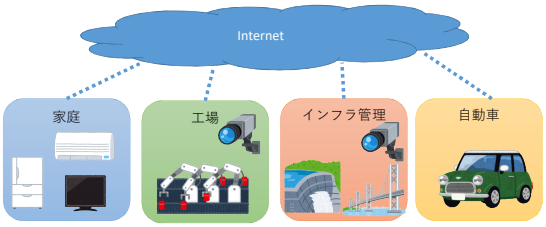
IoT時代に向けた攻撃検知技術

大阪大学
大下　裕一

IoT (Internet of Things)

様々なモノがインターネットに接続



IoT機器が増えるにつれて

IoT機器が攻撃者から狙われるように。。。

事例１：IoT機器を踏み台に
ウイルスに感染したネットワークカメラ等１０万台を超えるＩｏＴ機器を踏み台にネットワークインフラを攻撃[1]

事例２：停電
変電所システムにウイルスが感染し一部機能が停止[2]

セキュリティ面におけるIoT特有の性質１

脅威の影響範囲・影響度合いが大きい
- 攻撃をうけると機器単体ではなく、関連するIoTシステム、サービス全体へも影響を与える
- 自動車分野、医療分野等では、生命の危機さえ起こりうる
- IoT機器から個人の生活データ等の漏洩も起きうる

出典：[3] IoT 推進コンソーシアム “IoTセキュリティガイドライン”

セキュリティ面におけるIoT特有の性質２

IoT 機器のライフサイクルが長い
- 自動車のライフサイクル：12～13年
- 工場の制御機器：10～20年

↓

セキュリティ的に不十分になった機器が使い続けられる可能性が高い

出典：[3] IoT 推進コンソーシアム “IoTセキュリティガイドライン”

セキュリティ面におけるIoT特有の性質３

IoT 機器に対する監視が行き届きにくい
- IoT機器は、PCやスマートフォンと違い画面がない

↓

従来と違う挙動をしていても気づきにくい。

出典：[3] IoT 推進コンソーシアム “IoTセキュリティガイドライン”

セキュリティ面におけるIoT特有の性質４

IoT 機器側とネットワーク側の環境や特性の相互理解が不十分

- IoT機器側とネットワーク側の特性の理解が不十分

↓

- IoT機器で前提としている要件・性能をネットワーク側が満たさない可能性

出典：[3] IoT 推進コンソーシアム “IoTセキュリティガイドライン”

セキュリティ面におけるIoT特有の性質５

IoT 機器の機能・性能が限られていること

- マルウェア対策ソフトの導入が困難
- 機器よっては、暗号化すら困難なことも

出典：[3] IoT 推進コンソーシアム “IoTセキュリティガイドライン”

セキュリティ面におけるIoT特有の性質６

開発者が想定していなかった接続が行われる可能性がある

- これまでネットワーク接続されていなかった新たな機器がネットワーク接続されることにより、サービス開始時には想定していなかったリスクが生じることも

出典：[3] IoT 推進コンソーシアム “IoTセキュリティガイドライン”

IoTのリスクに備えるには

- 設計時にリスクに備えた設計
  - リスクを考慮した機器・サービス・ネットワークの設計

機器側の設計のみでリスクの回避は困難
- 機器のライフサイクルが長い
- 機器の機能・性能が限られている

- 攻撃の検知・遮断
- 早期発見

} 異常検知

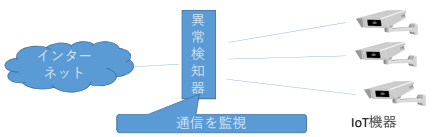
IoTへの異常検知

機器の機能・性能が限られている

↓

機器の外側で機器を監視

- IoT機器が出す通信を監視
- IoT機器宛の通信を監視
- IoT機器から情報を収集して監視



IoT機器で発生しうる異常

- 機器自体が異常
  - マルウェア感染
  - 不正侵入

↓

  - 機器の本来機能の停止
  - 機器で本来動作しない機能の開始
  - 機器を踏み台にした新たな攻撃
- 機器自体は正常に動作
  - 権限のない者からの機器の操作

我々の研究事例の紹介

事例１：
重要インフラにおけるIoTシステムの動作監視・解析技術[4,5]
　戦略的イノベーション創造プログラム（ＳＩＰ）
　「重要インフラ等におけるサイバーセキュリティの確保」における取組

事例２：
ホームネットワークにおける不正操作検知[6]
　大阪大学　三菱電機サイバーセキュリティ協働研究所における取組

重要インフラにおけるIoT機器の動作監視・解析技術

研究目標

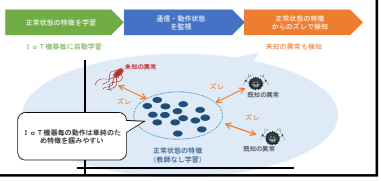
多種多様なIoT機器に自動適応して動作を監視・解析し、不正な動作を検知

- 未知なものを含む多様なIoT機器に対応したIoTシステムの動作監視・解析
  - 多種のIoT機器が次々と導入されたり、IoT機器の用途が多様化したとしても、自動的に適応して動作を監視・解析
- 膨大なIoT機器により構成されたIoTシステムの動作監視・解析
  - 多様な方法で接続される膨大なIoT機器を自動検出し、効率的に導入、監視・解析
  - 複数IoTシステムからの結果を安全に集約・解析することにより異常原因を特定

動作監視・異常検知の考え方

各IoT機器は限られた動作のみを行う
＝正常な挙動は分かりやすい

正常な状態を教師なし学習で学習
学習した正常状態からのずれで異常を検知



異常検知のみで十分か？

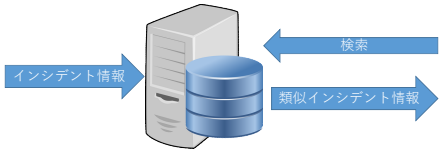
- 検知された異常により対応が異なる
  - IoT機器への侵入
    → 当該機器の取り外し、再設定　等
  - 外部から脆弱性のスキャンによる通信が到達
    → 当該機器に該当する脆弱性がなければ、対処不要　等

異常の検知のみでは、その異常に対応するためのヒントすら得られない

検知のみでなく分析も

インシデント情報データベース／解析システム

- これまでに対応したインシデントの情報を蓄積
- 新たなインシデントが発生した際には、蓄積されたインシデント情報から類似インシデント情報を検索



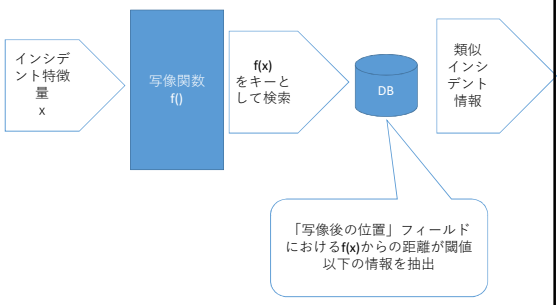
インシデント情報データベース／解析システム

以下のような検索が可能であること
- 検索クエリ：
  現在発生したインシデントの特徴量
  （インシデント発生時に 通信が行われていた宛先ポート番号や通信量等）
- 検索結果：
  現在発生したインシデントに類似したインシデントの詳細情報

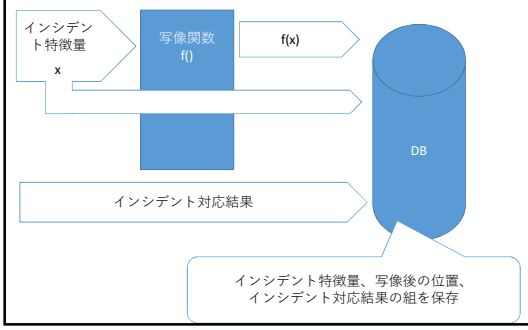
留意事項
- インシデントの特徴量が類似＝同種のインシデントではない
- どのような種類のインシデントが発生するかは、事前には分からない
  - 新種のインシデントの発生の可能性もあり

本研究のアプローチ
- インシデントの特徴量を写像する写像関数を導入
- 類似のインシデントを近い位置に写像することにより、過去の類似のインシデントの情報を取得

インシデント情報データベース／解析システム ～検索～



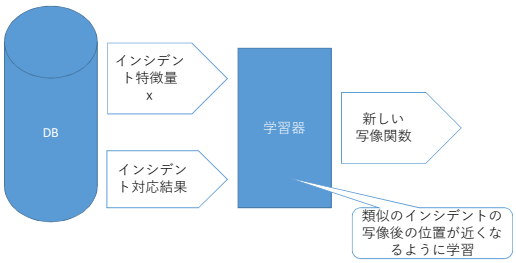
インシデント情報データベース／解析システム ～情報の蓄積～



写像関数の学習

- どのインシデント同士が類似であるかは、インシデント対応後に判明
  類似のインシデントと考えられる例
  - インシデント対応情報に含まれる、インシデントの種類に関する情報が同一
  - インシデント対応の手順が同じ
- インシデント情報を蓄積後に、その情報をもとに写像関数を逐次学習することが望ましい

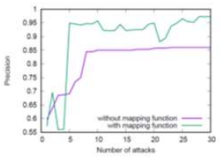
インシデント情報データベース／解析システム ～写像関数の学習～



※ 写像関数の更新後は、必要に応じてDBの写像後の位置も更新

試作・評価

- 評価環境
  - 写像関数をニューラルネットワークで構成・学習
  - 検索対象：フロー
    - 特徴量：継続時間、通信量、ポート番号、同時に発生した同じ宛先のフロー数等
    - 生成したフロー：SYN flood攻撃、Brute force攻撃　等
  - 評価指標
    - 新たに生成したフローの特徴量で検索
    - 得られた最も似ているフローが、同種のフローであった割合
- 結果
  - 写像関数を通すことにより、一定数の経験が蓄積されると、高い精度で分類が可能



ホームネットワークにおける
不正操作検知

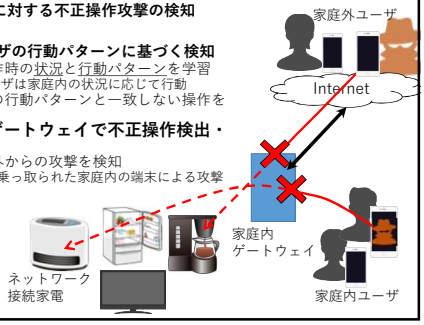
検知対象：IoT 機器の不正操作

- 第三者による操作トラヒックの送信
  - 機器の不正操作には重大な危険性
    - 例：ヒータの不正操作による火傷・火災
- 従来型の手法での検知が困難
  - パターンマッチングによる検知は困難
  - 正常な通信パターンとの比較による検知も困難
    - 正常な操作と同様の通信で不正操作
    - スマートフォンにマルウェアが侵入した場合などは、不正操作時の操作元のIPアドレスすら、普段と使用しているものと同一

研究目的とアプローチ

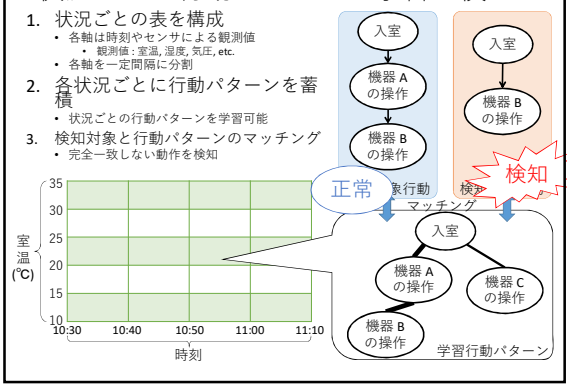
研究目的
- **ホーム IoT に対する不正操作攻撃の検知**

アプローチ
- **家庭内ユーザの行動パターンに基づく検知**
  - 機器操作時の状況と行動パターンを学習
    - ユーザは家庭内の状況に応じて行動
  - 各状況の行動パターンと一致しない操作を検知
- **家庭内のゲートウェイで不正操作検出・遮断**
  - 家庭内外からの攻撃を検知
    - 例：乗っ取られた家庭内の端末による攻撃



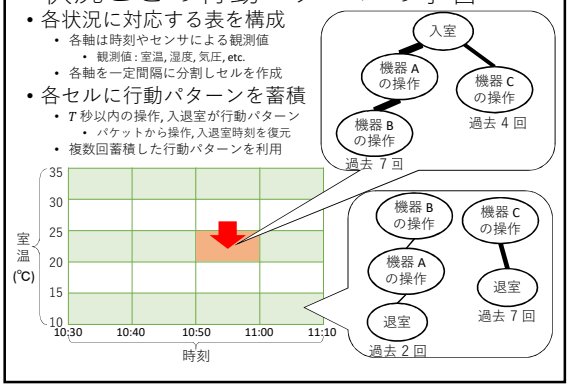
状況ごとの行動パターンの学習と検知

1. 状況ごとの表を構成
   - 各軸は時刻やセンサによる観測値
     - 観測値：室温, 湿度, 気圧, etc.
   - 各軸を一定間隔に分割
2. 各状況ごとに行動パターンを蓄積
   - 状況ごとの行動パターンを学習可能
3. 検知対象と行動パターンのマッチング
   - 完全一致しない動作を検知



状況ごとの行動パターンの学習

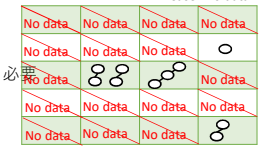
- 各状況に対応する表を構成
  - 各軸は時刻やセンサによる観測値
    - 観測値：室温, 湿度, 気圧, etc.
  - 各軸を一定間隔に分割しセルを作成
- 各セルに行動パターンを蓄積
  - T 秒以内の操作, 入退室が行動パターン
    - パケットから操作, 入退室時刻を復元
  - 複数回蓄積した行動パターンを利用



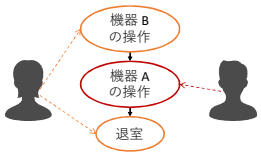
学習の問題点 – small and mixed data

- スモールデータの学習
  - 機械学習を用いるにはデータ数が不足
    - 各セルへの蓄積データ数が少ない
    - データ : 機器の操作, ユーザの入退室
  - 少ないデータ数でも学習できる工夫が必要

Case : 10 data

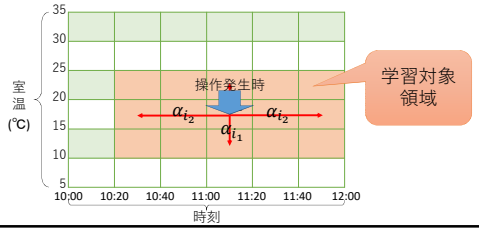


- 他のユーザの行動が混在
  - 家庭内には複数のユーザ
  - 他のユーザの操作を取り除く必要



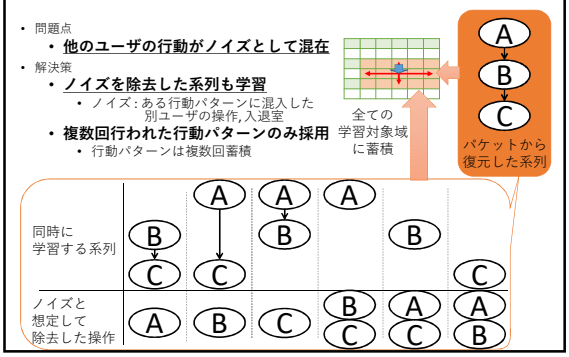
スモールデータの学習手法

- 問題点
  - **スモールデータの学習**
- 解決策
  - **学習対象領域の拡大**
    - 操作発生時の軸 i の値 x_i に対して $x_i - \alpha_i$ から $x_i + \alpha_i$ のセルに学習
    - 類似した状況でも同じ操作が行われる可能性
      - 例 : 18 時の機器操作は 17 時や 19 時にも発生



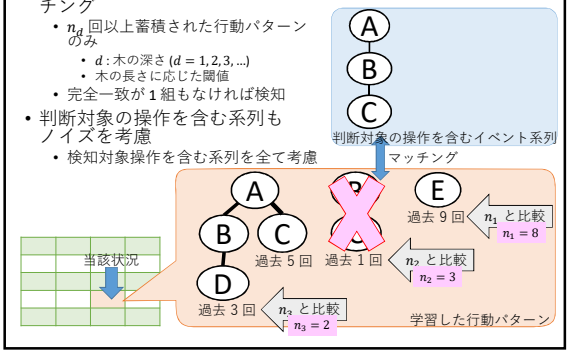
他ユーザの行動が混在する状況での学習手法

- 問題点
  - **他のユーザの行動がノイズとして混在**
- 解決策
  - **ノイズを除去した系列も学習**
    - ノイズ : ある行動パターンに混入した別ユーザの操作, 入退室
  - **複数回行われた行動パターンのみ採用**
    - 行動パターンは複数回蓄積



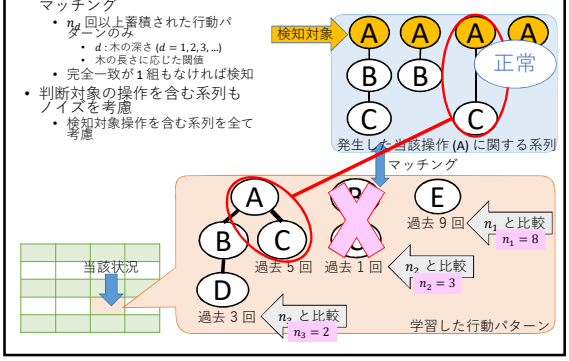
不正操作パケットの検知

- 蓄積された行動パターンとのマッチング
  - n_d 回以上蓄積された行動パターンのみ
    - d : 木の深さ ($d = 1, 2, 3, \ldots$)
    - 木の長さに応じた閾値
  - 完全一致が 1 組もなければ検知
- 判断対象の操作を含む系列もノイズを考慮
  - 検知対象操作を含む系列を全て考慮



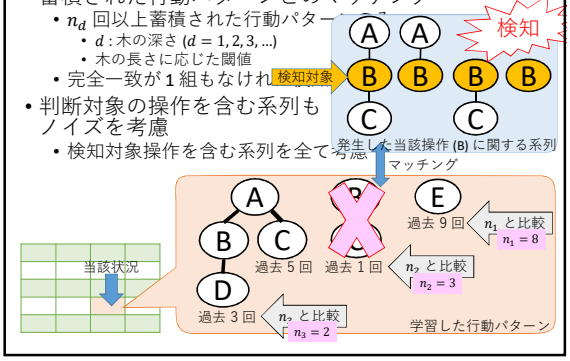
不正操作パケットの検知

- 蓄積された行動パターンとのマッチング
  - n_d 回以上蓄積された行動パターンのみ
    - d : 木の深さ ($d = 1, 2, 3, \ldots$)
    - 木の長さに応じた閾値
  - 完全一致が 1 組もなければ検知
- 判断対象の操作を含む系列もノイズを考慮
  - 検知対象操作を含む系列を全て考慮



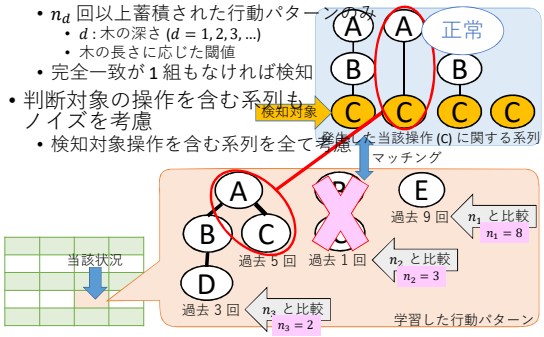
不正操作パケットの検知

- 蓄積された行動パターンとのマッチング
  - n_d 回以上蓄積された行動パターンのみ
    - d : 木の深さ ($d = 1, 2, 3, \ldots$)
    - 木の長さに応じた閾値
  - 完全一致が 1 組もなければ検知
- 判断対象の操作を含む系列もノイズを考慮
  - 検知対象操作を含む系列を全て考慮



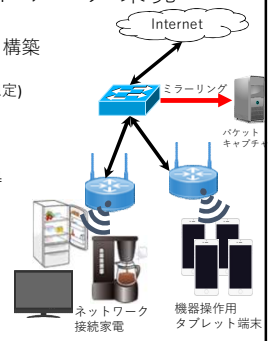
不正操作パケットの検知

- 蓄積された行動パターンとのマッチング
  - n_d 回以上蓄積された行動パターンのみ
    - d : 木の深さ ($d = 1, 2, 3, \ldots$)
    - 木の長さに応じた閾値
  - 完全一致が 1 組もなければ検知
- 判断対象の操作を含む系列もノイズを考慮
  - 検知対象操作を含む系列を全て考慮



評価用ホームネットワーク環境

- 研究室内にホームネットワーク構築
  - 設置 IoT 機器 : 15 種類
  - 被験者 : 4 名 (複数人が住む家を想定)
  - 実験期間 : 1 ヵ月 (2017 年 1 月)
  - データセット数 : 29 日分
- パケットとセンサ観測値を取得
  - パケットから復元
    - 被験者の入退室時刻
    - 機器操作の時刻
    - 機器の mac アドレスは既知
- 環境変数に時刻のみ利用
  - 1 か月間の気温変化は微小



評価方法

- 評価用データセット
  - パケットデータに不正操作を混入したもの
  - 正常操作 : パケットデータに元から含まれる機器操作
  - 不正操作 : ランダムな時刻に 1 日あたり 100 回分混入
- 評価手法
  - LOO-CV (Leave-One-Out Cross-Validation) による評価
    - 学習データ : 特定の 1 日分を除くデータ
    - テストデータ : 特定の 1 日分のデータに不正操作を混入
    - 各日の検知した不正操作数と誤って検知した正常操作数を合算
- 評価指標
  - 混入した全不正操作のうち不正操作であると検知した割合

$$\text{検知率} = \frac{\text{検知した不正操作数}}{\text{混入させた不正操作数}}$$

  - 被験者が実際に行った全操作のうち誤って不正操作と検知した割合

$$\text{誤検知率} = \frac{\text{検知した正常操作数}}{\text{パケットデータに含まれる正常操作数}}$$

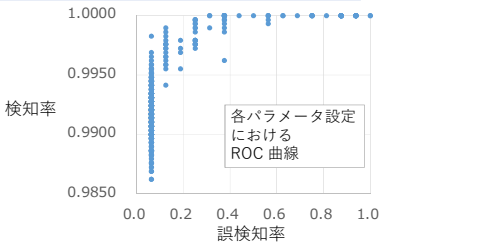
評価結果 - 検知率・誤検知率

- ヒータに対する不正操作攻撃の検知が可能
  - 検知率 : **99.6** %　誤検知率 : **6.25** %

検知率 (検知数/混入不正操作総数)	誤検知率 (誤検知数/正常操作総数)
0.9962 ($^{2889}/_{2900}$)	0.0625 ($^{1}/_{16}$)

パラメータ設定

T	n_1	$n_{2,3,\ldots}$	α_i
300 (sec)	9 (回)	3 (回)	31200 (sec)



参考文献

1. "IoT マルウエア大量感染の現状と対策," 日経コンピュータ, 2016年 8 月 18 日号, pp.78–81, 2016年8月.
2. "ウクライナの「ハッキングによる大規模停電」で強まる、サイバー戦争への懸念" Wired, 2016年1月7日
3. IoT 推進コンソーシアム,"IoTセキュリティガイドライン"
4. NTT、三菱電機,"多様な IoT 機器に自動適応してサイバー攻撃を検知," SIP/重要インフラ等におけるサイバーセキュリティの確保 シンポジウム, 2017年10月.
5. 大下裕一, 村田正幸, "特徴量写像関数の学習による類似インシデント検索," 電子情報通信学会技術研究報告(IN2017-57), pp. 67-72, 2017年12月.
6. 山内雅明, 大下裕一, 村田正幸, 上田健介, 加藤嘉明, "スマートホームIoTにおけるユーザ行動の学習に基づく異常検知手法," 電子情報通信学会技術研究報告(IN2017-58), pp. 73-78, 2017年12月.

謝辞

本講演で紹介した研究開発の一部は、総合科学技術・イノベーション会議の戦略的イノベーション創造プログラム（SIP）「重要インフラ等におけるサイバーセキュリティの確保」（管理法人：NEDO）にて実施されております。