


# スマートホーム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

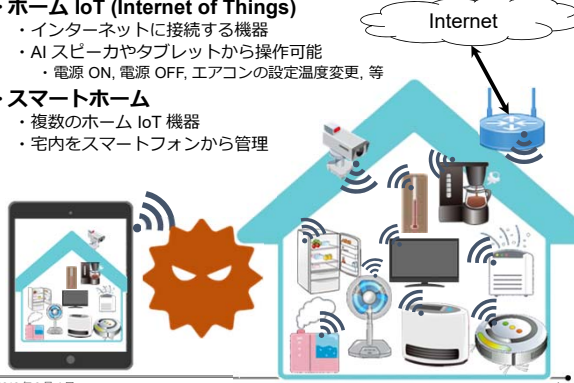
大阪大学 大学院情報科学研究科  
情報ネットワーク学専攻 村田研究室  
博士後期課程 1年

 大阪大学  
OSAKA UNIVERSITY

2019年9月4日 FIT 2019

## IoT, スマートホーム


- ・**ホーム IoT (Internet of Things)**
  - ・インターネットに接続する機器
  - ・AI スピーカやタブレットから操作可能
  - ・電源 ON, 電源 OFF, エアコンの設定温度変更, 等
- ・**スマートホーム**
  - ・複数のホーム IoT 機器
  - ・宅内をスマートフォンから管理



2019年9月4日 FIT 2019

## IoT 機器への不正操作攻撃


- ・**第三者による操作バケットの送信**
  - ・ユーザへの重大な被害
    - ・例：ヒータの不正操作による火傷や火災
    - ・例：ペースメーカーへの不正な設定変更による心停止
- ・**バケットの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通信
  - ・乗っ取ったタブレット端末、AI スピーカを経由する可能性 [1]
  - ・統計的手法や深層学習などの手法の適用が困難
    - ・宅内での機器操作の回数が少なく、学習データが不足



2019年9月4日 [1] armis, "Airborne Cyber Threats Reach Amazon Echo and Google Home," <https://www.armis.com/airborne-cyber-threats-reach-amazon-echo-google-home/news/air-security.com/2018/08/01/> 2

## 研究目的とアプローチ

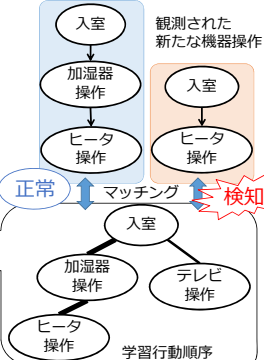
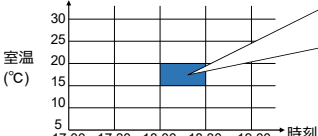
- ・[研究目的]
  - ・IoT 機器に対する不正操作攻撃の検知
- ・[アプローチ]
  - ・**ユーザが機器を操作する際の「癖」をもとに検知**
    - ・ユーザは状況に応じて一定の行動
      - ・例：部屋が寒い時、ヒータ ⇒ 加湿器と操作
  - ・**機器が操作される状況と行動順序を学習**
    - ・状況：時刻や、室温などの観測値
    - ・行動順序：機器操作や入退室などが行われる順序
    - ・学習結果と一致しない機器操作を異常として検知
  - ・**家庭内のゲートウェイで学習、検知**
    - ・ゲートウェイは宅内外の機器操作バケットを監視可能



2019年9月4日 FIT 2019

## 行動パターンの学習と検知

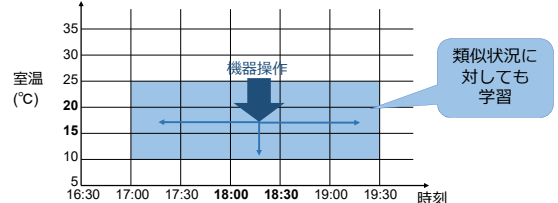
1. **宅内状況の分類表を定義**
  - ・各軸は時刻やセンサの観測値
  - ・センサ：室温、湿度 など
  - ・一定間隔に分割し、セルを生成
  - ・各セルが各状況に相当
2. **行動順序の蓄積**
  - ・行動：機器操作、ユーザの入退室
  - ・木構造で保存
3. **検知対象とのマッチング**
  - ・完全一致しない場合は検知

2019年9月4日 FIT 2019

## スモールデータの学習

- ・[問題点]
  - ・**各状況における機器操作回数が非常に少ない**
    - ・ユーザが機器を操作する回数が限られている
    - ・1日に家で機器を操作することは、数回程度
- ・[アイデア]
  - ・**類似した状況に対してと同時に学習**
    - ・ユーザは似たような状況下で同じ行動を行う



2019年9月4日 FIT 2019

### ノイズとなる他ユーザの機器操作の除去

- ・[問題点]
  - ・他のユーザによる操作がノイズとして混入
    - ・家庭内にはユーザが複数いることが多い
- ・[アイデア]
  - ・ノイズを除去した多種の行動パターンを同時に蓄積
    - ・ユーザが実際に行う行動パターンが蓄積可能
  - ・複数回蓄積された行動パターンのみを検知に利用
    - ・ユーザが実際に行う行動パターンは何度も行われる

蓄積する行動パターン	入室	入室	TV	入室	ヒータ	TV	入室
ゲートウェイで観測した行動順序	None	入室	TV	ヒータ	入室	ヒータ	TV
除去した行動							

2019年9月4日 FIT 2019 6

### 実験環境

- ・研究室内にホームネットワークを想定した環境を構築
  - ・IoT家電：13種類
    - ・ヒータ、コーヒーマカ、など
  - ・IoTセンサ：9種類
    - ・室温センサ、湿度センサ、など
  - ・実験期間：1か月間を繰り返し実施
  - ・被験者：学生4名
- ・全パケットとセンサ観測値を取得
  - ・機器操作、ユーザの入退室時刻を記録
- ・時刻のみで宅内状況を分類
  - ・研究室内の室温変化は微小

2019年9月4日 FIT 2019 7

### 評価方法

- ・データセット
  - ・キャプチャしたパケットから抽出した、機器操作および入室の履歴
  - ・正常操作：パケットに含まれる、被験者による機器操作履歴
  - ・不正操作：偽の操作履歴をランダムな時刻に100(回/日)分混入
- ・パラメータ設定
  - ・月の最初の1週間分のデータを利用
  - ・残りの3週間分のデータをLOO-CVに利用
- ・評価手法
  - ・LOO-CV (Leave-One-Out Cross-Validation)
    - ・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
    - ・各日の結果を合算し検知率、誤検知率を算出
- ・評価指標
  - ・検知率 =  $\frac{\text{検知した不正操作数}}{\text{混入した不正操作数}}$
  - ・誤検知率 =  $\frac{\text{誤検知した正常操作数}}{\text{正常操作の総数}}$

2019年9月4日 FIT 2019 8

### 1か月間のデータに対する評価結果

1. 検知率 95--100%、誤検知 数回程度 (誤検知率 19% 未満)
  - ・各状況に対する行動パターンが十分に学習されたため
  - ・検知漏れは、実際に機器操作が行われうるタイミングに存在
2. 「単発操作」を含む機器は検知が困難
  - ・「単発操作」：前後に他の行動が観測できない機器操作
  - ・操作順序の情報が使えず、状況の情報のみで検知するため
  - ・特定の時間帯に操作が集中していない
3. 同じ行動パターンの蓄積回数が少ない場合は検知が困難
  - ・「レアな操作」 2017年1月(1か月間)

	検知率	検知数 / 合計	誤検知率	誤検知数 / 合計
ヒータ	0.959	2110/2200	0.182	2/11
テレビ B	1.000	2200/2200	0.000	0/2
コーヒーマカ	0.157	346/2200	0.000	0/48
加湿器	0.080	176/2200	0.000	0/38
テレビ A	1.000	2200/2200	1.000	8/8

2019年9月4日 FIT 2019 9

### ヒータに関して学習された行動パターン

- ・他の機器操作とユーザの退室と関連した行動パターン
- ・特定の時間帯に操作が集中

2019年9月4日 FIT 2019 10

### 1か月間のデータに対する評価結果

1. 検知率 95--100%、誤検知 数回程度 (誤検知率 19% 未満)
  - ・各状況に対する行動パターンが十分に学習されたため
  - ・検知漏れは、実際に機器操作が行われうるタイミングに存在
2. 「単発操作」を含む機器は検知が困難
  - ・「単発操作」：前後に他の行動が観測できない機器操作
  - ・操作順序の情報が使えず、状況の情報のみで検知するため
  - ・特定の時間帯に操作が集中していない
3. 同じ行動パターンの蓄積回数が少ない場合は検知が困難
  - ・「レアな操作」 2017年1月(1か月間)

	検知率	検知数 / 合計	誤検知率	誤 /
ヒータ	0.959	2110/2200	0.182	
テレビ B	1.000	2200/2200	0.000	
コーヒーマカ	0.157	346/2200	0.000	0/48
加湿器	0.080	176/2200	0.000	0/38
テレビ A	1.000	2200/2200	1.000	8/8

2019年9月4日 FIT 2019 11

### 1 か月間のデータに対する評価結果

- 検知率 95~100%、誤検知 数回程度 (誤検知率 19% 未満)
  - 各状況に対する行動パターンが十分に学習されたため
  - 検知漏れは、実際に機器操作が行われうるタイミングに存在
- 「単発操作」を含む機器は検知が困難
  - 「単発操作」: 前後に他の行動が観測できない機器操作
  - 操作順序の情報が使えず、状況の情報のみで検知するため
  - 特定の時間帯に操作が集中していない
- 同じ行動パターンの蓄積回数が少ない場合は検知が困難
  - 「レアな操作」

2017年1月(1か月間)

	検知率	検知数 / 合計	誤検知率	誤検知数 / 合計
ヒータ	0.959	2110/2200	0.182	2/11
テレビ B	1.000	2200/2200	0.000	0/2
コーヒーメーカー	0.157	346/2200	0.000	0/48
加湿器	0.080	176/2200	0.000	0/38
テレビ A	1.000	2200/2200	1.000	8/8

2019年9月4日

FIT 2019

12

### 3 か月間のデータに対する評価結果

- 検知率は 99% 以上で、誤検知は 18% 未満
- テレビの誤検知率が 1 か月間の結果よりも低下
  - より多くの行動パターンが蓄積されたため
- 単発操作を含む機器は検知が困難

2017年4,6,8月(3か月間)

	検知率	検知数 / 合計	誤検知率	誤検知数 / 合計
扇風機 A	0.998	6384/6400	0.000	0/9
扇風機 B	0.999	6399/6400	0.000	0/6
テレビ A	0.996	6377/6400	0.171	7/41
テレビ C	0.999	6397/6400	0.000	0/10
テレビ D	0.999	6398/6400	0.111	1/9
コーヒーメーカー	0.611	3908/6400	0.058	3/52

2019年9月4日

FIT 2019

13

### 3 か月間のデータに対する評価結果

- 検知率は 99% 以上で、誤検知は 18% 未満
- テレビの誤検知率が 1 か月間の結果よりも低下
  - より多くの行動パターンが蓄積されたため
- 単発操作を含む機器は検知が困難

2017年4,6,8月(3か月間)

	検知率	検知数 / 合計	誤検知率	誤検知数 / 合計
扇風機 A	0.998	6384/6400	0.000	0/9
扇風機 B	0.999	6399/6400	0.000	0/6
テレビ A	0.996	6377/6400	0.171	7/41
テレビ C	0.999	6397/6400	0.000	0/10
テレビ D	0.999	6398/6400	0.111	1/9

2017年1月(1か月間)

	検知率	検知数 / 合計	誤検知率	誤検知数 / 合計
テレビ A	1.000	2200/2200	1.000	8/8

2019年

14

### 3 か月間のデータに対する評価結果

- 検知率は 99% 以上で、誤検知は 18% 未満
- テレビの誤検知率が 1 か月間の結果よりも低下
  - より多くの行動パターンが蓄積されたため
- 単発操作を含む機器は検知が困難

2017年4,6,8月(3か月間)

	検知率	検知数 / 合計	誤検知率	誤検知数 / 合計
扇風機 A	0.998	6384/6400	0.000	0/9
扇風機 B	0.999	6399/6400	0.000	0/6
テレビ A	0.996	6377/6400	0.171	7/41
テレビ C	0.999	6397/6400	0.000	0/10
テレビ D	0.999	6398/6400	0.111	1/9
コーヒーメーカー	0.611	3908/6400	0.058	3/52

2019年9月4日

FIT 2019

15

### まとめと今後の課題

- 行動パターンが蓄積できた機器に関しては、95-100%の不正操作を検知可能で誤検知率は 19% 未満
  - 「単発操作」を含む機器の検知が困難
  - 「レアな操作」を含む機器の検知が困難

#### 今後の課題

- 「単発操作」の検知
- 「レアな操作」による誤検知の低減



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

本研究成果は、三菱電機サイバーセキュリティ協働研究所での成果の一部です。

2019年9月4日

FIT 2019

16