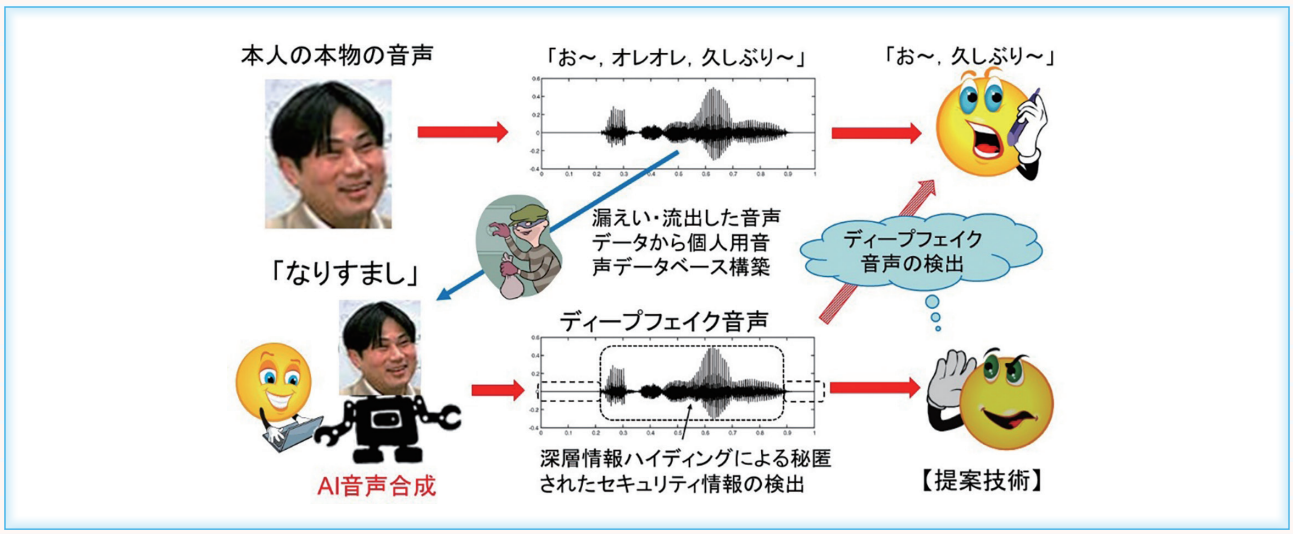


図７　ディープフェイク合成音声の検出

6 今後の展開

前述したように ADD 2022/2023 では，「深層学習ベースの音声合成で作成されたもの＝ディープフェイク音声」として検出対象とした．ASVspoof は音声合成のほか，音声変換や収録音声（リプレイ攻撃）をなりすまし音声の対象とした．いずれも重要な課題を検討するためのシナリオに沿ったものである．しかし，この方針でいくと，元々，利用目的のあった音声合成が全て「悪」として捉えられてしまうことになる．

筆者がこれまでに力を入れて取り組んできたマルチメディアセキュリティの一つである「音声情報ハイディング技術」の立場からすると，むしろ安全な合成音声を利用するならば，「合成音声マーク」のような安全装置を合成音声に秘匿しておくことがよいのではないかと考える．例えば，図７に示すように，ネットを流通するようなディジタル音声（あるいは合成音声）に事前に秘密情報を秘匿しておき，それが悪意ある操作を受けることで，秘匿情報が壊され，それを検出器で確認することで改ざんの有無を確認できる．また，データの真正性を担保したい場合は，同様に秘密情報を秘匿しておき，それをもって正しい利用である宣言をすれば，ディープフェイク音声のような攻撃に対しては一定の防御策になり得るものと考える．

既にこのようなアプローチもみられるようで，深層学習ベースの音声合成器に知覚不可能で頑健な音声情報ハイディング技術を組み込むことが，今後のなりすまし音声に対する大きな防御策になるものと予想する．

7 おわりに

本稿では，ディープフェイク音声といった音声なりすましによる脅威を紹介し，その対策と課題を概説した．また，最新の研究成果を紹介するとともに，筆者の取組みと現在の動向，残された課題を紹介した．

■ 文献　（サイト閲覧日は 2025 年 6 月 16 日）

(1) 鵜木祐史，“聴覚特性に基づいた音響情報ハイディング技術，”信学 FR 誌，vol. 13, no. 4, pp. 284-293, April 2020.
(2) J. Yamagishi, X. Wang, M. Todisco, M. Sahidullah, J. Patino, A. Nautsch, X. Liu, K. A. Lee, T. Kinnunen, N. Evans, and H. Delgado, “ASVspoof 2021: accelerating progress in spoofed and deepfake speech detection,” Proc. Workshop Automatic Speaker Verification and Spoofing Coutermeasures Challenge, Sept. 2021.
(3) https://ja.wikipedia.org/wiki/%E9%9F%B3%E5%A3%B0%E5%90%88%E6%88%90
(4) 毎日新聞，“小池知事そっくり「AI ゆりこ 2.0」新動画公開　11 カ月ぶり更新，”June 1, 2025. https://mainichi.jp/articles/20250530/k00/00m/040/323000c
(5) 日経ビジネス，“AI で故人を再現，倫理的に許される？，”June 12, 2025. https://business.nikkei.com/atcl/gen/19/00648/060900008/
(6) Automatic speaker verification and spoofing countermeasures challenge（ASVspoof）. https://www.asvspoof.org/
(7) マカフィー PR Times，“AI（人工知能）を悪用した音声詐欺が世界で増加中，”May 17, 2013. https://prtimes.jp/main/html/rd/p/000000032.000033447.html
(8) 日本経済新聞，“AI 音声を詐欺に悪用　3 秒の素材で親族らになりすまし，”Sept. 11, 2023. https://www.nikkei.com/article/DGXZQOUE104850Q3A710C2000000/
(9) FNN プライムオンライン，“動画と声で「堀江貴文」「前澤友作」かたる投資勧誘，”April 1, 2024. https://www.fnn.jp/articles/-/685739?display=full
(10) CNN，“AI 生成のクローン音声使ったなりすまし詐欺，”Sept. 19, 2024. https://www.cnn.co.jp/tech/35224072.html

(11) NHK, “AI で音声など無断生成 声優らがルール作りを訴え啓発動画,” Oct. 21, 2024. https://www3.nhk.or.jp/news/html/20241021/k10014615291000.html
(12) 日経 BOOK PLUS, “生成 AI が「オレオレ詐欺」を助長？　技術革新が犯罪を高度に,” Jan. 22. https://bookplus.nikkei.com/atcl/column/011400459/011400002/
(13) 朝日新聞, “バイデン大統領の AI「なりすまし音声」、男を起訴　罰金 9 億円も,” May 24, 2025. https://www.asahi.com/articles/ASS5S03CSS5SBQBQ0GLM.html
(14) J. Yi, R. Fu, J. Tao, S. Nie, H. Ma, C. Wang, T. Wang, Z. Tian, Y. Bai, C. Fan, S. Liang, S. Wang, S. Zhang, X. Yan, L. Xu, Z. Wen, H. Li, Z. Lian, and B. Liu, “ADD 2022: the first audio deep synthesis detection challenge,” Proc. ICASSP2022, pp. 9216–9220, Feb. 2022.
(15) J. Yi, J. Tao, R. Fu, X. Yan, C. Wang, T. Wang, C. Y. Zhang, X. Zhang, Y. Zhao, Y. Ren, L. Xu, J. Zhou, H. Gu, Z. Wen, S. Liang, Z. Lian, S. Nie, and H. Li, “ADD 2023: the second audio deepfake detection challenge,” Proc. IJCAI2023 Workshop on Deepfake Audio Detection and Analysis, pp. 125-130, May 2023.
(16) T. Chen, A. Kumar, P. Nagarsheth, G. Sivaraman, and E. Khoury, “Generalization of audio deepfake detection,” Proc. Odyssey, pp. 132–137, Nov. 2020.
(17) R. Wang, F. J-Xu, Y. Huang, and Q. Guo, “Deepsonar: Towards effective and robust detection of ai-synthesized fake voices,” Proc. 28th ACM international conf. multimedia, pp. 1207–1216, May 2020.
(18) J. Yi, C. Wang, J. Tao, X. Zhang, C. Y. Zhang, and Y. Zhao, “Audio deepfake detection: A survey,” ArXiv Preprint, ArXiv:2308.14970, Aug. 2023.
(19) M. Todisco, X. Wang, V. Vestman, M. Sahidullah, H. Delgado, A. Nautsch, J. Yamagishi, N. Evans, T. Kinnunen, and K. A. Lee, “ASVspoof2019: Future horizons in spoofed and fake audio detection,” arXiv preprint arXiv:1904.05441, April 2019.
(20) X. Wang, J. Yamagishi, M. Todisco, H. Delgado, et al., “ASVspoof2019: A large-scale public database of synthesized, converted and replayed speech,” Comput. Speech & Language, vol. 64, no. 2, p. 101114, May 2020.
(21) K. Zaman, M. Sah, C. Direkoglu, and M. Unoki, “A survey of audio classification using deep learning,” IEEE Access, vol. 11, pp. 106620-106649, Sept. 2023.
(22) H. Cheng, C. O. Mawalim, K. Li, L. Wang, and M. Unoki, “Analysis of spectro-temporal modulation representation for deep-fake speech detection,” Proc. APSIPA ASC 2023, pp. 1822-1829, Nov. 2023.
(23) K. Li, X. Lu, M. Akagi, and M. Unoki, “Contributions of jitter and shimmer in the voice for fake audio detection,” IEEE Access, vol. 11, pp. 84689-84698, Aug. 2023.
(24) A. Chaiwongyen, N. Songsriboonsit, S. Duangpummet, J. Karnjana, W. Kongprawechnon, and M. Unoki, “Contribution of timbre and shimmer features to deepfake speech detection,” Proc. APSIPA2022, pp. 97-103, Nov. 2022.
(25) A. Chaiwongyen, s. Duangpummet, J. Karnjana, W. Kongprawechnon, and M. Unoki, “Potential of speech-pathological features for deepfake speech detection,” IEEE Access, vol. 12, pp. 121958-121970, Aug. 2024.
(26) K. Zaman, I. J. A. M. Samiul, K. Li, Y. Uezu, S. Kidani, and M. Unoki, “Ability of human auditory perception to distinguish human imitated speech,” IEEE Access, vol. 13, pp. 6225–6236, Jan. 2025.

鵜木祐史（正員：フェロー）

北陸先端大・先端科学技術研究科・教授，生体機能・感覚研究センター長．専門は聴覚科学・音声科学・音信号処理．1999 北陸先端大・情報科学・博士後期課程了．博士（情報科学）．2023-2025 日本音響学会副会長．2025 EMM 研究会委員長．現在，音声情報ハイディング技術を利用した音声なりすまし対策の課題に鋭意取り組んでいるところである．

暗号の安全性評価の重要性

國廣　昇　Noboru Kunihiro　筑波大学

1　あなたの暗号は安全ですか？

身の回りを見渡してみると，世の中には暗号があふれていることが分かる．インターネットバンキングは既に広く利用されており，オンラインストアによって，クレジットカードを用いた商取引が一般的になっている．これらのサービスを安全に利用するためには，当然，安全性が担保されていなければならない．しかし，広く認識されているとおり，我々の周りには多くの脅威があふれている．これらの脅威の詳細な内容については，情報処理推進機構（IPA）が毎年公表している「情報セキュリティ 10 大脅威」[(1)] などを参照されたい．2025 年 6 月現在でも，証券会社のインターネット取引への不正アクセス・不正取引が問題になっており，それを受けて，多くのサービスで多要素認証の必須化などの対策が進んでいる．その際，その対策が本当に有効であるか，安全であるかの厳密な評価が必要となる．せっかく対策を施しても，安全でなければ，意味がない．たとえ高度な暗号を使って安全な通信路を設けたとしても，暗号とは関係ないところで安全性が破られることも多い．セキュリティ技術は，一番弱いところが狙われるため，穴を埋めることが重要である．そのような状況ではあるものの，セキュリティの基盤となる暗号技術がぜい弱であれば，その暗号技術によって立つセキュリティシステムもぜい弱なものとなる．セキュリティシステムの構築において，使用している暗号は「当然安全である」という前提で議論することになる．

ケルクホフスの原理というものが知られている．「攻撃者は仕組みを知っている」という前提で対策をすべきである，という原理である．そのため，暗号方式自体を隠すことにより安全性を担保しようとすることは，この原理の下では許容されない．

秘匿通信を考えてみよう（図 1）．ケルクホフスの原理の下では，暗号化及び復号のアルゴリズムを公開することになる．復号する権限を持つ者だけが所有する復号鍵のみを隠すことにより，安全性が保証されなくてはならない．この枠組みにおいては，暗号化鍵は必ずしも秘密にする必要はないことに注意されたい．復号鍵は秘密にしなければならないため，暗号化鍵と復号鍵が同一である場合には，暗号化鍵も秘密にしなければならない．そのため，暗号化鍵と復号鍵は，送信者と受信者の共通の秘密となる．この方式は，「共通鍵暗号」や「対称暗

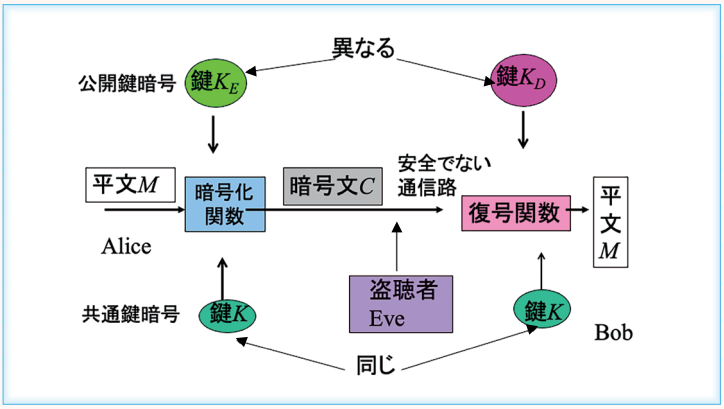


図 1　公開鍵暗号と共通鍵暗号

号」と呼ばれることが多い．暗号化鍵と復号鍵が異なる場合には，暗号化鍵を秘密にせず，公開しても問題ない．更に，利便性のために，積極的に暗号化鍵を公開することが多い．そのため，「公開鍵暗号」と呼ばれる．

暗号化鍵で暗号化し，復号鍵で復号処理をすると平文が復元されるため，暗号化鍵と復号鍵の間には強い関係がある．そのため，公開されている暗号化鍵を利用すると，その関係性を用いて復号鍵が分かってしまうかもしれない．公開鍵を公開しても復号鍵が分からないようにするために，何らかの数学的な困難な問題（RSA 暗号*においては素因数分解問題，ElGamal 暗号においては離散対数問題）の困難さを仮定し，その仮定の下で安全性が担保されることになる．ここで，適切な設定（RSA 暗号においては，2,048 bit の合成数を利用するなど）を行うことにより，現実的な時間では解読できないようにすることが重要である．適切な鍵長の設定については，CRYPTREC 暗号技術報告書[(2)] などを参照されたい．

2 暗号の安全性評価

暗号の安全性評価において，攻撃者は，公開されている情報や通信を盗聴することによって得られる情報は自由に使ってよい．まずは，このように受動的に得られる情報のみを使った攻撃を考える．

現在広く利用されている RSA 暗号[(3)] は素因数分解の困難さに安全性の根拠を置いている．素因数分解は，15=3×5 というように，合成数が与えられたときに，素数の積に分解するものである．合成数が小さい場合には，2，3，5 と小さい順に素数で割り切れるかを確認することにより，容易に素因数分解を行うことが可能である．しかし，この単純な方法では，大きい合成数の場合には，膨大な計算時間が必要である．それでは，もっと効率的なアルゴリズムではどうであろうか？ 素因数分解は，古くから研究されているが，現在のところ，素因数分解を行う多項式時間アルゴリズムは知られていない．その一方で，現在知られている最適なアルゴリズムは，指数関数時間を必要とせず，多項式時間と指数関数時間の間に属する準指数関数時間というクラスに属している．2025 年 6 月現在で，素因数分解された最も大きな RSA 型の合成数（同じサイズの二つの異なる素数の積）は，RSA-250 という 250 桁・829 bit の合成数である．図 2 に RSA-250 の素因数分解を示す．RSA 型の合成数の場合，これより大きい合成数は，依然，素因数分解はされていない．

この攻撃では，RSA 暗号を真正面から解読する場合を考えている．つまり，公開されている情報のみから暗号を解読する状況である．これ以降では，攻撃者は，公開されている情報だけでなく，様々な副次的な情報を入手できる状況を考え，「サイドチャネル攻撃」として総称される攻撃を取り上げる．この攻撃では，例えば，消費電力などの物理的な何らかの量を観測し，その情報を用いることによって，秘密情報を復元することがゴールとなる．

以降の説明のために，教科書的な RSA 暗号を簡単に説明する．

鍵生成：

Step1: nビットの異なる素数 p,q を生成．$N=pq$ を計算．

Step2: $ed \equiv 1 \pmod{(p-1)(q-1)}$ となる e, d を設定

暗号化鍵を (N, e)，復号鍵を d とする．

暗号化：平文を m として，暗号文 $C=m^e \bmod N$ を計算する．

復号：暗号文を C として，平文 $m=C^d \bmod N$ を計算する．

＊　提案者 3 人 Rivest，Shamir，Adleman の頭文字から名付けられた．

RSA-250=
21403246502407449612644230728393335630086147151447550177977549
20881418023447140136643345519095804679610992851872470914587687
39626192155736304745477052080511905649310668769159001975940569
34574522305893259766974716817380693648946998715784949759374979 37
=
64135289477071580278790190170577389084825014742943447208116859 6
32024532344630238623598752668347708737661925585694639798853367
×
33372027594978156556226010605355114227940760344767554666784520 9
87023841729210037080257448673296881877565718986258036932062711

図 2　RSA-250 の素因数分解

RSA 暗号の復号処理は，二乗計算と乗算計算により，実現される．この計算を実現する最も基本的な計算法は，「バイナリー法」（若しくは「Square-and-Multiply 法」）と呼ばれる方法である．秘密のべき指数（RSA 暗号においては，秘密鍵 d）を 2 進数展開した上で，上位ビットから値を読んでいき，

1. ビットが 1 であるときには，二乗演算（Square）をした上で乗算演算（Multiply）を行い，
2. ビットが 0 であるときには，二乗演算（Square）を行う，

ことにより行われる．

「単純電力解析」と呼ばれる攻撃を説明する．復号処理中の電力波形を観測することにより，二乗演算と乗算演算の区別ができる状況を考える．多くの場合，二乗演算と乗算演算は，異なる処理であるため，原理的にもこの二つの演算は区別可能である．この観測により，二乗演算と乗算演算がどの順番で行われたかの演算系列を知ることができる．更に，この演算系列を基に秘密鍵を復元することが可能である．これは，秘密鍵の各ビットが 0 であるのか，1 であるのかによって行う演算が異なること，そして，それを攻撃者が識別可能であることから攻撃に成功する [(4)]．

3 ダメな ATM の事例

暗号の安全性を評価する際に，仕様どおりであれば安全であるが，仕様どおりに実装をしないために，想定よりも弱くなってしまうことがある．このような実装を用いたために，秘密情報に依存した挙動が観測されることにより，ぜい弱性が生じることがある．多くの人になじみのある銀行 ATM での認証を基に，説明をしたい．

ATM では，通常キャッシュカードと 4 桁の暗証番号を用いた認証が行われる．（なお，キャッシュカードは所持情報，暗証番号は知識情報による認証であるため，多要素認証とみなすことができる．）キャッシュカードを ATM に挿入した上で，4 桁の暗証番号を入力し，暗証番号が事前に設定したものと一致すれば，正しい利用者であると確認され，現金の引出し等の処理が行われる．一致しない場合には，単に認証エラーであることを返す（図 3）．暗証番号を適切に設定した上で，たまたま，暗証番号を当てられてしまう確率は，1/10,000 であり，十分小さい確率である．そのため，高々 10,000 回，平均でも 5,000 回程度の繰返しで暗証番号を当てることができる．多くの場合，3 回暗証番号の入力を間違えると，取引が停止されるため，適切に暗証番号が設定されていれば，3 回以内に暗証番号が当たる確率は極めて小さく，安全であるとみなすことができる．しかし，暗証番号を適切に設定していない場合（例えば，1234 などの簡単なものや誕生日に設定するなど．）は，当然，安全性は保証されない．

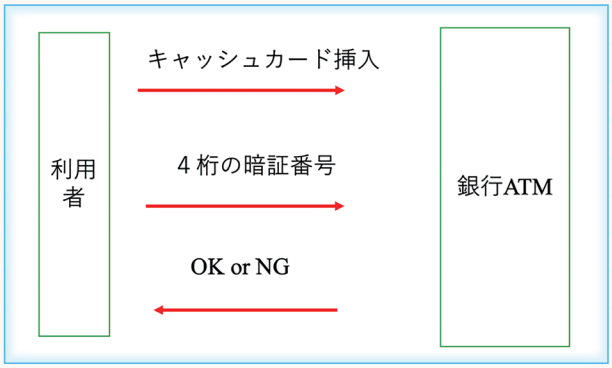


図 3　通常の銀行 ATM の認証

以下の二つの少し変わった（ダメな）ATM を考えてみる．いずれも，サイドチャネル攻撃とみなすことができる攻撃により，安全性が損なわれる．

一つ目は「おせっかい」な ATM である．この ATM は，暗証番号を間違えた場合，単に間違えていることを伝えるだけでなく，4 桁の暗証番号を入力し終えた段階で，何桁目を間違えているかのエラーメッセージを返してくれるものである．このような状況は考えにくいかもしれないが，利用者からの苦情（間違えた場合でも，ヒントを教えるべきであるなど．）や（間違った）利便性の向上を理由として，このようなシステムを導入してしまうということがあるかもしれない．この ATM の場合では，エラーメッセージを利用することで，高々 10 回の暗証番号の入力で認証に成功してしまう．正しい実装の場合は平均 5,000 回程度必要であったものが，格段に減少しており，ぜい弱性が生じている．この ATM の教訓としては，エラーが生じた（暗証番号が事前に設定したものと異なる）場合，エラーが生じている事実だけを伝えるべきであり，暗証番号のヒントとなり得る情報を与えてはならないということである．この ATM の例では，いかにもヒントになりそうな情報をエラーメッセージとして回答しているが，ヒントになりそうもないが，本来不要なエラーメッセージを回答してしまうために，ぜい弱性が生じた例がある．PKCS #1 v1.5 に対する Bleichenbacher によるミリオンメッセージ攻撃が，その代表例である．

二つ目の例は「せっかちな」ATM である．この ATM は，4 桁全ての暗証番号の入力を待たずに，暗証番号を間違えた段階で，エラーが生じたことを報告するものである．暗証番号を 1 桁でも間違えた段階で，既に，この試行は失敗することが確定しており，このタイミングでエラーを返しても，認証システムとしては表面上何も問題はない．それどころか，認証に要するトータ

ルの時間を考えると，この実装の方が有用であると考えて，このような実装をしてしまうことがあるかもしれない．この例では，明示的に，一つ目の例のようにどのようなエラーが生じたかを返しているわけではなく，単に，エラーが生じたことだけを返していることに注意されたい．しかし，どのタイミングでエラーが返されるかの情報を利用して，何桁目でエラーが生じたかの情報を得ることができる．この情報を用いることにより，5,000 回よりもずっと少ない回数で暗証暗号を当てることができる．まず，1 桁目が 0 であると予想し，ここでエラーが出れば，0 ではないことが確定する．次の試行では 1 桁目を 1 としてみる．エラーが生じなければ，1 桁目の正しい値が確定する．次いで 2 桁目の入力を，全ての桁を確定していく．この単純な方法により，高々 37 回の試行回数で暗証番号の復元に成功する．間違いが生じたタイミングでエラーを出力しなくても，何桁目でエラーが生じたかにより，エラーを出力するまでの時間がごく僅かに異なる場合も同様の攻撃が可能である．

ここで示した例は極端なものであり，このような ATM は当然のことながら存在しない．ここで示した攻撃は，秘密の情報を推測する際に，正解であるか不正解であるかの条件分岐により挙動が異なり，そのことを外部から観測可能であることを原因として成立する．つまり，安全性を担保するためには，秘密の情報に依存した情報を一切，漏らさないことが必要となる．

この種の攻撃を防ぐ技術として，「定数時間実装」が知られている．どのような入力（この場合は，暗証番号などの秘密情報）に対しても，計算時間が一定となるというものであり，（少なくとも計算時間からは，）何も情報が漏れず，安全性が担保されるというものである．

これまで説明してきたように暗号技術の安全性を考える場合には，様々な状況下での安全性を考慮する必要がある．安全性評価においては，使える道具は何でも使ってよい．これ以降では，RSA 暗号の安全性評価において，単に素因数分解の困難さを評価するという教科書的な評価だけでなく，攻撃者にとって有利な状況下での RSA 暗号の安全性について考える．**4.** 以降では，これまでに筆者が行ってきた誤り訂正符号，格子理論，量子アルゴリズムといった数理的な手法を用いた解析を紹介する．

具体的には，

1. RSA 秘密鍵の各ビットが漏えいする場合の安全性評価（**4.**）
2. RSA 秘密鍵 d が小さいときの攻撃（**5.**）
3. 量子計算機を用いた場合の RSA 暗号の安全性評価（**6.**）

を詳しく説明する．

4 RSA 暗号の秘密鍵が誤り付きで漏えいする場合の攻撃

RSA 暗号の安全性は巨大な合成数の素因数分解の困難さに基づいているが，これは単に「素因数分解を経由して RSA 暗号を破ることは難しい」ということにすぎない．攻撃者にとって有利なヒントが与えられた場合に，どの程度安全性が低下するかの評価が重要となる．ここでは，秘密鍵の部分情報が得られたときに，鍵全体を復元する攻撃を考える．

何らかの物理的な攻撃により，RSA 暗号の秘密鍵 (p,q,d,d_p,d_q) に関連した値が観測される状況を考える．サイドチャネル情報として，正しい秘密鍵の値に小さなエラーが乗った情報が得られる状況である．このような状況として，「Coldboot 攻撃」[5] などがある．ここでは，RSA 暗号の秘密鍵の各ビットが独立に漏えいする状況を想定する．まず，離散的な漏えいについて考える．これまでの研究では，秘密鍵の各ビットが独立に，

- 確率 δ で消失するモデル [6]
- 確率 ϵ でビット反転するモデル
- 非対称な確率でビット反転するモデル
- 確率 δ で消失しかつ確率 ϵ でビット反転するモデル [7]

などが検討されている．

より現実的な環境下での物理的な攻撃として，連続量のデータが得られるという状況も考えられている．観測値 y が，対応する正しい秘密鍵ビットに依存した確率分布に従うというモデルが提案されている [8], [9]．具体的には，秘密鍵のビットが 0 のときには確率密度関数 $f_0(y)$，ビットが 1 のときには確率密度関数 $f_1(y)$ に従って，実数値が観測される状況を考えている．このような誤り付きの連続量のデータから，

(1)RSA 暗号の秘密鍵を復元する効率的なアルゴリズムを提案すること，

(2) 攻撃に成功するための条件を導出すること，

が研究のゴールとなる．攻撃の成功条件の記述として，文献 (8)，(9) では，微分エントロピー（Differential Entropy）とカルバック・ライブラー情報量（KL-divergence）などの連続量に関する情報量を測る概念を用いている．ここで，カルバック・ライブラー情報量は二つの確率密度関数の間の距離とみなすことができる量である．

文献 (8)，(9) の成果を簡単に説明する．f_0 と f_1 に関する事前知識の程度により，四つのタイプに分類して，それぞれに対して有効なアルゴリズムを提案している．これらのタイプに共通するフレームワークとして，Tree-Based Approach [6] を用いている．このフレームワー

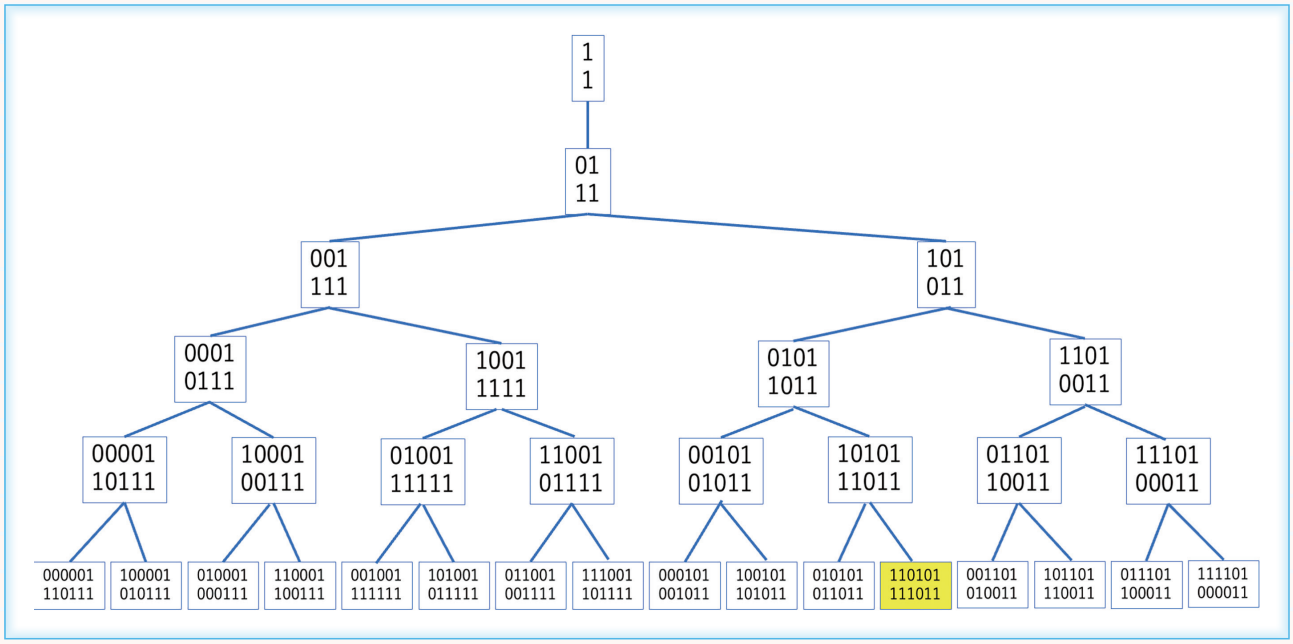


図４　Tree-Based Approach
(N=3127(=53×59)であり，pとqの情報だけを用いている．黄色い葉が正しいp (=110101) とq (=111011) に対応している．)

クでは，秘密鍵を
slice(i) := ($p[i]$), $q[i]$, $d[i+\tau(k)]$, $d_p[i+\tau(k_p)]$, $d_q[i+\tau(k_q)]$)
のように表現している．秘密鍵の slice($i-1$) までの部分情報を知っていると仮定すると，秘密鍵の各ビット間には五つの変数に対して四つの制約条件が存在するため，これにより二つの候補ビット列が生成される．Tree-Based Approach では，二分木（図４）を用いて，slice(i) を表現することになる．公開鍵が定まると二分木全体が一意に定まり，一つの葉が正しい秘密鍵に対応する．（図４中で，黄色くマークされている葉．）ここで，葉の総数は$2^{n/2}$となる．鍵復元アルゴリズムでは，適切なルールを設定した上で，観測された系列と木により管理される候補系列に基づき，それぞれのノードを捨てるか残すかを決めている．

具体的なアルゴリズムは，Expansion（拡張）and Pruning（枝刈り）戦略を用いる．各フェーズで，L個のノードそれぞれに対して2^t個のノードを生成し、$L*2^t$個のノードを得る Expansion フェーズと，その中からスコア関数の上位L個を残す Pruning フェーズを$n/2t$回繰り返すことになる．

このアルゴリズムにおいて，スコア関数をどのように設定するかが重要である．１個の観測系列yと多くの候補系列x_iに対して，argmax Score (x_i,y) を計算することで，最ももっともらしい秘密鍵を推定する．

以下，四つの状況のうち三つについて簡単に整理する．

状況１：f_0と f_1が完全に分かっている場合

スコア関数として対数ゆう度比を用いる．このスコア関数がうまく機能する直感的な理由は，正しくない候補xに対するスコアの期待値が０であるのに対し，正しい候補xに対するスコアの期待値が正であるためである．

状況２：f_0とf_1の近似が分かっている場合

現実には正確な漏えい分布が分からないという問題に対処するため，まず分布f_0とf_1を推定し，推定した分布を用いてスコア関数を設定することにする．成功条件は，正確な分布を知っている場合と比較して，推定のずれによる情報損が加わる形（KL Divergence により評価される）で悪化することになる．

状況４：f_0とf_1に関する情報が何もない場合

漏えい分布の知識を一切使用せずに計算可能な関数をスコア関数として用いる．これは，共通鍵暗号に対する差分電力解析で用いられる技術と類似している．

5　秘密鍵が小さいときの攻撃

RSA 暗号では，秘密鍵dを小さく設定することは実装上意味を持つ．復号処理は，$C^d \bmod N$で行われるため，dが小さければ，復号時間を短くすることが可能である．

RSA 暗号の秘密鍵が非常に小さいときには，容易に解読可能である．例えば，dが 1,000 以下と設定されている場合には，高々 1,000 回の試行で解読が可能である．複数の暗号文が得られているときに，その復号結果が常に意味を持つ場合，正しい鍵であると判定できる．また，正しいdが与えられたときに，ed-1は，(p-1)(q-1)の倍数であることを利用して，Nの素因数分解が可能である．

逆に，秘密鍵 d が非常に大きい場合，公開鍵のみから d を求めることは素因数分解と同じ計算量が必要であることが知られている．そのため，秘密鍵 d の大きさが，その間に属している場合，つまり，d がそれほど小さくなく，大きくもない場合の安全性に興味がある．

秘密鍵 d が $N^{0.25}$ よりも小さいときには，連分数展開を用いることにより，素因数分解できることが 1990 年に Wiener により提案されている．この d の限界は，1999 年に Boneh と Durfee により，$N^{0.292}$ にまで拡張されている[10]．この攻撃は，格子理論に基づくアルゴリズムであり，多項式時間で動作する格子簡約アルゴリズム，LLL (Lenstra–Lenstra–Lovasz) アルゴリズムをサブルーチンとして用いている．

格子理論に関する問題として最も代表的な問題は，最短ベクトル問題である．この問題は，m 本の線形独立なベクトルが与えられたときに，このベクトルの整数結合のうちで非自明な最も短いベクトルを求める問題である．(0 ベクトルが自明な最短ベクトルである．) この問題は，次元が大きい場合には，現実的な計算時間で解くことができないことが知られている．最短ベクトルを求めることが困難であることを利用して，暗号の構成に用いられている．その一方で，最短ではないが，そこそこ短いベクトルを求めることは現実的な時間で可能である．暗号の解読問題を，そこそこ短いベクトルを求めることに帰着することによって一連の攻撃は設計されている．この分野全般の解説として，文献 (11) などがある．

解読問題（RSA 暗号の秘密鍵 p と q を求めること）は，Small Inverse Problem として定式化されている．RSA 暗号の鍵設定 $ed \equiv 1 (\mathrm{mod}\,(p-1)(q-1))$ は，ある自然数 k が存在して，$ed-1=k(p-1)(q-1)$ と記述できる．更に，$k(N+1-(p+q))+1=0\,(\mathrm{mod}\,e)$ と変形可能である．$f(x,y)=x(A+y)+1$ とおくと，$f(x,y)=0\,(\mathrm{mod}\,e)$ の解の一つは $(x,y)=(k,-(p+q))$ となる．細かい係数を無視すると，$k<e^{\delta}, p+q<N^{1/2}\approx e^{\alpha}$ の範囲で解 $(k, -(p+q))$ を求める問題と考えることができる．$p+q$ を求めることができれば，$pq=N$ であるため，二次方程式 $x^2-(p+q)x+N=0$ と解くことによって，p と q を求めること（素因数分解）が可能である．

多項式時間で Small Inverse Problem の解を求めることができるための解の上限を導出することにより，解読できるための条件を定めることが可能である．$\alpha=1/2$ の場合，$\delta<0.292$ のときに多項式時間で解けることが示されている．

RSA 暗号における低指数べき攻撃の限界 $d<N^{0.292}$ は，提案当時は，0.292という値は不自然であり，限界が更新されることが予想されていた．しかし，その提案から 25 年以上たつが，いまだ更新されていないというのが現状である．筆者らは，0.292は，$1-1/\sqrt{2}$ であり，ある意味自然であるという考えに基づき，適切に限定したクラスにおいては，この0.292が最適であることを示している[12]．もちろん，このクラスに属さないアルゴリズムが，この限界を破る可能性は残されている．

6 量子計算機を用いた場合の RSA 暗号の安全性評価

古典計算機を用いた場合，素因数分解は多項式時間で解くことはできないと強く信じられている．RSA 暗号において，現在標準的に用いられている公開鍵 N は 2,048 bit に設定されているが，このアルゴリズムの計算量評価及び解読実験の結果を踏まえて設定されている．

量子 Turing 機械上では，RSA 暗号，だ円曲線暗号といった現在広く使われている公開鍵暗号方式を多項式時間で破ることが可能である．1994 年に提案されたショアのアルゴリズム[13] により素因数分解や離散対数問題の効率的な解法が可能になるためである．

量子計算の歴史としては，1985 年に Deutsch による量子 Turing 機械の導入をきっかけとし，1994 年に Simon のアルゴリズム（周期関数の位数発見），そして Shor のアルゴリズムが提案されている．更に，1996 年には Grover のアルゴリズム（探索問題）が提案されている．このアルゴリズムは，共通鍵暗号の秘密鍵探索やハッシュ関数の衝突発見に利用される．Shor のアルゴリズムは理論的な拡張が進み，可換群上での隠れ部分群問題（素因数分解や離散対数問題を含むより広いクラスの問題）を解くことが可能であることが知られている．

Shor の素因数分解アルゴリズムの戦略は，ターゲットとする合成数 N に対して，N と互いに素な自然数 a を選び，$a^r \bmod N=1$ となる自然数 r（位数）を求めることにある．r を求めることができれば，古典的な計算により（一定の確率で）N の因数を多項式時間で見つけることが可能である．Shor のアルゴリズム全体の計算時間のオーダは $\log N$ の多項式であることが知られているが，現代暗号に対する実際の影響を考えると，素因数分解に要する詳細なリソース評価が必要である．

素因数分解回路（図 5）の主要なステップは以下のとおりである．

Step 1: 初期状態 $|0\rangle^m \otimes |1\rangle$ を準備する．ここで $m=2\log N+1$ と設定する．

Step 2: 第 1 レジスタ（先頭 m-qubit）にアダマール変換を施し，均一な重ね合わせ状態を作る．

Step 3: べき乗剰余を適用し，重ね合わせの各状態に対して $|x\rangle\ |a^x \bmod N\rangle$ と変換する．

Step 4: 第 1 レジスタに逆量子フーリエ変換を作用さ

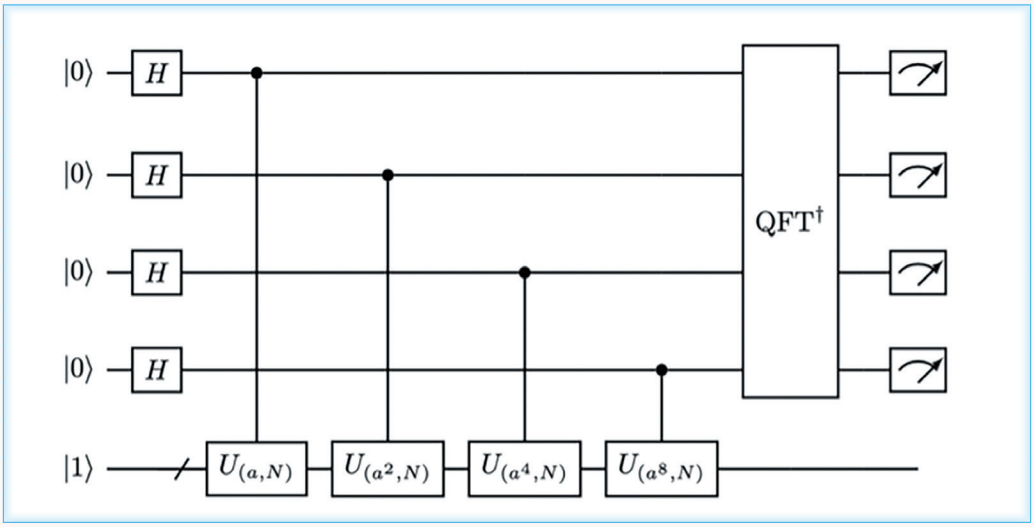


図５　Shor の素因数分解アルゴリズム

せる．

Step 5: 第 1 レジスタを観測する．

量子計算パートはここで終了である．量子計算回路構成において，$|y\rangle \rightarrow |by \bmod N\rangle$という，量子的に格納されている$y$という値を，古典的に定められた$b$倍を行い，同じく古典的に定まっている値$N$で割り算した余りを求める演算の実装が重要となる．

次は古典計算パートである．観測により得られた値を連分数展開することにより，rを復元し，このrの値を用いて素因数分解を行う．

前述のように，Shor アルゴリズムは量子パートと古典パートに分かれる．量子パートの核となるべき乗剰余演算は，更に幾つかの小さい演算に分割され，標準的には，加算回路にまで分解される．

加算回路の構成法として，様々な回路構成が知られており，古典計算の足し算を可逆化したもの，一般化 Toffoli ゲートを用いたもの，量子フーリエ変換を用いた量子加算などの構成法が知られている．どの加算を用いるかによって，量子ビット数，回路サイズ，回路深さなどの量子的なリソースが異なる．全てに優れた方法は存在せず，少ない量子ビット数で動作する回路構成は，回路サイズ，回路深さが大きくなる傾向がある．

文献 (14) では，量子誤りが一切ないという理想的な状況下でのリソース評価が与えられている．2,048 bit の合成数の素因数分解を行うには，4,000〜6,000 程度の量子ビット，10^{11}〜10^{13} 程度の量子ゲートが必要である．その後の研究の進展により，量子誤りがある状況下でのリソースの評価や，より効率的な量子回路の構成が提案されている．

量子計算機に対しても安全な耐量子計算機暗号の研究・開発が世界中で進んでいる．例えば，米国 NIST では，標準化プロジェクトが進んでおり，既に三つの標準化文書が公開されている．日本でも，耐量子計算機暗号への移行に向けて準備が進められている．移行を円滑に進めるためにも，RSA 暗号など現在利用されている方式に対する安全性をできるだけ正確に理解することが必要となる．

7　まとめ

本稿では，主に RSA 暗号に対する様々な安全性評価について紹介した．特に，教科書的ではない攻撃手法について詳しく説明を行った．**5.**，**6.** における発展した内容を理解したい場合には，文献（15）などの書籍を参照されたい．本稿により，暗号の安全性評価の重要性について一層理解が深まることを期待する．

■ 文献

(1) 情報処理推進機構，“情報セキュリティ 10 大脅威 .” https://www.ipa.go.jp/security/10threats/index.html
(2) CRYPTREC 暗号技術評価委員会報告 . https://www.cryptrec.go.jp/eval_cmte.html
(3) R. Rivest, A. Shamir, and L. Adleman, “A method forobtaining digital signatures and public-key cryptosystems,” Commun. ACM, vol. 21, no.2, pp. 120–126, 1978.
(4) P. Kocher, “Differential power analysis,” Proc. CRYPTO99, pp. 388-397, Dec. 1999.
(5) J. A. Halderman, S. D. Schoen, N. Heninger, W. Clarkson, W. Paul, J. A. Calandrino, A. J. Feldman, J. Appelbaum, and E. W. Felten, “Lest we remember: cold-boot attacks on encryption keys,” Proc. 17th USENIX Security Symposium, 2008.
(6) N. Heninger and H. Shacham, “Reconstructing RSA private keys from random key bits,” Proc. CRYPTO2009, 2009.
(7) N. Kunihiro, N. Shinohara, and T. Izu, “Recovering RSA secret keys from noisy key bits

with erasures and errors," IEICE Trans. Fundamentals, vol. E97-A, no.6, pp. 1273-1284, June 2014.
(8) N. Kunihiro and J. Honda, "RSA meets DPA: recovering RSA secret keys from noisy analog data," Proc. CHES2014, LNCS 8731, pp. 261-278, June 2014.
(9) N. Kunihiro and Y. Takahashi, "Improved key recovery algorithms from noisy RSA secret keys with analog noise," Proc. CT-RSA2017, LNCS 10159, pp. 328-343, Jan. 2017.
(10) D. Boneh and G. Durfee, "Cryptanalysis of RSA with private key d less than $N^{0.292}$," IEEE Trans. Inf. Theory, vol. 46, no. 4, pp. 1339–1349, July 2000.
(11) 國廣　昇, "格子理論を用いた暗号解読の最近の研究動向," 信学 FR 誌, vol. 5, no. 1, pp. 42–55, July 2011.
(12) N. Kunihiro, N. Shinohara, and T. Izu, "A unified framework for small secret exponent attack on RSA," IEICE Trans. Fundamentals, vol. E97-A, no. 6, pp. 1285-1295, June 2014.
(13) P. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," Proc. 35th Annual Symposium on Foundations of Comput. Sci., pp. 124–134, Nov. 1994.
(14) N. Kunihiro, "Exact analyses of computational time for factoring in quantum computers," IEICE Trans. Fundamentals, vol. E88-A, no. 1, pp. 105-111, Jan. 2005.
(15) 國廣　昇, 安田雅哉, 水木敬明, 高安　敦, 高島克幸, 米山一樹, 大原一真, 江村恵太, 暗号の理論と技術 量子時代のセキュリティ理解のために, 國廣　昇（編）, 講談社, 東京, 2024.

國廣　昇（正員：フェロー）

筑波大システム情報系教授．1996 東大大学院工学系研究科修士課程了．同年，日本電信電話株式会社入社．2001 博士（工学）．電通大，東大を経て，現職．一貫して，暗号理論，特に公開鍵暗号の安全性評価に従事．編著書に，「暗号の理論と技術 量子時代のセキュリティ理解のために」（講談社，2024-05），「データサイエンスはじめの一歩」（講談社，2024-08）などがある．