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

山内 雅明[1], 大下 裕一[1], 村田 正幸[1], 上田 健介[2], 加藤 嘉明[3]

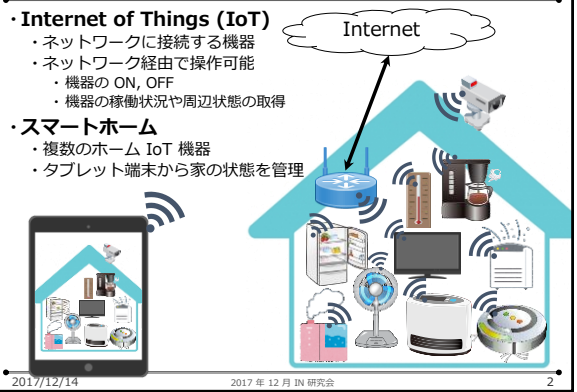
[1]大阪大学 大学院情報科学研究科
[2]三菱電機株式会社 先端技術総合研究所
[3]三菱電機株式会社 情報技術総合研究所

大阪大学 OSAKA UNIVERSITY

Mitsubishi Electric Cybersecurity Research Alliance Laboratories
三菱電機サイバーセキュリティ協働研究所

IoT , スマートホーム

- **Internet of Things (IoT)**
  - ネットワークに接続する機器
  - ネットワーク経由で操作可能
    - 機器の ON, OFF
    - 機器の稼働状況や周辺状態の取得
- **スマートホーム**
  - 複数のホーム IoT 機器
  - タブレット端末から家の状態を管理

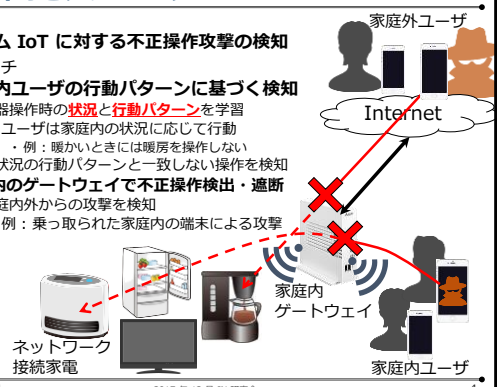


IoT 機器の不正操作

- **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
- **パターンマッチングによる検知が困難**
  - 正常な操作に用いるプロトコルに従って不正操作

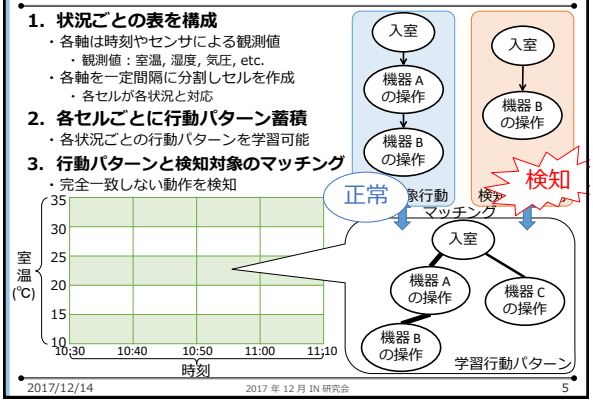
研究目的とアプローチ

- 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
        - 例：暖かいときには暖房を操作しない
    - 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



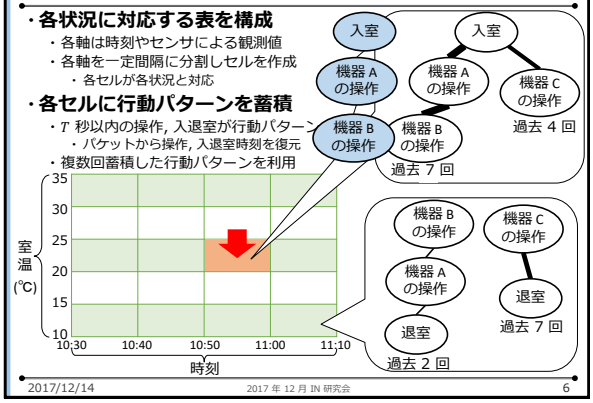
状況ごとの行動パターンの学習と検知

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しセルを作成
     - 各セルが各状況と対応
2. **各セルごとに行動パターン蓄積**
   - 各状況ごとの行動パターンを学習可能
3. **行動パターンと検知対象のマッチング**
   - 完全一致しない動作を検知



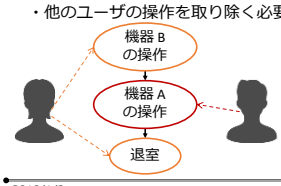
状況ごとの行動パターンの学習

- **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
  - 各軸を一定間隔に分割しセルを作成
    - 各セルが各状況と対応
- **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

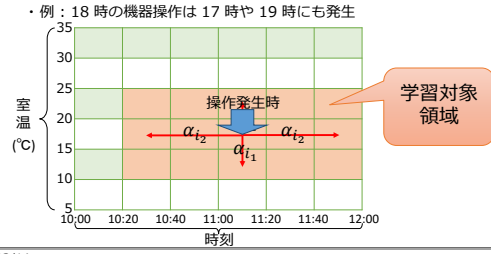
- **スモールデータの学習**
  - 機械学習を用いるにはデータ数が不足
    - データ：機器の操作, ユーザの入退室
    - データが存在しない領域
    - 蓄積データ数が少ない領域
  - 少ないデータ数でも学習できる工夫が必要
- **他のユーザの行動が混在**
  - 家庭内には複数のユーザ
  - 他のユーザの操作を取り除く必要



Case : 10 data

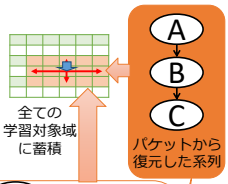
スモールデータの学習手法

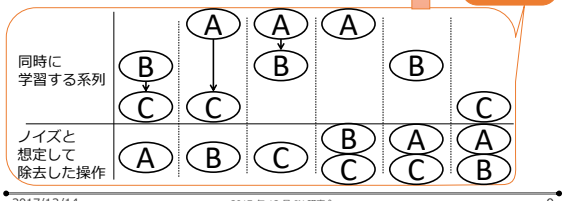
- 問題点
  - **スモールデータの学習**
- 解決策
  - **学習対象領域の拡大**
    - 操作発生時の軸 i の値 x_i に対して $x_i - \alpha_i$ から $x_i + \alpha_i$ のセルに学習
    - 類似した状況でも同じ操作が行われる可能性
      - 例：18 時の機器操作は 17 時や 19 時にも発生



他ユーザの行動が混在する状況での学習手法

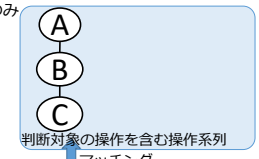
- 問題点
  - **他のユーザの行動がノイズとして混在**
- 解決策
  - **ノイズを除去した系列も学習**
    - ノイズ：ある行動パターンに混入した別ユーザの操作, 入退室
  - **複数回行われた行動パターンのみ採用**
    - 正しい行動パターンは複数回蓄積

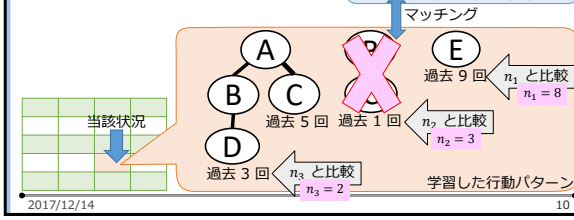




不正操作パケットの検知

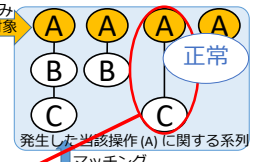
- **蓄積された行動パターンとのマッチング**
  - n_d 回以上蓄積された行動パターンのみ
    - d：木の深さ ($d = 1, 2, 3, \ldots$)
    - 木の長さに応じた閾値
  - 完全一致が 1 組もなければ検知
- **判断対象の操作を含む系列もノイズを考慮**
  - 検知対象操作を含む系列を全て考慮

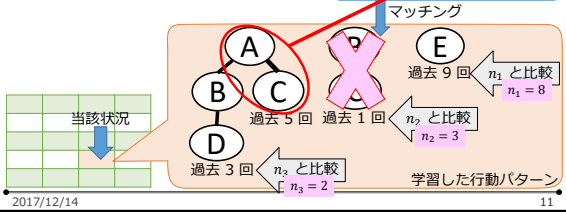




不正操作パケットの検知

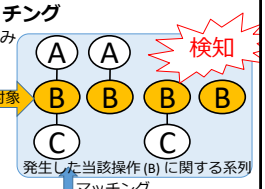
- **蓄積された行動パターンとのマッチング**
  - n_d 回以上蓄積された行動パターンのみ
    - d：木の深さ ($d = 1, 2, 3, \ldots$)
    - 木の長さに応じた閾値
  - 完全一致が 1 組もなければ検知
- **判断対象の操作を含む系列もノイズを考慮**
  - 検知対象操作を含む系列を全て考慮

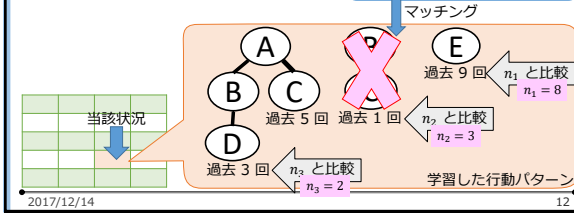




不正操作パケットの検知

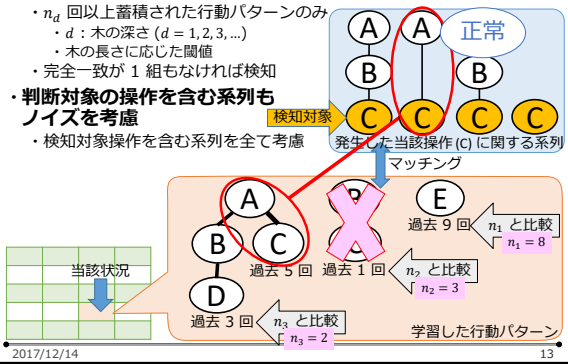
- **蓄積された行動パターンとのマッチング**
  - n_d 回以上蓄積された行動パターンのみ
    - d：木の深さ ($d = 1, 2, 3, \ldots$)
    - 木の長さに応じた閾値
  - 完全一致が 1 組もなければ検知
- **判断対象の操作を含む系列もノイズを考慮**
  - 検知対象操作を含む系列を全て考慮





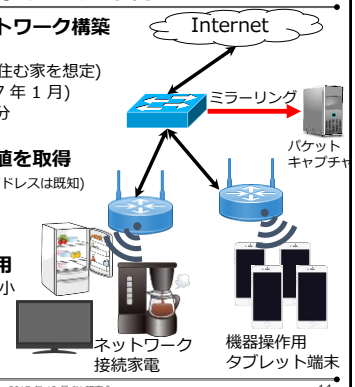
不正操作パケットの検知

- **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 カ月 (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 (macアドレスは既知)
    - 被験者の入退室時刻
    - 機器操作の時刻
- **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

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

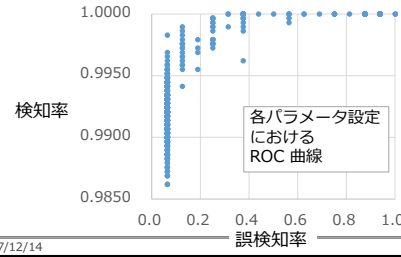
評価結果 - 検知率・誤検知率

- **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)

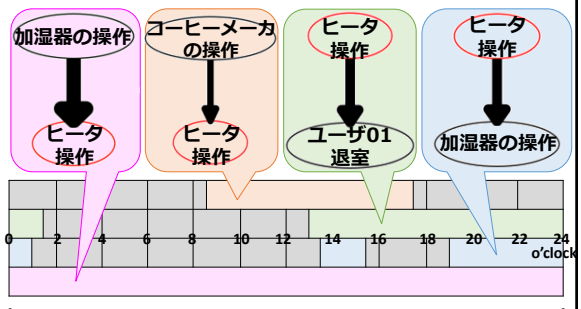
パラメータ設定

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



評価結果 – 状況ごとの行動パターン

- **ヒータの操作が他の機器操作やユーザの退室と関連**
- **特定の時間帯に操作が集中**



まとめと今後の課題

- **不正操作攻撃は状況ごとの行動パターン学習により検知可能**
  - 家庭内ユーザは状況に応じて行動
  - 誤検知率を 6.25% に抑え 99.6% の不正操作を検知
- **今後の課題**
  - 複数の機器に対する不正操作攻撃
  - より長期間のデータを使って学習した場合の精度
  - 学習するのに最低限必要となるデータ数
  - 家庭内以外の環境