情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

属性情報と履歴情報を用いた不正アクセスの分析

大久保研究室 修士2年
邦本 理夫
2021/2/20

アジェンダ

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

- 研究の背景と目的
- ログの分析
- 検知手法の評価
- 考察
- まとめと今後の課題

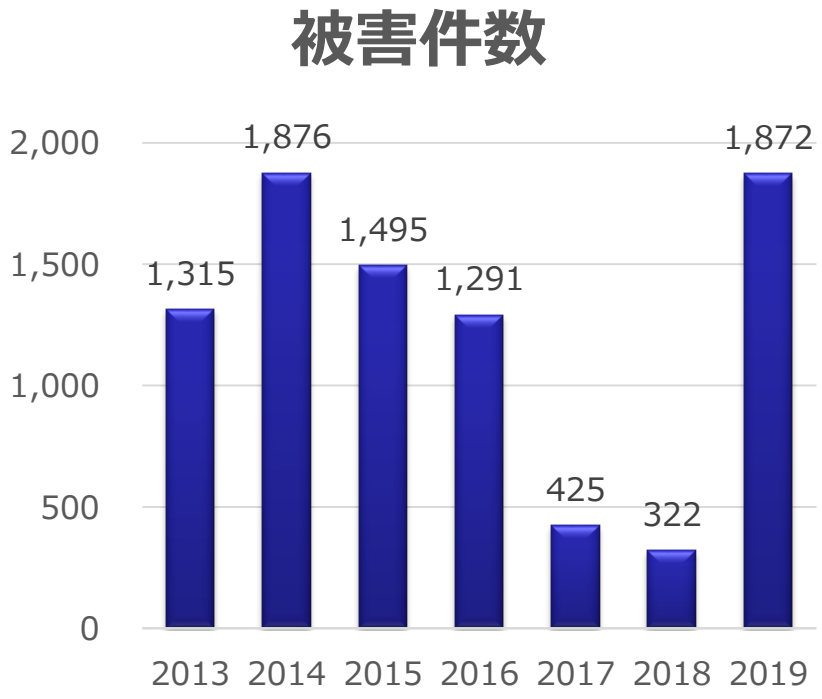
研究の背景と目的

- 成りすましによる不正アクセス（不正送金等）の被害が後を絶たない
  - フィッシング、マルウェア感染、リスト型攻撃等
    - ◆フィッシングサイト作成ツールが容易に入手可能
    - ◆不正取引用の口座やアカウント情報がSNSやダークウェブで売買されている
- こうした被害の完全な対策は困難
  - セキュリティ対策を行わない利用者は一定数存在（利用者のリテラシーに依存）
  - 各種の対策製品・サービスが存在するが、いずれも完全な対策は難しい

研究の背景と目的

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

金融機関をターゲットとした不正送金の被害件数、被害額は2018年まで減少傾向だったが、**2019年後半に急増**

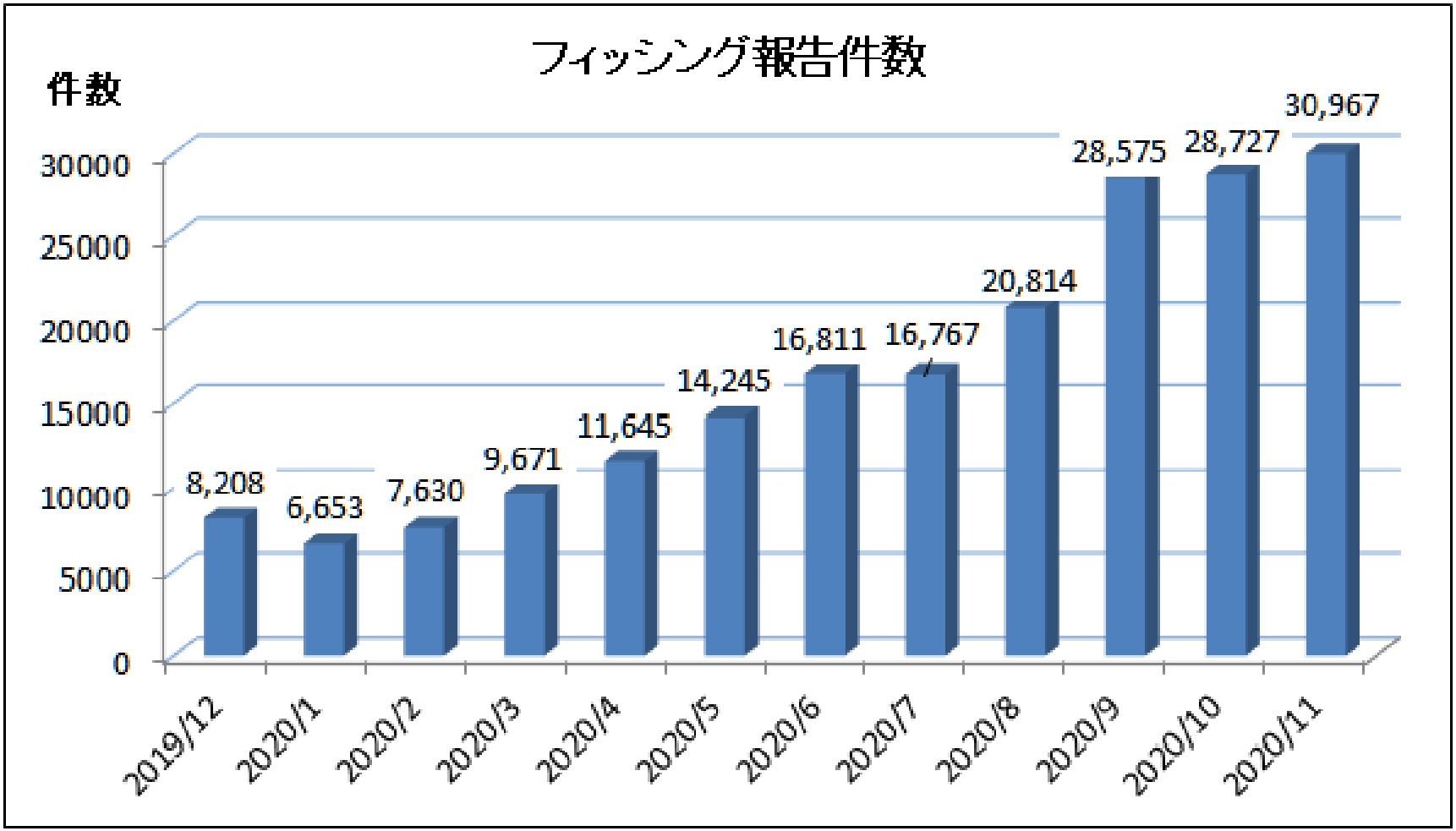


http://www.npa.go.jp/publications/statistics/cybersecurity/data/R01_cyber_jousei.pdf より

研究の背景と目的

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

フィッシング報告件数は近年急増



* https://www.antiphishing.jp/report/monthly/202012.html

研究の背景と目的

■ 着眼点

- 成りすまし犯人は最終的に盗んだ認証情報を使って対象のウェブシステムにログインし、送金等の不正行為を行う

↓

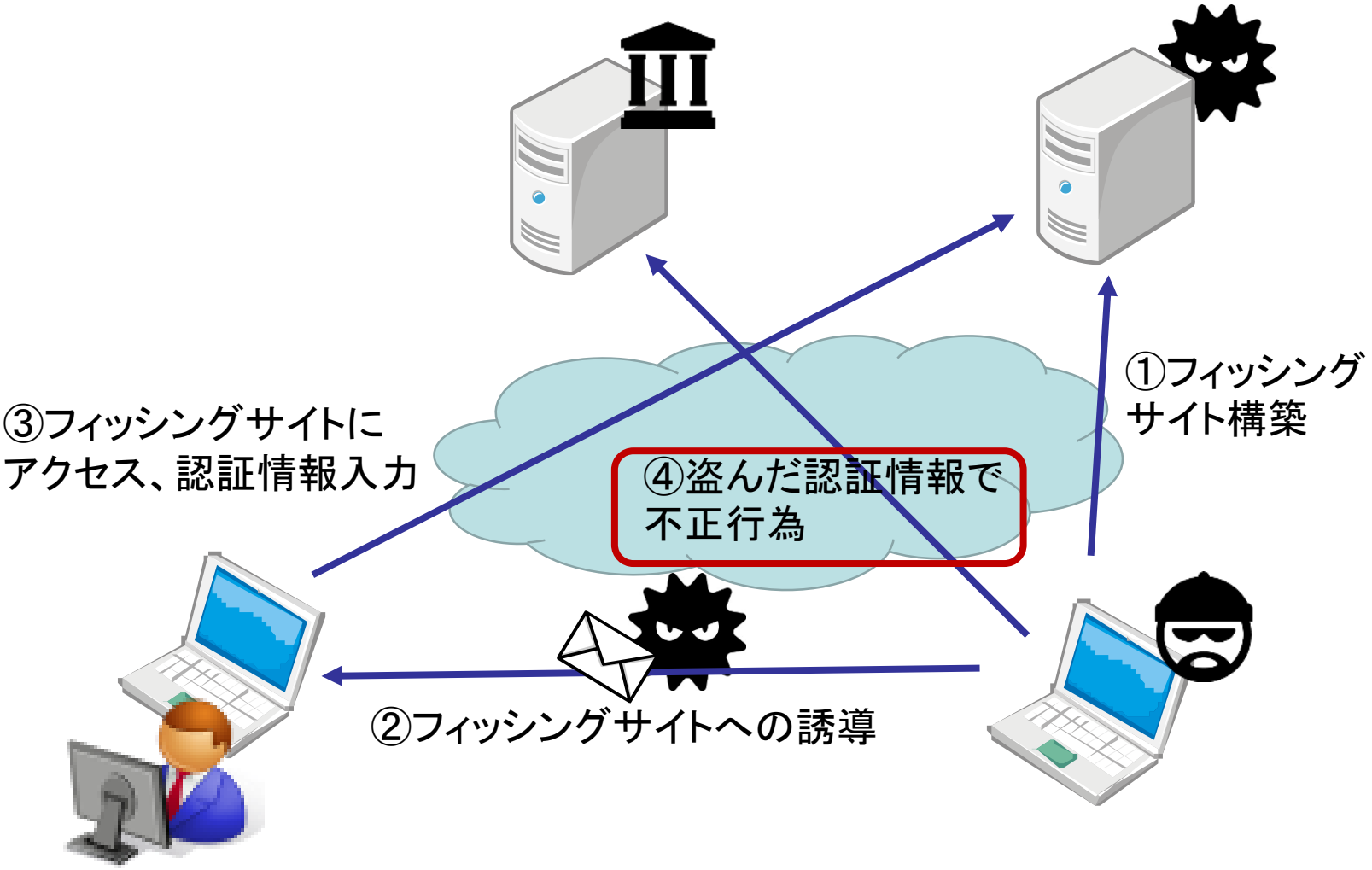
ウェブシステムにログが残る

↓

ログから正規利用者と犯人を区別できないか？

本研究の位置付け

■ フィッシングによる攻撃の流れと本研究の対象



本研究の位置付け

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

攻撃の段階	対策	課題
①フィッシングサイト構築	(1)コンテンツのクロールを検知して複製抑止 (2)類似ドメイン取得拒否 (3)サイトが立ち上がったことを検知して停止依頼	(1)作成ツールを特定できれば抑止できる場合もあるが限定的 (2)登録審査は業者依存、特に海外の場合困難 (3)検知までのタイムラグあり
②フィッシングサイトへの誘導	(1)クライアント側でメール/SMSフィルタ (2)プロバイダ、キャリア側でブロック	(1)クライアントおよびサービスの仕様依存 (2)ブロックの効果/範囲が限定的
③フィッシングサイトにアクセス、認証情報入力	(1)フィッシング検知ソフト (2)プロバイダ側でアクセス禁止	(1)インストールしない利用者 (2)IP、ドメインのブラックリスト追加までのタイムラグあり
④盗んだ認証情報で不正行為	(1)多要素認証 (2)不正アクセス検知	(1)全ユーザへの適用困難 (2)検知精度、ルールの運用

研究内容

■ ウェブシステムのログで収集できる情報の分析観点

- 端末の属性情報
  - ◆HTTPヘッダに含まれる情報（IP、UA、accept-language、etc…）
  - ◆JavaScriptで収集可能な端末情報
- 利用者の履歴情報
  - ◆普段使用しているOS/ブラウザ
- 利用者の振る舞い
  - ◆キーボード、マウスの挙動
  - ◆画面遷移のパターン、速度
  - ◆アクセス元情報（IPの位置情報など）の変化

分析対象のログの内容を踏まえ、
属性情報および履歴情報を使用した
不正アクセスの検知手法を検討

研究内容

■ 先行研究

- 属性情報
  - ◆Browser Fingerprint関連の研究
    →端末特定の精度が主眼で、不正検知への応用は少ない
  - ◆不正検知：Picasso（2016）
    →UserAgent偽装検知に特化
- 履歴情報
  - ◆現状見当たらず（実サービスのログが必要なためと推測）
- 振る舞い
  - ◆特定のシステムに特化した不正検知
    （BankSealer(2015)、FraudBuster(2018)など）
    →インターネットバンキングの振込額などの振る舞いによる検知

属性情報と履歴情報を組み合わせた不正検知の先行研究は発表されていない

ログの分析

■ 分析対象のログについて

- 業務で提供している不正検知サービスのログ
  （顧客名を出さないことを条件に使用可）
- アクセス元端末の属性情報（JavaScriptで取得可能なもの）を収集
- サービスを利用している管理者が、実際に不正のあった端末の属性情報をブラックリスト登録
- 分析対象の期間と件数：

期間	総件数	分析対象件数	不正判定数
2019/1/1~12/31	171,631,789	93,311,194	1,725

◆分析対象件数：外部サービスのAPIからのアクセスなどを除外した件数

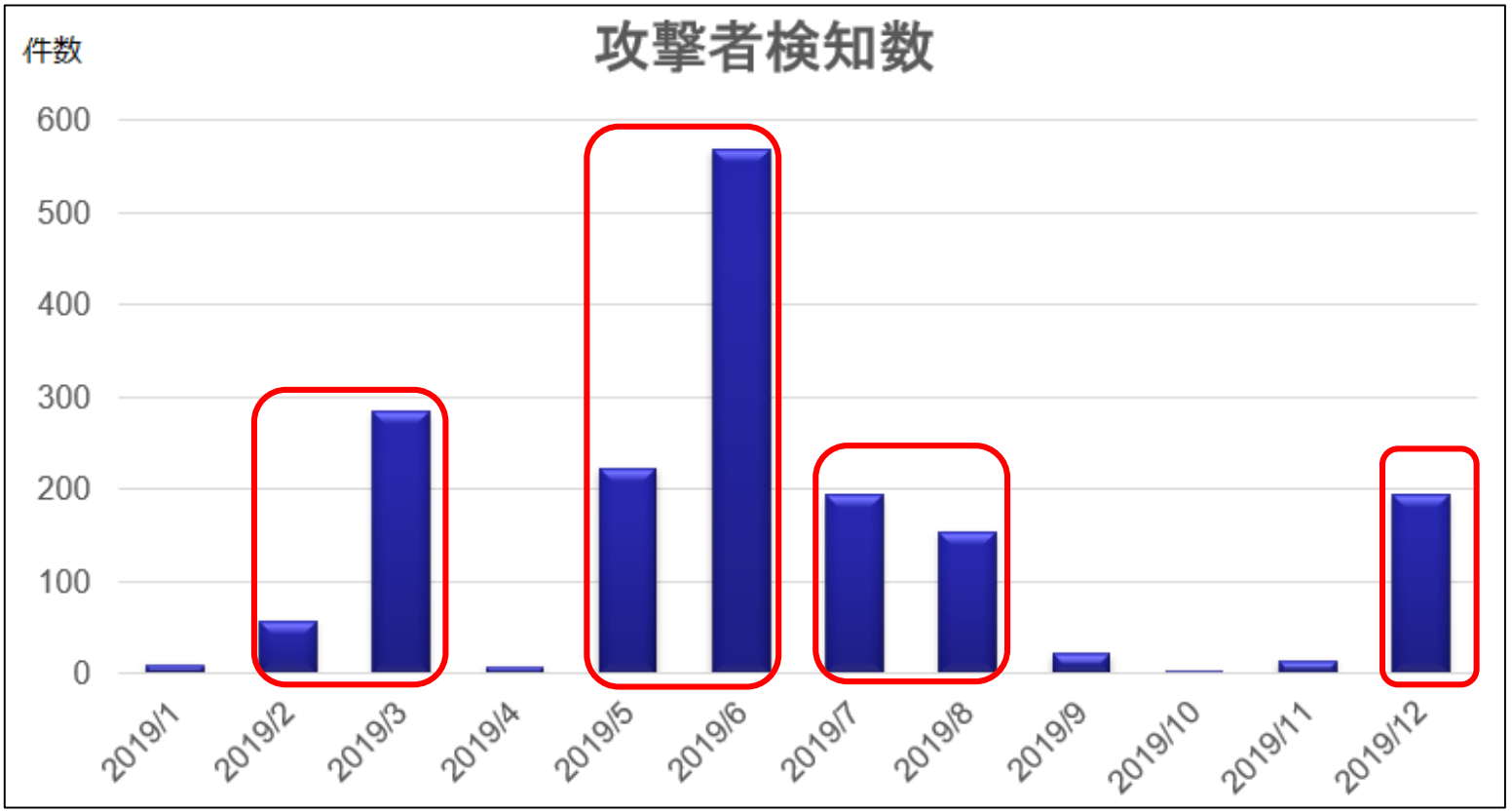
◆不正判定数：
サービスを利用しているシステムの管理者が不正と判定した端末からのアクセス（述べ件数）
※誤検知と判明しているアクセスについては分析前に件数を補正済み

ログの分析

- 属性情報の評価
  - 不正アクセス判定されたログの属性情報を分析する
  - 不正アクセスの検知条件として有効な組合せを検討する
- 履歴情報の評価
  - 顧客ID単位でログを分析し、端末情報の変化を分析する
  - 不正アクセスを受けたときのブラウザ/OSや属性情報の変化と、通常利用時の変化を比較する
  - 不正アクセスの検知条件として有効な変化パターンを検討する
- それぞれの条件について定量評価を行う

分析結果(属性情報)

■ 攻撃者アクセスの統計



- 月によって増減あり
- アクセス元環境の特徴が異なる4つのグループが存在

分析結果(属性情報)

■ 攻撃者の属性情報の特徴

- 高性能GPUからのアクセス（2月～3月）
- 仮想マシンからのアクセス（5月～6月）
- 国外IP、クラウド環境からのアクセス（7月～8月）
- iPhoneからのアクセス（12月）

・特徴的に表れた端末の属性情報（国外IPを除く）について対象期間と2019年全体のそれぞれで定量評価を実施

■ 正規利用者の属性情報との比較

- 言語設定、タイムゾーンなどが異なる傾向

・正規利用者と異なる傾向を持つ言語とタイムゾーンを用いて定量評価を実施

分析結果(属性情報)

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

■ 高性能GPU利用

- WebGLRenderer = ANGLE (NVIDIA GeForce GTX750 Ti Direct3D11 vs 5 0 ps 5 0)

全期間を対象
- 検知率（TPR）　：　**12%**　（207 / 1,725）
- 誤検知率（FPR）　：　**0.1%**　（96,698 / 93,309,404）

3月のみ対象
- 検知率　：　**66.9%**（190 / 284）
- 誤検知率　：　**0.2%**　（9,343 / 4,818,202）

特定の期間を対象にした場合約2/3の攻撃者を検知可能
誤検知も0.2%程度発生

※検知率　：攻撃者アクセス全体のうち指定の条件を満たすものの割合
　誤検知率：正規利用者アクセス全体のうち指定の条件を満たすものの割合

分析結果(属性情報)

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

■ 高性能GPU利用（条件組み合わせ）

- WebGLRenderer = ANGLE (NVIDIA GeForce GTX750 Ti Direct3D11 vs 5 0 ps 5 0)
  HardwareConcurrency >= 16
  language = zh-CN

  全期間を対象
  - 検知率　：　**10.8%**（187 / 1,725）
  - 誤検知率　：　**0.001%**（1,132 / 93,309,404）

  3月のみ対象
  - 検知率　：　**60.5%**（172 / 284）
  - 誤検知率　：　**0.003%**　（133 / 4,818,202）

検知率は微減したが誤検知を約1/100に低減

分析結果(履歴情報)

■ 履歴情報の分析における仮説

- 正規利用者が普段ウェブシステムを利用する環境は
  ある程度固定されているはず

- 攻撃者が攻撃対象の正規利用者のアカウント毎に
  利用環境を調査して同じ環境を用意する可能性は低い
  - ◆フィッシングなどバラマキ型の攻撃でそれをやるとコストがかかる
    (標的型攻撃などであればそこまでやる可能性あり？)
  - ◆最近はワンタイムパスワード突破のためリアルタイムで攻撃が
    行われており、環境を準備している時間は無い

**普段と異なる環境からのアクセスは、
第三者による不正アクセスの可能性が高いと推測**

分析結果(履歴情報)

■ 履歴分析の観点

① OSの差異

a. OSの種類が異なる
b. 攻撃者の方がバージョンが古い（バージョンダウン）
c. 攻撃者の方がバージョンが新しい（バージョンアップ）

② ブラウザの差異

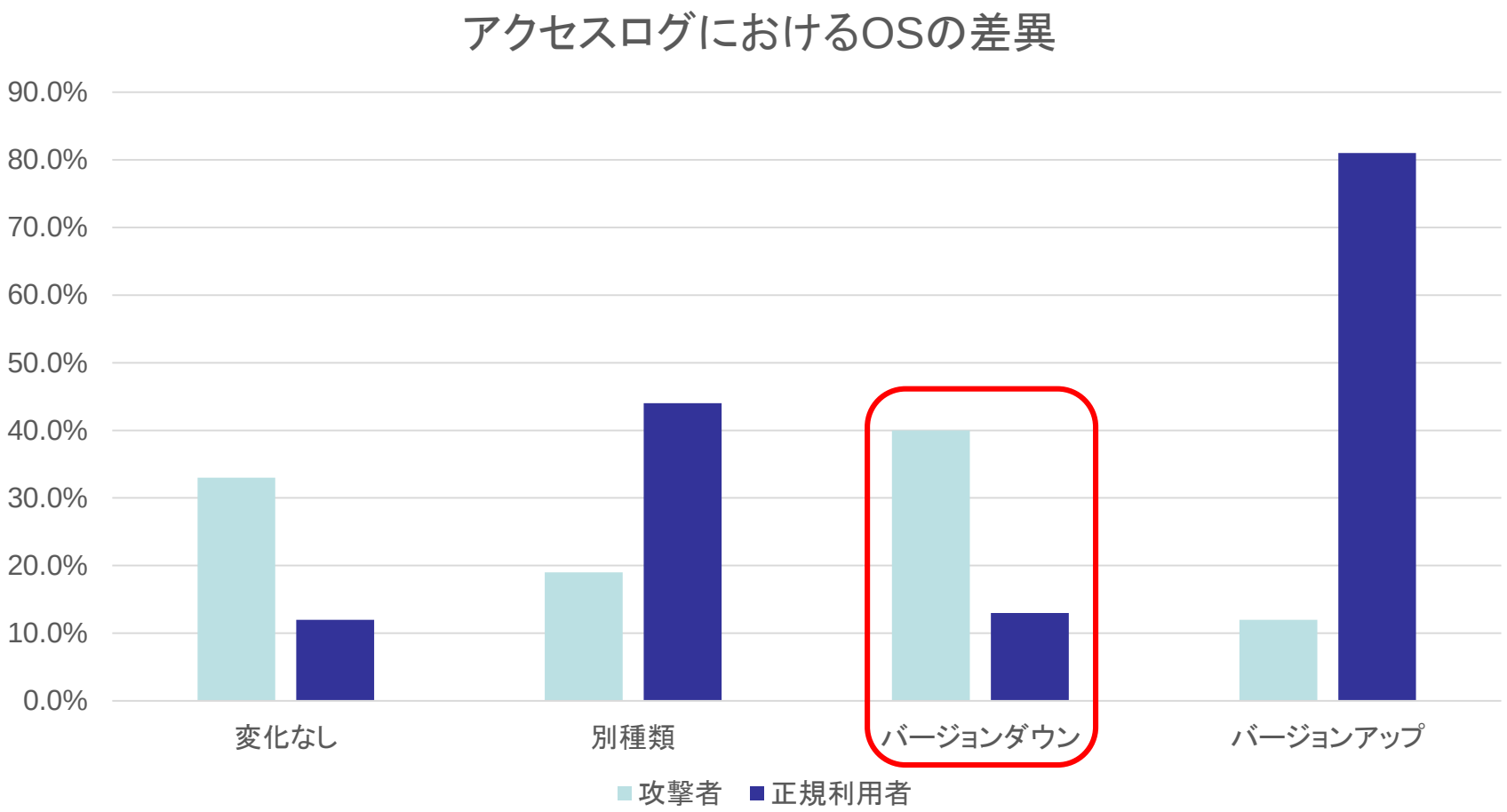
a. ブラウザの種類が異なる
b. 攻撃者の方がバージョンが古い（バージョンダウン）
c. 攻撃者の方がバージョンが新しい（バージョンアップ）

③ 言語・タイムゾーンの差異

a. 言語設定が異なる（完全一致）
b. 言語設定が異なる（第一言語が異なる）
c. タイムゾーンが異なる

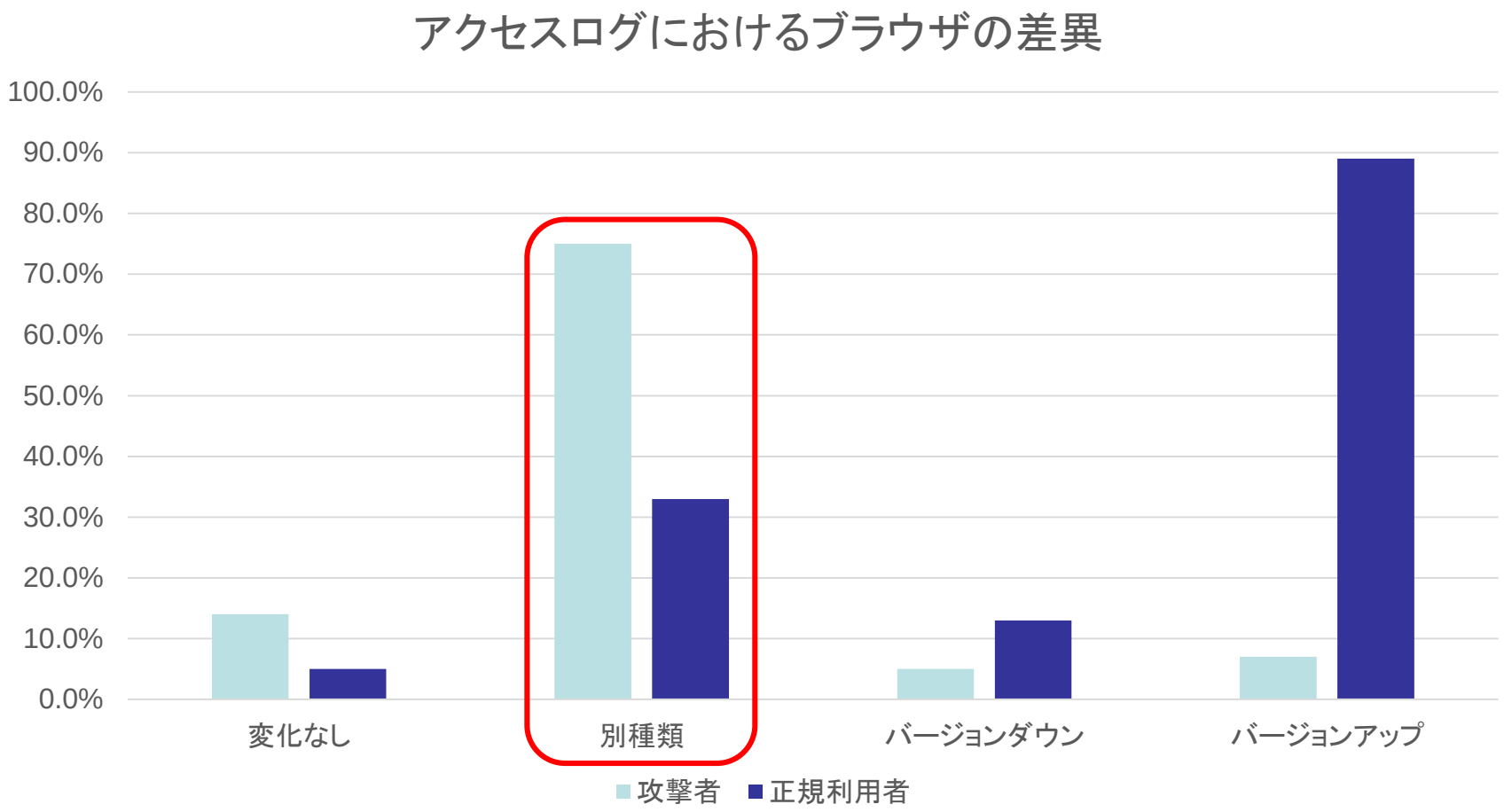
分析結果(履歴情報)

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY



分析結果(履歴情報)

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY



分析結果(履歴情報)

■ OS・ブラウザ変化の分析

- 攻撃者アクセス時、正規利用者と比較して
  OSがバージョンダウンしている場合と
  ブラウザが変化している場合が多い
  - OSバージョンダウン　：　攻撃者**40件** / 正規利用者13件
  - ブラウザ変化　：　攻撃者**75件** / 正規利用者33件

- 上記のAND条件（OSバージョンダウンかつ別ブラウザ）
  を取ると検知率/誤検知率ともに低減
  - 攻撃者　：　30件
  - 正規利用者　：　7件

分析結果(履歴情報)

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

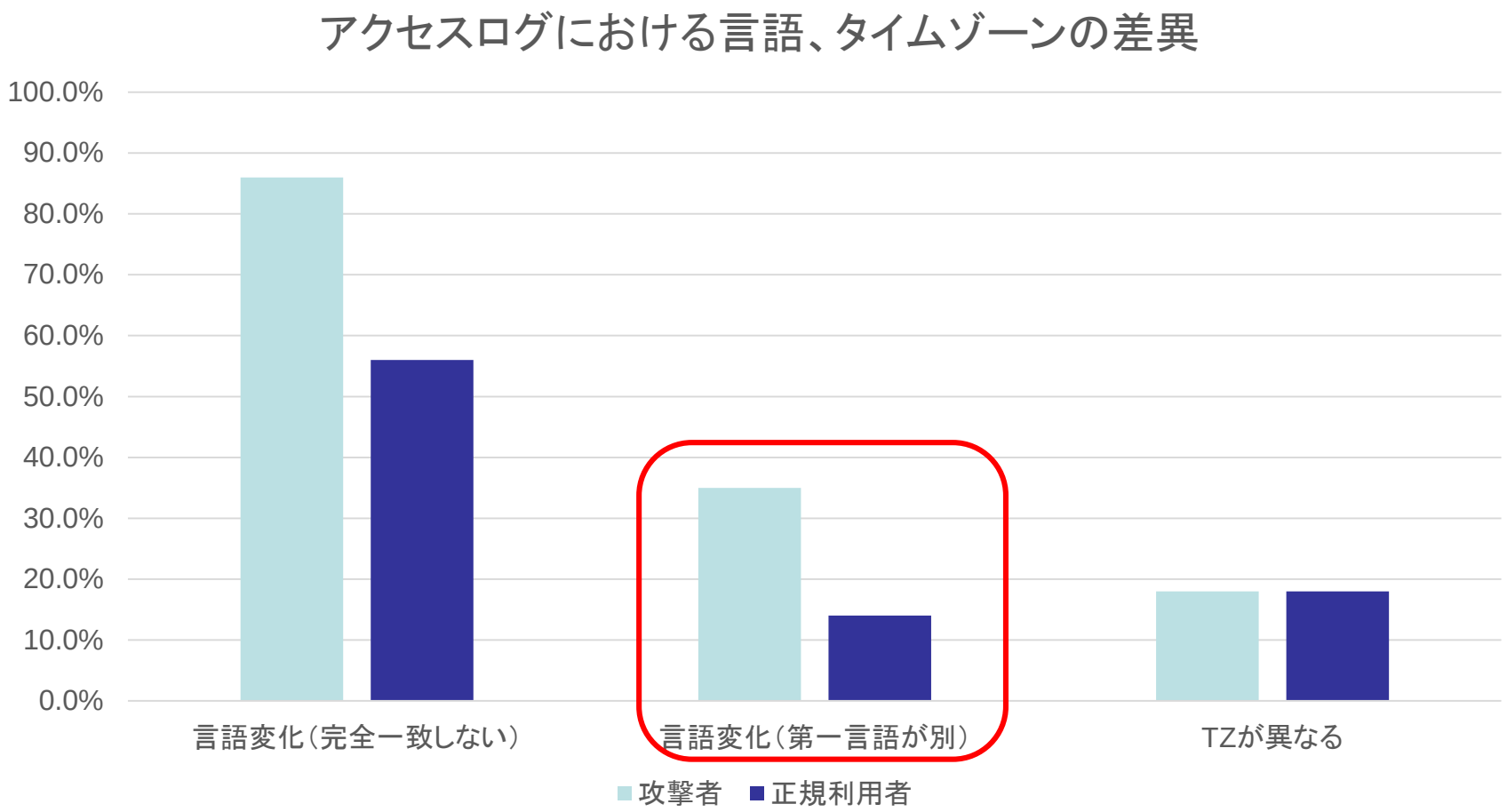
■ 言語・タイムゾーンの変化分析

● 言語・タイムゾーンへの着目理由

- フィンガープリントの大半は ”端末特定” に用いられる
- 言語環境は利用者に紐づくもので、端末やブラウザを変えても変化しにくいはず
  ※細かい値が変わることはあると想定し、完全一致と第一言語の差異をそれぞれ調査
  （例）ja,en-US;q=0.7,en;q=0.3 と ja-JPは
  完全一致ではないが、第一言語の差異は無しと判断
- タイムゾーンは出張・旅行等での変化は有り得るがそれほど多くないと想定

分析結果(履歴情報)

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY



分析結果(履歴情報)

情報セキュリティ大学院大学
INSTITUTE of INFORMATION SECURITY

■ 言語・タイムゾーン変化の分析

- 攻撃者アクセス時の第一言語変化（全35件）
  - ja/ja-JP → zh-CN : **30件**
  - ja/ja-JP → en-US : 5件
- 正規利用者の第一言語変化（全14件）
  - ja/ja-JP → en-US : 11件
  - ja/ja-JP → en-AU : 2件
  - ko-KR → en-US : 1件

第一言語が日本語→中国語に変化した場合は攻撃の可能性が高い
（サンプルでは検知率30%、誤検知0）

検知手法の評価

- 2019年の攻撃者アクセスの分析結果から、攻撃者検知に有効と思われる下記の条件について2020年のログを用いて有効性を検証

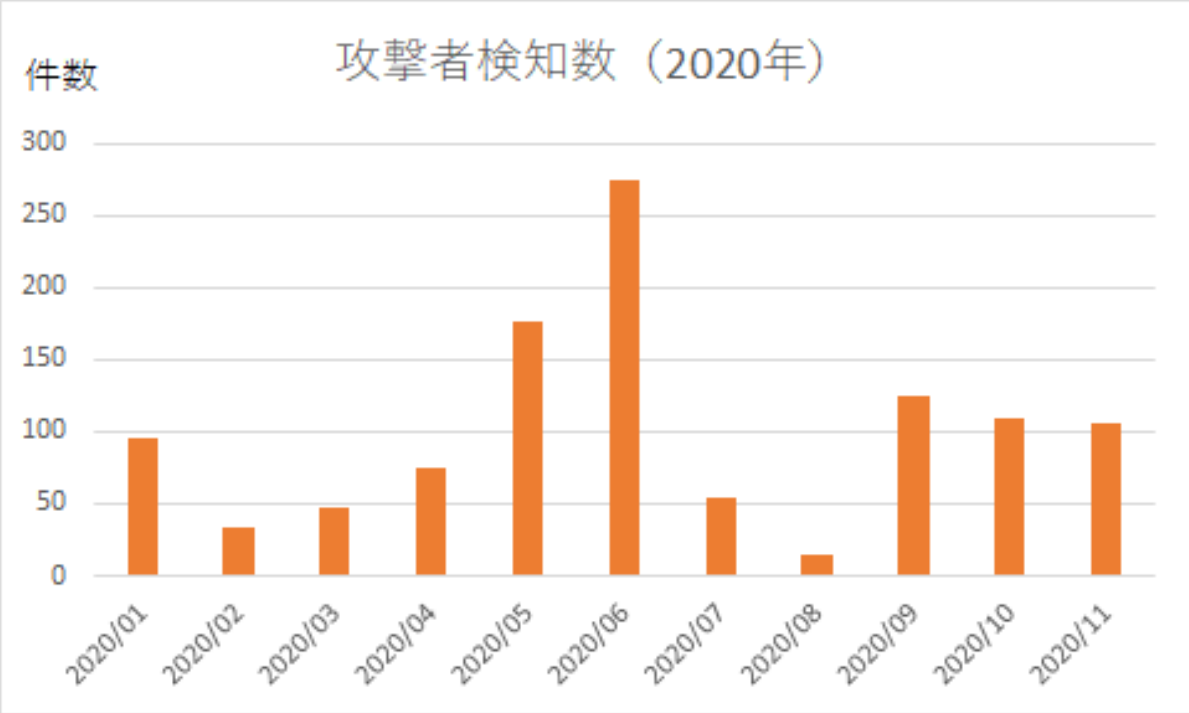
種別	検知条件
属性評価	高性能GPUからのアクセス
	仮想マシンからのアクセス
履歴評価	OSのバージョンダウン
	別ブラウザの利用
	第一言語の日本語から中国語への変化

検知手法の評価

■ 評価対象のログ

- 評価対象の期間と件数：

期間	総件数	分析対象件数	不正判定数
2020/1/1~11/30	158,989,711	127,100,947	1,111



検知手法の評価

■ 評価結果

- 攻撃者の端末属性情報の特徴が2019年と2020年で異なっており、属性評価による検知は有効性が低い
- 履歴評価は条件によって一定の有効性がある

種別	検知条件	有効性
属性評価	高性能GPUからのアクセス	低
	仮想マシンからのアクセス	低
履歴評価	OSのバージョンダウン	低
	別ブラウザの利用	中
	第一言語の日本語から中国語への変化	高

検知手法の評価

■ 属性情報による評価

- 高性能GPUからの攻撃者アクセスは、2020年のログには存在せず
  - ◆ 2019年と別種のGPUからのアクセスは数件存在したが単発のアクセスのみで有効性が低い
- 仮想マシンからのアクセスも存在せず

■ 履歴情報による評価

- OSバージョンダウンの割合は正規利用者とほぼ同一
- 別ブラウザ利用の割合は正規利用者と比較して高い
- 第一言語が変化した場合について、**19件中18件**が日本語→中国語の変化であり、検知条件として有効

検知手法の評価

■ 攻撃者環境の傾向変化分析

- 言語環境、タイムゾーンが2019年と2020年で大きく変化
  - 言語環境

設定値	2019年	2020年
ja / ja-JP	1023	451
zh-CN	517	167
en / en-US	61	22
vi / vi-VN	27	**471**

  - タイムゾーン

設定値	2019年	2020年
-540	1304	754
-480	373	126
-420	8	**220**
420	0	10
480	0	1

検知手法の評価

■ 攻撃者環境の傾向変化分析

- モバイル端末からのアクセスが大きく増加
  - モバイルからのアクセス数

	2019年	2020年
攻撃者アクセス総数	1725	1111
iPhoneからのアクセス	207	565
Androidからのアクセス	1	260

- 2020年の攻撃者によるiPhoneからのアクセスのうち、8割以上（471件）はベトナム語環境からのアクセス

考察

- 属性評価および履歴評価による不正検知は、検知精度と長期的な有効性が一長一短
  - 属性情報は検知精度が高いが変化しやすい
  - 履歴評価は精度に劣るが長期的に有効
- 属性評価は、攻撃者に特有の属性が特定できれば高精度で効果を発揮
  - 攻撃の兆候を素早く掴めればその後の被害を低減可能（企業間の情報連携も可能）
  - 検知条件の見直しが随時必要
- 履歴評価は、攻撃者の初回アクセスも検知できる可能性あり
  - 履歴は企業間での共有が困難

考察

■ 将来の環境変化への対応

- 本研究の検知手法は、利用者の環境の多様性に基づいている
  - ◆PC・ブラウザの属性情報
  - ◆履歴（普段使っている環境）が個々人で異なる
- 利用ユーザは将来的にスマホが増加すると想定
  - ◆ブラウザベースであれば現状の提案手法である程度対応できるが、精度が低下する懸念有り
  - ◆ネイティブアプリの場合はSDK等による情報収集の仕組みが必要
    ※アプリの場合は不正の手口も変化すると想定され、別の切り口での対応が必要となる可能性

モバイル端末への対応は今後の課題

考察

- 不正検知のエコシステム
  - 今回の提案手法で用いるパラメータはすべてJavaScriptで収集可能
  - 属性情報による検知は、情報を各企業で共有することで、被害を未然に防ぐ効果が期待できる
    - ◆金融ISACでは不審IPの共有などの取組みが既に行われており、その発展形
  - 完全な対策は困難でも、攻撃者に「すぐに検知される」「手間が掛かる」と思わせることが出来れば良い
    - ◆金銭目的の攻撃者は、割に合わない相手を狙わない

まとめ

- 実サービスのログの分析から、攻撃者アクセスの特徴を明らかにした
  - 月毎に攻撃者アクセスが増減し、時期によってアクセス元の特徴が異なる
  - 同一の属性情報を持つ端末からのアクセスが一定期間継続
  - 攻撃者と正規利用者の履歴で異なる傾向が見られる
  - 攻撃者の端末属性が2019年から2020年で変化している

- 不正検知手法の有効性評価を行い、以下の点を示した
  - 属性情報による検知は精度が高いが有効期間が限られる
  - 履歴情報による検知は精度に劣るがある程度長期的に有効

今後の課題

- 攻撃者と正規利用者の端末情報が類似していた場合の検知手法の検討
  - 利用者の振る舞いなどを用いた検知手法について今後有効性を検討

- 端末環境の変化への対応
  - ネイティブアプリで収集可能な属性情報と、不正検知への有効性評価、分析手法の検討
  - 利用環境の変化に伴う攻撃の手口とその対策の検討

ご清聴ありがとうございました