通信トラヒック分析に基づく IoTデバイスの発動機能推定手法の検討

小池 大地[1] 石田 繁巳[1] 荒川 豊[1]

概要：近年，IoT デバイスが多くの場面で利用されるようになっており，今後もデバイスの数は増え続けると考えられる．IoT デバイスは外部ネットワークに接続されていることによりハッキングの対象にもなり，プライバシー流出の原因となる可能性が指摘されている．現在の IoT デバイスは動作がブラックボックスであることから，デバイスがユーザの意図しない通信を行っていた場合に気づく術がない．そこで，我々は IoT デバイスの動作状況の可視化システム（IoT 活動量計）の実現を目指している．その実現のため，本研究では，IoT デバイスの通信トラヒックを分析し，どのような機能が使われているかを推定する手法を提案する．パケットキャプチャアプリである Wireshark を用いてスマートスピーカー Amazon Echo Spot から送出されるデータに対して初期的評価を行い，通信トラヒックから発動した 10 種類の機能を精度 56.1％で推定できることを確認した．

Called Function Identification of IoT Devices by Network Traffic Analysis

Daichi Koike[1] Shigemi Ishida[1] Yutaka Arakawa[1]

1. はじめに

近年，IoT デバイスが多くの場面で利用されるようになっている．総務省によると 2018 年の世界の IoT デバイスの数は約 307 億台であり，2021 年には約 448 億台にのぼると予測される [1] など，今後もデバイスの数は増え続けると考えられている．代表的な家庭内 IoT デバイスとしては，Google Home や Amazon Echo などのスマートスピーカーが挙げられる．これらの機器は，VUI（Voice User Interface）を備え，音声コマンドによって，ネット検索やタイマーと言った様々な機能を利用することができる．デバイスによっては，カメラを備え，機器間でビデオ電話が可能なスマートスピーカーもある．日本では一般的ではないが，海外では防犯のためのネットワークカメラも代表的な IoT デバイスであり，ホームセンターで多数販売されるほど，一般的に普及している．このような IoT デバイスの特徴として，クラウドおよびスマートフォンとの連携がある．大半の IoT デバイスは，家庭内の WiFi ネットワークを通じて専用のクラウドシステムに接続され，クラウド上にデータが集まり，クラウド上でデータ分析がなされる．そして，クラウドを介して，ユーザのスマートフォンアプリと連携しており，スマートフォンアプリに通知を送ったり，スマートフォンアプリから宅内の IoT 機器の制御や設定を行ったりすることができる．

このように，IoT デバイスは外部ネットワークに接続することを前提として設計されており，その結果として，ハッキングの対象にもなり，種々のプライバシー流出事件が起きている．例えば，スマート掃除機に付属するカメラが簡単にハッキング可能であることが判明したこと [2] や，カジノの水槽に設置されたスマート水温計からカジノの顧客情報が盗み出された事件 [3]，IP カメラがハックされ全世界のカメラにアクセス可能になった事件 [4] などが実際に起きている．IoT デバイスが普及することによって，我々の生活が便利になる一方で，IoT デバイスのプライバシーについても考えることは非常に重要である．

現在，IoT デバイスは，次々と新しいデバイスが発売されている一方で，パソコンのようなファームウェアの継続的なアップデートが不十分な状況である．大手企業によって発売される高価なデバイスであればある程度のサポートは期待できるが，安価な IoT デバイスは今後も売りっぱなしの状況のままであることが予想される．また，IoT デバ

[1] 九州大学大学院システム情報科学研究院

イスは動作がブラックボックスであるということも問題点である．初期設定の際にネットワーク接続をしたら，ネットワーク接続が切れるまでこのIoTデバイスがどのサーバと, どのような通信を, どの程度の頻度で行っているかを知る術はなく，もし，デバイスがユーザの意図しない通信を行っていた場合にも気づくことはできない．昨今，Zoomで問題となったが，通信が特定の国を経由している可能性も大いに考えられる [5]．その経路は基本的に暗号化されている事になっているが，そのことも一般的な利用者が確認する方法はない．さらに，パソコンにおけるアンチウイルスソフトなどを追加で導入することもできない．

これまでに挙げた問題をまとめると，本稿で取り扱うIoTデバイスを取り巻く問題は，以下のようになる．

- IoTデバイスの動作はブラックボックスであり，ユーザが意図したとき以外にも動作しているかもしれない．あるいは，意図しない情報もクラウドに送っているかもしれない．
- IoTデバイスは多様性が高く，すべての機器がきちんと継続的なセキュリティアップデートがなされるとは考え難い．
- IoTデバイスは，パソコンのようにあとからアンチウイルスソフトのような不正通信検知機能を追加することができない．

我々はこれらの問題を解決し，どのようなIoTデバイスにおいてもその動作をきちんと確認して，信頼して家庭内に設置できる仕組みを作りたいと考えている．例えば，スマートフォンの場合，アプリケーションをインストールした際にカメラやマイクなどのリソースに対して，それぞれ利用の可否をユーザに対して確認するようになっている．また，設定画面において各機能へアクセスしているアプリを確認ができるなど，リソースに対するアクセス情報が可視化されている．

本稿では，スマートフォンのリソース制御と同様に，IoTデバイスがどのような通信を行っているかを検知や理解することを可能にする動作状況の可視化システム（IoT活動量計）を提案する．IoT活動量計その実現手段として我々は，IoTデバイスの通信トラヒックパターンに着目し，そのパターンからデバイス及び，デバイスのどのような機能が使われているかを推定する．IoTのトラヒック分析に基づき，IoTデバイスの識別に関する研究はこれまでにいくつか提案されている [6,7]．しかし，スマートスピーカーなどのIoTデバイスは機能が多岐に渡っており，デバイスの分類ができただけでは，不正な動作を検出することができない．そこで，IoTデバイスの種類だけではなく，個々のデバイスの機能まで識別することを目指す．前提として，家庭内のルータがすべての通信トラヒックを収集できるものとし，ルータが通過するトラヒックを観察し，どのデバイスがどの機能を使用しているのかを識別する．

本稿においては，パケットキャプチャアプリであるWiresharkをインストールしたパソコンをルータ化し，そこにIoTデバイスを接続する形でトラヒックを観察した．今回，IoTデバイスとしては，スマートスピーカーAmazonEchoSpotを用い，10通りの機能を各3〜5回ずつ繰り返して発動させ，データ収集を行った．収集したデータは，48.9MBになった．

初期的評価として，教師あり学習のランダムフォレストアルゴリズムを用い，10分割交差検証で発動機能の推定精度を評価した．このとき，特徴量としては，パケットの平均，分散，標準偏差，各通信プロトコル (21種類) の出現回数という24個を用いている．機能を使用している際のトラヒックデータを固定幅ウィンドウで区切り，ウィンドウ内のデータから特徴量を抽出して発動機能を推定した．ウィンドウサイズは，データ間の秒数が均一でない (通信が起こった時にデータを取得する) という理由からデータ30個としている．その結果，10通りの機能について，平均精度56.1％で発動機能を推定できることを確認した．

本稿の構成は以下の通りである．2.ではIoTトラヒック分析の関連研究を述べ，3.ではIoTデバイスの機能推定手法を示す．4.でデータ取得方法を示し，その初期的評価を5.で示す．最後に6.でまとめとする．

2. 関連研究

2.1 IoT通信のプライバシーに関する脆弱性を示す研究

Noahら [8] は暗号化されたIoTトラヒックのプライバシーに関する脆弱性を示している．4つの商用利用可能なスマートホームデバイス (睡眠モニターSense，屋内セキュリティカメラNest Cam，リモートスイッチWeMo switch，スマートスピーカーAmazon Echo) を分析することで，通信トラヒックレートがユーザの活動を明らかにしてしまうことを示している．これはIoTデバイスが実世界の情報を通信トラヒックに変換するため，暗号化されたトラヒックの送受信レートのみによりユーザの行動を予測できてしまうからであり，新しいプライバシーの脅威をユーザに与えることを警鐘している．行動予測を可能とするトラヒック情報を攻撃者から守ることが重要である一方で，ユーザにわかりやすく可視化した状態で伝えることはユーザが安心してデバイスを利用することの助けとなる．筆者らは，このようなプライバシーの問題を解決する手段の1つとして，本研究，すなわちIoTデバイスのトラヒックを分析し，利用状況をユーザに提示することで利用状況を可視化し，不審な通信を発見しやすくできると考えている．

2.2 通信トラヒック分析によるIoTデバイスの推定手法

本研究はデバイスの通信トラヒック分析から機能を推定しているが，先行研究としてIoTデバイスの識別手法が研究されている．

Yairら [6] は機械学習を用いたトラヒック分析によるIoTデバイス及び非IoTデバイスの識別手法を示している．WiFiに接続されたデバイスの通信トラヒック情報が保存されているファイルを分析して送信アドレス，受信アドレ

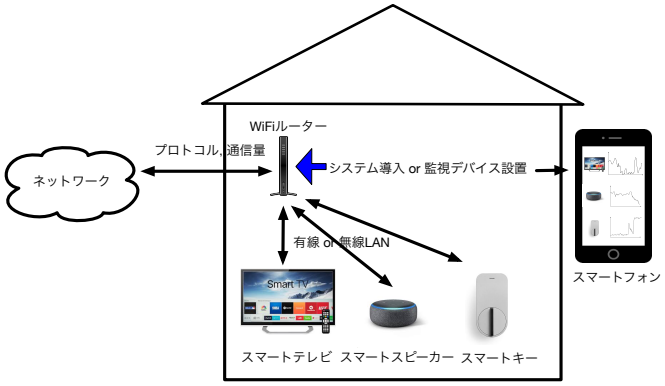

図 1 「IoT 活動量計」プラットフォームの概要

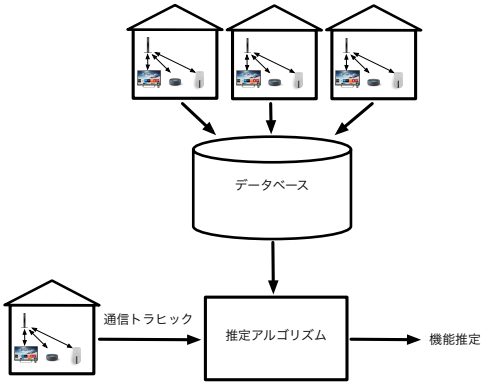

図 2 実現したいシステムの概要

ス，ポート番号などを抽出し，これらを特徴量として教師あり学習により 2 段階でデバイスを識別する．まず IoT デバイスであるかを識別し，IoT デバイスと識別されたデバイスがあらかじめ登録されたデバイスクラスのうちどれに属するかを識別する．この手法では，99.281 %という高精度で IoT デバイスの種類を識別しているものの，IoT デバイスが使用している機能までは識別していない．

Arunan ら [7] はスマートシティ，キャンパスにおける IoT デバイスの特徴付け及び分類手法を示している．21 個の IoT デバイスをキャンパスに設置し，通信トラヒックデータを 3 週間に渡って取得する．そして幅広い通信トラヒック (トラヒックのロードや信号の種類，アクティブ時間の分布など) を分析し，教師あり学習によりデバイスを識別する．この手法においても，95 %という高精度で IoT デバイスの種類は識別しているが，使用されている機能は識別していない．

3. 通信トラヒック分析に基づく発動機能推定手法の検討

3.1 実現したいシステム概要

図 1 に実現したいプラットフォームを示す．本研究では，ある宅内で複数の IoT デバイスが利用されており，それらが有線もしくは無線で WiFi ルータに接続されていることを前提としている．全てのデバイスがインターネットと通信する際に WiFi ルータなどのホームルータに着目し，ホームルータにおいて IoT デバイスから送出される通信情報を監視，記録して発動機能を推定する．その上で，IoT 活動量計として発動機能をユーザに提示することで IoT デバイスに対するトラストを形成する．

筆者らは「IoT 活動量計」を「所有する IoT デバイスを簡単に登録でき，その動作状況を確認可能にするプラットフォーム」と定義している．スマートフォンにおいては，アプリケーションをインストールした際に “アプリケーションが機能 (例えばマイクなど) へのアクセスを求めています” などというポップアップが表示され，許可するかどうかを選択できる．また，スマートフォンの設定画面から各機能へアクセスしているアプリケーションを一覧で確認できる．このように，スマートフォンではアプリケーションが利用する機能が可視化されており，ユーザが安心して利用できる．そこで筆者らは，スマートフォンのように IoT デバイスの動作を可視化し，スマートフォン等のモニタに IoT デバイスが利用する機能及び通信パケットの時間変位を表示させ，ユーザが安心して IoT デバイスを利用できるようにしたいと考えている．

IoT デバイスの機能を推定する際に，直接 IoT デバイスにアクセスして情報を得てデバイスの使用機能を推定することは現実的ではない．デバイスによってシステムの規格が異なり，個々のデバイスに対応した機能推定を行うことは莫大な手間がかかる．また，新しいデバイスが出現するたびに推定手法を追加しなければならないという問題も生じる．

本研究では，IoT デバイスが WiFi ルータなどのホームルータを介してクラウドシステムと接続されていることに着目し，すべての通信が経由するホームルータ内で通過トラヒックを観察することで，IoT 機器の動作機能を推定する．推定した機能をクラウドを介してユーザのアプリに通知することで，ユーザは外出先からも IoT 機器の動作状況を確認することが可能になる．

図 2 に実現したいシステムの概要を示す．上述のプラットフォームを各家庭に導入し，IoT デバイスの発動機能を推定するアルゴリズムを共用することで十分な量の通信トラヒックデータを収集し，高い精度での発動機能推定を実現する．アルゴリズムは教室あり機械学習の利用を想定している．ユーザに対して低頻度に使用機能に関するアンケートを行うことでラベル付けを行い，収集データとラベルを用いてモデルを定期的に更新することを想定している．

本稿では提案するシステムの中でも，IoT デバイスの発動機能を推定する部分の検討について報告する．

3.2 基本アイデア

通信トラヒック分析に基づく IoT デバイスの発動機能推定手法の基本アイデアは，すべての通信が経由するホームルータ内において IoT デバイスの通信トラヒックデータを取得し，それらを特徴量とする機械学習アルゴリズムを用いて IoT デバイスの発動機能を推定することである．ホームルータで観測した通信トラヒックデータから通信の特徴

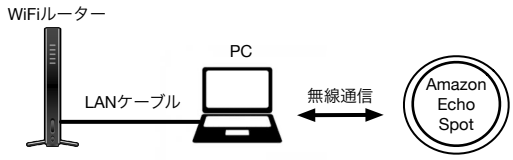

図 3　データ収集環境

図 4　実験環境

量を取得し，発動機能推定アルゴリズムによって使用されている機能を推定する．

通信トラヒックの時間変動が各機能に特徴的なものであることから，通信トラヒックデータ 30 個を 1 ウィンドウとし，ウィンドウごとのデータの変動を表す平均，分散，標準偏差のような特徴量を抽出する．そして，抽出した特徴量を用いて各種の教師あり機械学習によってデバイスの使用機能を推定する．

3.3　発動機能推定手法

提案方式は，通信トラヒックのパターンを機械学習によって学習することで発動機能を推定する．推定モデルの構築に向けて，スマートスピーカー Amazon Echo Spot を用いてデータの収集を行い，いくつかの機械学習アルゴリズムを適用して，精度の評価を行った．

具体的な機能としては，Kindle, Amazon Music, ビデオ通話，質問, ニュース, Amazon Prime Video, レストラン検索, Spotify，TuneIn，音声通話を使用した．機能呼び出し 1 回ごとに通信トラヒックデータを切り出し，切り出したデータブロックごとに，ビデオ通話，音声通話，動画再生などのラベルを付与した．特徴量として，エラーパケットを除いたパケットの平均, 分散, 標準偏差と各プロトコルでの通信回数を計算し，教師あり機械学習のいくつかのアルゴリズムをもとに推定モデルを作成し，テストデータを用いてモデルを評価する．最後にそれぞれのモデルの比較を行い，比較の結果最も精度の高いモデルを選定し，推定モデルとする．

4.　データ取得

今回，パケットキャプチャ機能を有する WiFi アクセスポイントを構築する手法として，macOS のインターネット

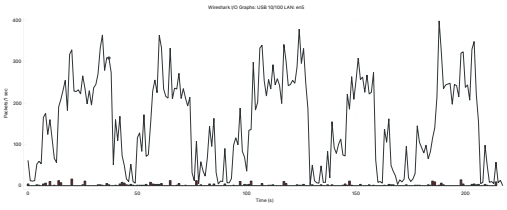

図 5　Kindle

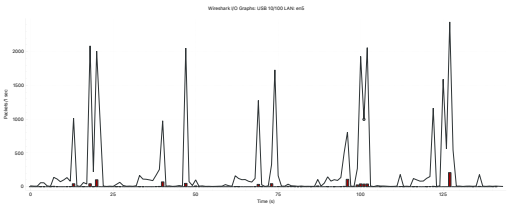

図 6　AmazonMusic

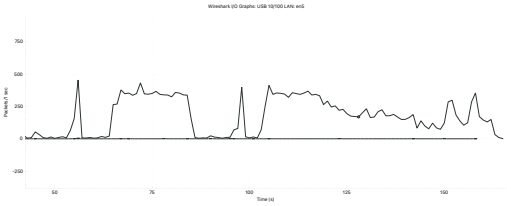

図 7　ビデオ通話

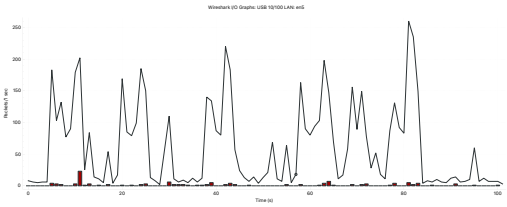

図 8　質問

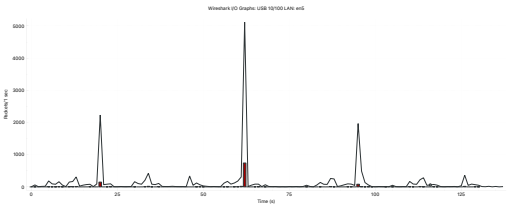

図 9　ニュース

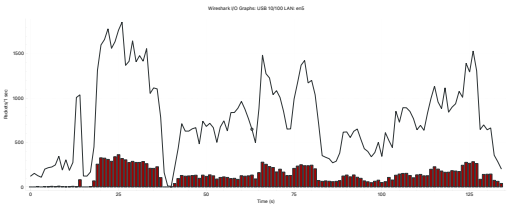

図 10　Amazon Prime Video

共有機能を用いた．パケットキャプチャアプリケーションである Wireshark をインストールしたパソコンをアクセスポイントとして機能させることで，そのパソコンを通過した通信トラヒックをすべてキャプチャすることができる．

このアクセスポイントにスマートスピーカーを接続し，

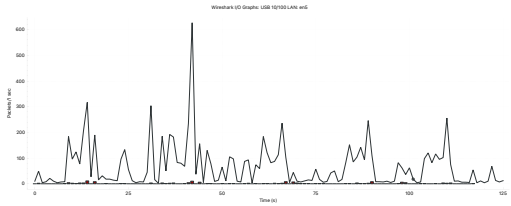

図 11 レストラン

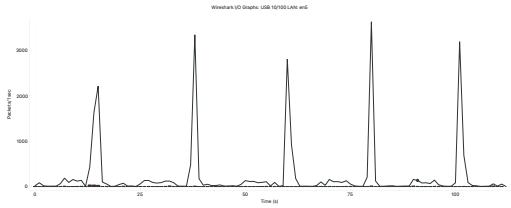

図 12 Spotify

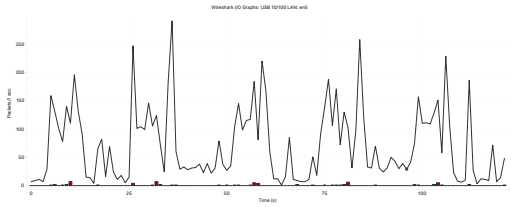

図 13 TuneIn

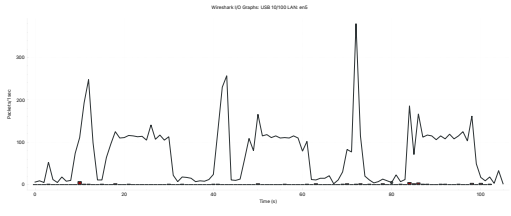

図 14 音声通話

そのデバイスから送信されたすべての通信を記録する．データ収集実験では，スマートスピーカーの種々の機能を約 2 分間ずつ利用し，各機能を 3~5 回程度呼び出す．そして利用時の通信トラヒックデータ（プロトコル，通信量）を記録する．**図 3** にデータ収集環境を，**図 4** に実験環境を示す．そして**図 5** に Kindle，**図 6** に Amazon Music，**図 7** にビデオ通話，**図 8** に質問，**図 9** にニュース，**図 10** に Amazon Prime Video，**図 11** にレストラン，**図 12** に Spotify，**図 13** に TuneIn，**図 14** に音声通話を呼び出した時のパケットグラフをそれぞれ示す．AmazonMusic のパケットグラフでは機能を呼び出した瞬間の通信量が高く，呼び出し後に急激に低くなり非常に低い値で推移している．一方ビデオ通話のパケットグラフは機能呼び出し時の通信量が低く，呼び出し後もほぼ一定の通信量で推移している．以上より機能ごとにパケットの時間変位の特徴が異なることが分かる．

5. 評価

発動機能推定手法の有効性を検証するためにランダムフォレストアルゴリズムを用いて，10 分割交差検証により評価を行った．上記のパケットグラフの特徴より，各機械学習アルゴリズムの特徴量は，データ 30 個ごとのパケットの平均，分散，標準偏差と各プロトコルでの通信回数を用いた．機械学習アルゴリズムについては，ランダムフォレスト，SGD，SVC，K 近傍，線形 SVM，カーネル近似を用いて精度を測定し，最も精度の高かったランダムフォレストを推定モデルとした．

図 15 にランダムフォレスト，**図 16** に SGD，**図 17** に SVC，**図 18** に K 近傍，**図 19** に線形 SVM，**図 20** にカーネル近似の発動機能推定結果の混同行列を，**表 1** にランダ

表 1 発動機能推定結果の混同行列
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

	0	1	2	3	4	5	6	7	8	9
0	57	1	0	20	3	0	5	1	2	5
1	2	42	0	22	11	0	2	3	3	6
2	1	0	157	10	0	0	1	0	1	13
3	8	3	5	212	8	6	16	5	8	46
4	1	22	0	41	12	0	3	1	2	8
5	0	1	0	21	1	56	2	1	1	3
6	2	4	0	51	1	4	3	0	5	15
7	2	4	0	19	3	1	0	53	5	4
8	2	2	1	28	3	2	2	6	21	5
9	3	2	5	76	0	6	4	4	8	228

表 2 発動機能推定結果の混同行列
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

	0	1	2	3	4	5	6	7	8	9
0	57	1	0	20	3	0	5	1	2	5
1	2	42	0	22	11	0	2	3	3	6
2	1	0	157	10	0	0	1	0	1	13
3	8	3	5	212	8	6	16	5	8	46
4	1	22	0	41	12	0	3	1	2	8
5	0	1	0	21	1	56	2	1	1	3
6	2	4	0	51	1	4	3	0	5	15
7	2	4	0	19	3	1	0	53	5	4
8	2	2	1	28	3	2	2	6	21	5
9	3	2	5	76	0	6	4	4	8	228

表 3 発動機能推定結果の混同行列
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	87	7	0	0
1	0	0	0	0	0	0	62	28	0	1
2	0	0	0	0	0	0	179	4	0	1
3	0	0	0	0	0	0	243	74	0	0
4	0	0	0	0	0	0	66	24	0	0
5	0	0	0	0	0	0	83	3	0	0
6	0	0	0	0	0	0	71	14	0	0
7	0	0	0	0	0	0	64	27	0	0
8	0	0	0	0	0	0	64	8	0	0
9	0	0	0	0	0	0	146	128	0	0

表 4　発動機能推定結果の混同行列
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

	0	1	2	3	4	5	6	7	8	9
0	1	0	11	73	0	0	0	0	0	9
1	0	0	21	43	0	0	0	0	0	27
2	2	0	164	10	0	0	0	0	0	7
3	4	0	60	134	0	0	0	0	0	119
4	1	0	20	41	0	0	0	0	0	28
5	0	0	37	42	0	0	0	0	0	7
6	0	0	20	40	0	0	0	0	0	25
7	0	0	20	49	0	0	0	0	0	22
8	3	0	20	38	0	0	0	0	0	11
9	1	0	48	70	0	0	0	0	0	217

表 5　発動機能推定結果の混同行列
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

	0	1	2	3	4	5	6	7	8	9
0	41	2	7	24	4	3	3	2	2	6
1	8	36	4	21	7	3	1	1	0	10
2	3	3	139	27	1	2	1	1	1	5
3	30	13	43	132	4	21	9	4	8	53
4	7	21	6	26	8	3	3	3	2	11
5	9	5	20	29	3	12	2	1	3	2
6	3	6	14	44	2	3	1	1	0	11
7	16	2	12	17	1	4	3	28	1	7
8	8	2	14	24	2	7	0	4	4	7
9	22	5	25	106	3	3	3	0	4	165

表 6　発動機能推定結果の混同行列
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

	0	1	2	3	4	5	6	7	8	9
0	0	83	0	11	0	0	0	0	0	0
1	0	81	0	6	0	0	0	0	0	4
2	0	172	0	10	0	0	0	0	0	1
3	0	152	0	151	0	0	0	0	0	14
4	0	67	0	20	0	0	0	0	0	3
5	0	73	0	13	0	0	0	0	0	0
6	0	43	0	40	0	0	0	0	0	2
7	0	79	0	10	0	0	0	0	0	2
8	0	49	0	22	0	0	0	0	0	1
9	0	100	0	71	0	0	0	0	0	165

ムフォレスト，**表 2** に SGD，**表 3** に SVC，**表 4** に K 近傍，**表 5** に線形 SVM，**表 6** にカーネル近似のヒートマップを示す．発動機能推定手法の推定精度は 56.1 %である．

ビデオ通話, 質問, 音声通話は比較的高い精度で推定できいる．これはビデオ通話と音声通話に関しては機能発動中のパケットが一定に近いからだと考えられる．またビデオ通話, 質問, 音声通話の全てにおいて, 他の機能に比べて多くのデータを取得したからだと考えられる．(通話においては発信と着信を別々に取り, 質問においては天気検索や計算などを別々に取り, 後に結合したのでデータ量は多くなる.)

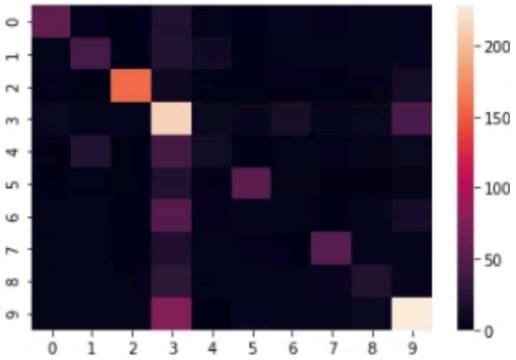

図 15　発動機能推定結果の混同行列のヒートマップ (ランダムフォレスト)
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

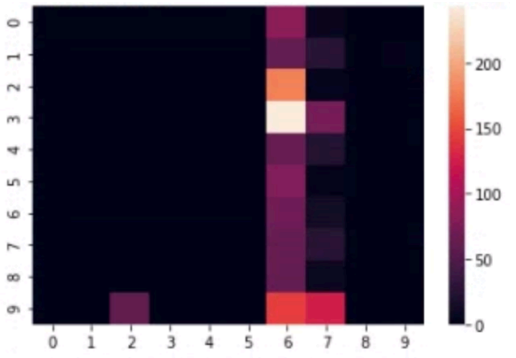

図 16　発動機能推定結果の混同行列のヒートマップ (SGD)
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

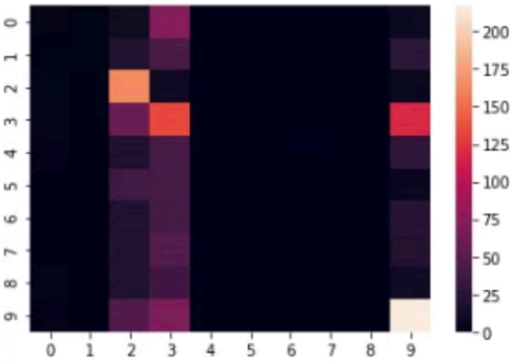

図 17　発動機能推定結果の混同行列のヒートマップ (SVC)
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

精度があまり良くない原因として分類対象が 10 種類と多いこと，データ数が少ないこと，データ 30 個を 1 ウィンドウとしたために送信元アドレス，送信先アドレスなどを特徴量として扱えなかったこと，また，機能呼び出し間の何も通信していないアイドリング時間があるにもかかわ

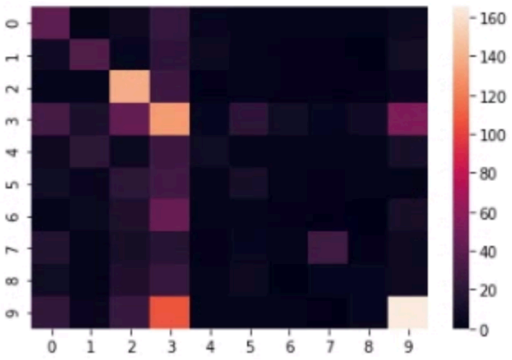

図 18　発動機能推定結果の混同行列のヒートマップ (K 近傍)
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

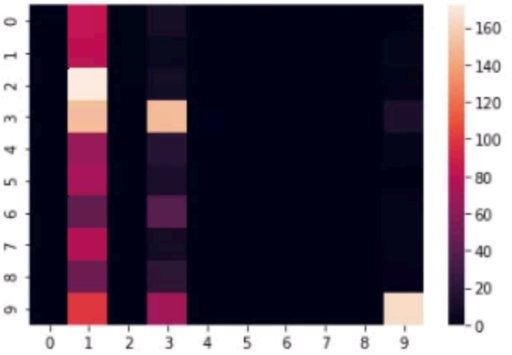

図 19　発動機能推定結果の混同行列のヒートマップ (線形 SVM)
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

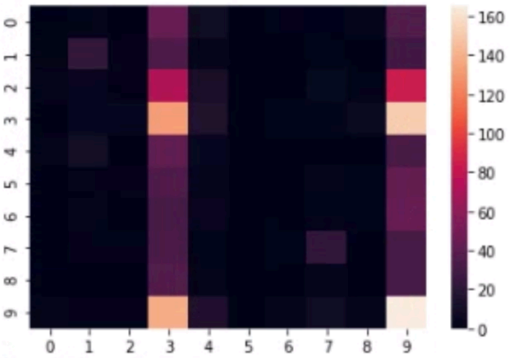

図 20　発動機能推定結果の混同行列のヒートマップ (カーネル近似)
0:Kindle，1:AmazonMusic，2:ビデオ通話，3:質問，4:News，5:AmazonPrimeVideo，6:レストラン検索，7:Spotify，8:TuneIn，9:音声通話

らず，その時間を各機能としてラベル付けしていることなどが挙げられる．

今後はデータ数を増やし，アイドリング時間を 1 つのラベルとして設定するなどの前処理を検討し，さらなる精度の向上を目指す．

6. おわりに

本稿では，IoT デバイスの活動量の可視化に向けて，通信トラヒック分析に基づく IoT デバイスの発動機能推定手法を示した．スマートスピーカー Amazon Echo Spot を用いてトラヒックデータを取得し，10 種類の機能について推定精度 56.1 %で発動機能を推定できることを確認した．推定精度は現時点では高くはないが，アイドリング時間を設定することや特徴量を増やすことで精度の向上ができると考えている．また，今後はスマートスピーカーに限らず，様々な IoT デバイスに対して機能を推定する予定である．そして WiFi を用いた行動認識などの他の推定技術と組み合わせて通信がユーザの意図したものかどうかを識別し，結果を可視化することでユーザがより安心して IoT デバイスを利用できる世界の実現を目指す．

謝辞　本稿で示した研究の一部は，科研費（JP19KT0020）の助成で行われた．

参考文献

[1] 総務省：IoT デバイスの急速な普及. https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nd112120.html.

[2] Forbes: Time To Update Your Vacuum Cleaner – Hack Turns LG Robot Hoover Into A Spy (2017). https://www.forbes.com/sites/thomasbrewster/2017/10/26/lg-hom-bot-robot-hoover-hacked-into-surveillance-device/#71ccc44bf042.

[3] Forbes: Criminals Hacked A Fish Tank To Steal Data From A Casino(2017). https://www.forbes.com/sites/leemathews/2017/07/27/criminals-hacked-a-fish-tank-to-steal-data-from-a-casino/#7faff4e832b

[4] Techcrunch: 世界中の無防備な Web カメラを見せる Insecam... パスワードに無関心なアドミンが多い (2014). https://jp.techcrunch.com/2014/11/08/20141107insecam-displays-insecure-webcams-from-around-the

[5] Techcrunch: Zoom，北米の Web 会議の暗号キーを誤って中国データセンター経由にした問題について説明 (2020). https://www.itmedia.co.jp/news/articles/2004/05/news010.html/.

[6] Meidan, Y., Bohadana, M., Shabtai, A., Guarnizo, J. D., Ochoa, M., Tippenhauer, N. O. and Elovici, Y.: ProfilIoT: a machine learning approach for IoT device identification based on network traffic analysis, *Proceedings of the symposium on applied computing*, pp. 506–509 (2017).

[7] Sivanathan, A., Sherratt, D., Gharakheili, H. H., Radford, A., Wijenayake, C., Vishwanath, A. and Sivaraman, V.: Characterizing and classifying IoT traffic in smart cities and campuses, *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 559–564 (2017).

[8] Apthorpe, N., Reisman, D. and Feamster, N.: A smart home is no castle: Privacy vulnerabilities of encrypted iot traffic, *arXiv preprint arXiv:1705.06805* (2017).