Basic 認証応答を用いた IoT 機器識別技術の大規模適用実験

大塚 瑠莉[1,2,a)]　インミンパパ[3]　吉岡 克成[4]

概要：インターネット初期から現在に至るまで，IoT 機器や汎用的なウェブサーバなど幅広い分野で利用されている認証方式の一つに Basic 認証がある．広域スキャンシステムである Shodan では，Basic 認証応答をするホストが，2025 年 8 月時点で全世界において 245 万件以上，同じく広域スキャンシステムである Censys では，2025 年 7 月時点で 460 万件以上確認されている．我々は LLM を活用してこれらのホストの機器識別を行う手法を提案し，ネットワーク機器や産業用機器を含む 434 機種の推定が可能であることを示したが，この実証実験では，Shodan の利用制限から 1,759 件のホストのみを対象にしていた．本報告では，より広い範囲のホストを対象とする．広域スキャンシステムである Censys において，HTTP/1.0 で稼働している 48 万件以上のホストを対象として提案手法を適用した．対象ホストのうち，先行研究で調査済みの 24 種類の Basic realm をもつホストを除外した 254,016 ホストへアクセスしたところ，16,796 種類の Basic realm を確認できた．この Basic realm を用いて機器推定をした結果，264 種類のメーカーの 934 種類の機器が推定された．

キーワード：Basic 認証, IoT, デバイスフィンガープリント, HTTP

Large-Scale Experiment on IoT Device Identification Using Basic Authentication Responses

Ruri Otsuka[1,2,a)]　Yin Minn Pa Pa[3]　Katsunari Yoshioka[4]

Abstract: Basic authentication is one of the authentication methods that has been widely used from the early days of the Internet to the present, ranging from IoT devices to general-purpose web servers. Shodan, a global scanning system, has confirmed over 2.45 million hosts responding with Basic authentication worldwide as of August 2025, while Censys, another global scanning system, has confirmed over 4.6 million hosts as of July 2025. We proposed a method for device identification of these hosts using LLMs and demonstrated that it is possible to estimate 434 device models including network equipment and industrial equipment. However, this proof-of-concept experiment only targeted 1,759 hosts due to Shodan's usage limitations. In this report, we target a broader range of hosts. We applied the proposed method to over 480,000 hosts running on HTTP/1.0 in Censys. After excluding hosts with 24 types of Basic realms that had already been investigated in previous research, we accessed 254,016 hosts and confirmed 16,796 types of Basic realms. As a result of device estimation using these Basic realms, 934 types of devices from 264 manufacturers were estimated.

Keywords: Basic Authentication, IoT, device fingerprinting, HTTP

[1] 三菱電機株式会社
Mitsubishi Electric Corporation,Kamakura,Kanagawa 247-8501,Japan
[2] 横浜国立大学大学院環境情報学府/先端科学高等研究院
Graduate School of Environment and Information Sciences, Yokohama National University/ Institute of Advanced Sciences, Yokohama National University
[3] 横浜国立大学大学院先端科学高等研究院
Institute of Advanced Sciences, Yokohama National University
[4] 横浜国立大学大学院環境情報研究院/先端科学高等研究院

1. はじめに

Basic 認証 [1] は実装が容易で多くのウェブサービスに対応可能なことから，インターネット初期から現在に至るまで，IoT 機器や汎用的なウェブサーバなど幅広い分野で利用されている．インターネット接続機器の広域スキャンシステムである Shodan [2] では，スキャンシステムからの接続要求に対して Basic 認証応答を返信するホストが全世界で 2025 年 8 月時点で 245 万件以上，同じくインターネット接続機器の広域スキャンシステムである Censys [3] では，2025 年 7 月時点で 460 万件以上確認された．Basic 認証ではスキーム名「Basic」，認証パラメータ「realm」が必須であり，realm は HTTP 認証における保護領域を示す [1][4]．図 1 は Basic 認証における認証応答の一例である．Basic 認証を実装しているサービスの応答には，図 1 のように，realm にメーカーや型番と推定される文字列を含んでいるものが存在し，これらの情報から機器の種別を推定できる場合がある．昨今，IoT 機器を狙うサイバー攻撃が増大し，特に機器固有の脆弱性を狙う攻撃が増加していることから [5][6][7]，インターネット上に存在する機器の種別や型番を特定し，脆弱な機器に対しては適切な対策を行うことが重要となっている [8]. 我々は，Basic 認証の応答に含まれる情報を基に LLM により機器推定を行う手法を提案しており，Shodan から取得した Basic 認証が動作する 1,759 件のホストを対象に分析を行い，HTTP/1.0 について 225 機種，HTTP/1.1 について 215 機種の推定を行った [9].

本研究では，先行研究 [9] の調査範囲を拡大するために，インターネット接続機器の広域スキャンシステムである Censys [3] を新たに活用する．2025 年 7 月の調査時点で，HTTP/1.0 の Basic 認証応答を行うホストは Censys 上で 483,991 ホスト検知された．これらのホストのうち，先行研究 [9] で調査済みである 24 種類の Basic realm をもつホストを除外した 254,016 ホストを対象として，Censys が認識する全ての待ち受けポートに対して大規模ネットワークスキャンツールである zgrab [10] を用いてアクセスを行ったところ， 500,865 件の応答を入手した．入手した応答から「Basic realm」フィールドのみを抽出すると 16,796 種類の Basic realm が確認できた．そのうち，12,433 種類は 1 回ずつしか観測されておらず，その中には機器が使用する IP アドレスやランダムに設定された値などが含まれることが多かった．これらの情報は機器の推定に利用することが難しいため，2 回以上観測された 4,363 種類について LLM による推定を行った．具体的には，OpenAI の LLM [11] である gpt-4o に realm の情報を入力し，リアルタイムのウェブ検索機能を利用して当該機器の「メーカー名」「機器名」「信頼度」「参照した Web URL」を回答させた．実験の結果，「メーカー名」「機器名」が特定され，さらに LLM 自身が出力した判定の信頼度が 10 段階中 9 以上となる回答が 1,207 機種分得られた．このうち，同じ機器を推定した重複を除くと，934 種類の機器を推定した．

今回の調査で判定された 934 機種のうち，先行研究 [9] で発見した機器との重複は 93 種類のみであり，本実験によって新たに 841 種類の機器を推定することができた．

```
HTTP/1.0 401 Unauthorized
WWW-Authenticate: Basic realm=“ルータメーカー名 型番”
x-frame-options: SAMEORIGIN
Set-Cookie: XSRF_TOKEN=1222440606; Path=/
Content-type: text/html
```

図 1 Basic 認証応答の例

Fig. 1 Example of basic authentication response.

2. Basic 認証要求応答による機器推定

先行研究 [9] で我々は，インターネット接続デバイス検索エンジンである Shodan [2] を使用し，Basic 認証が動作し，かつ，インターネット上でアクセス可能な機器の Basic 認証応答を収集した．一般的に，Basic 認証で保護されている領域に対してクライアントがアクセスをすると，HTTP 認証を使用するサーバは 401 Unauthorized 応答を行う [1]. そこで，Shodan のウェブインタフェース上で，「Basic」「realm」「HTTP/XX 401 Unauthorized」[*1] を全て含む応答を行うホスト群を検索クエリで指定することで，Basic 認証応答を検索した．検索の結果，ホスト数が上位である Basic 認証応答を抽出し，HTTP/1.0 は 760 種類，HTTP/1.1 は 999 種類の Basic 認証応答を入手し，これを調査対象とした．

機器推定には ChatGPT search [12] を使用し，入手した Basic 認証応答を入力することで，該当する Basic 認証応答を行う機器の推定を行った．調査対象に対して本手法を適用した結果，HTTP/1.0 で 239 種類，HTTP/1.1 で 245 種類の機器を推定することができた．そのうち，8 種類については両方の HTTP バージョンではたらいていたため，重複を除くと，合計で 476 種類の機器を推定した．

3. 評価実験

推定が可能な機器の正解データ: Basic 認証が動作する機器を大量に入手することが困難であったため，2 章で得た調査結果に対して評価を行った．調査対象は，HTTP/1.0 の 760 種類，HTTP/1.1 の 999 種類の Basic 認証応答のうち，realm に続く文字列と一致する機器がウェブ検索で発

Faculty of Environment and Information Sciences, Yokohama National University / Institute of Advanced Sciences, Yokohama National University
a) Otsuka.Ruri@bp.MitsubishiElectric.co.jp

[*1] XX は HTTP のバージョンであり，1.0 もしくは 1.1

見でき，かつ，マニュアルが入手でき，マニュアル内でBasic 認証が使用されていると判断できるものに限定したところ，50 種類の機種を抽出できた．この 50 種類については機器推定の精度が十分に高いとみなし，正解データとして評価に使用した．

評価結果: 50 機種の正解データに対する，提案手法の正答数と正答率を表 1 に示す．ChatGPT search の出力結果で，「メーカー名」「機種名」が一致したものを正答とした．HTTP/1.0 の 20 種類の機種のうち，機種を推定できた機種数は 16 種類であり，正答率は 0.80 であった．HTTP/1.1 の 31 種類の機種のうち，機種を推定できた機種数は 29 種類であり，正答率は 0.94 であった．HTTP/1.0, HTTP/1.1 を合わせると，機種数は 44 種類であり，正答率は 0.88 であった．合計機種数が 44 種類であるのは，推定した機種のうち 1 種類については，HTTP/1.0, HTTP/1.1 どちらでも推定ができ重複を除いたためである．HTTP/1.0 で機種を推定できた 16 種類の信頼度の平均は 9.75，分散は 0.15 であった．HTTP/1.1 で機種を推定できた 29 種類の信頼度の平均は 9.48，分散は 0.22 であった．このように正しく機器を推定できた場合に提案手法は最大値の 10 に近い，高い信頼度を出力した．

表 1 ウェブ調査を根拠とする正解データでの評価結果
Table 1 Evaluation result with correct data based on web research.

HTTP version	正解データの数	正答数	正答率
HTTP/1.0	20	16	0.80
HTTP/1.1	31	29	0.94
合計	50[*2]	44[*3]	0.88

推定が困難な機器の正解データ: 推定が困難と判断された機器に対して，評価を行う．具体的には，2 章で得た，HTTP/1.0 の 760 種類，HTTP/1.1 の 999 種類の Basic 認証応答のうち，realm に続く文字列をウェブ検索した際に一致する文字列を含む機器がなく，かつ、回答項目のいかなる情報も入手できなかった Basic 認証応答 642 種類 (HTTP/1.0: 226 種類，HTTP/1.1: 416 種類) を，推定が困難な機器の正解データとして扱う．

評価結果: 642 種類の正解データに対する，提案手法の正答数と正答率を表 2 に示す．提案手法による出力結果のうち，回答項目の全てが「UNKNOWN」と回答されているものを正答とした．HTTP/1.0 の 226 種類の Basic 認証応答のうち，ChatCPT search でいずれの項目についても推定できなかった数は 149 種類であり，正答率は 0.66 であった．HTTP/1.1 の 416 種類の Basic 認証応答のうち，ChatCPT search でいずれの項目についても推定できなかった数は 238 種類であり，正答率は 0.57 であった．HTTP/1.0，HTTP/1.1 を合わせると，Basic 認証応答の種類は 642 種類であり，提案手法でいずれも推定できなかった数は 387 種類，正答率は 0.60 であった．HTTP/1.0 でいずれも推定できなかった 149 種類の信頼度の平均は 1.32，分散は 2.99 であった．HTTP/1.1 でいずれも推定できなかった 238 種類の信頼度の平均は 1.08，分散は 2.15 であった

また，642 種類の正解データに対して信頼度 9 以上を出力したのは 128 種類（HTTP/1.0 が 25 種類，HTTP/1.1 が 103 種類）であり，このなかに回答項目の全てが「UNKNOWN」と回答された正答はなかった．これらは全て誤推定であるが，誤推定した各項目を確認すると，入力した Basic 認証応答の Server フィールドの文字列から機器推定を行っていることが分かった．

表 2 機器推定できないと判断された正解データでの評価結果
Table 2 Evaluation results on correct data judged to be undetectable by the device

HTTP version	正解データの数	正答数	正答率
HTTP/1.0	226	149	0.66
HTTP/1.1	416	238	0.57
合計	642	387	0.60

4. Censys を活用した大規模調査

Shodan の仕様上、確認できる Basic 認証応答の種類には限りがある．そこで，同じくインターネット接続デバイス検索エンジンである Censys [3] に着目した．Censys では，Shodan と同様に，ユーザが検索条件を入力することで，該当する IP アドレスごとの解放ポートや，ポートに対する応答などの情報を得ることができる．Censys で「Basic realm」「HTTP/1.0」「401 Unauthorized」をバナーに含むホストは 483,991 件（2025 年 7 月時点）だった一方，Shodan では 259,392 件（2025 年 8 月時点）であった．より大規模に調査を行うため，Censys のデータを元として，調査を拡大した．

4.1 調査

Censys で取得したデータから，機器推定に使用する入力データを作成するフローを図 2 に示す．Censys で「Basic realm」「HTTP/1.0」「401 Unauthorized」をバナーに含むホストは 483,991 件（2025 年 7 月時点）あった．本実験の目的は調査対象を拡大することであるため，先行研究 [9] で機器推定済みである Basic realm については入力データから除外する[*4]．入力データから除外する前に，機器推定

[*2] HTTP/1.0, HTTP/1.1 の Basic 認証応答が確認できた機種について，重複を除いたため機種数合計が 50 となっている．
[*3] HTTP/1.0, HTTP/1.1 の Basic 認証応答が確認できた機種について正答し，重複を除いたため機種数合計が 44 となっている．
[*4] 機器推定に使用する gpt-4o は月毎の使用量制限があり，広く推

済みである Basic realm について，Censys のデータに含まれるホスト数を調査した．2 章で調査した HTTP/1.0 の Basic 認証応答 760 種類に対し，Basic realm のみを抽出したところ 479 種類の Basic realm が得られた．479 種類の Basic realm を含むホスト数をそれぞれ Censys で調査したところ 416 種類については Censys でも同じ Basic realm を含む Basic 認証応答をするホストが存在することが判明した．このうち，ホスト数上位 24 種類の Basic realm については，検索クエリで指定をすることで，入力データから除外することとした[*5]．なお，除外した Basic realm を含むホスト数の合計は 194,286 件であり，全体の 40.1%を占める．24 種類の Basic realm を除いた検索クエリを実行し，収集したホストの IP アドレスとその IP アドレスの解放ポートに対し，「IP アドレス：ポート番号」の組を作成する．作成した組は 724,798 組であり，そのうち，IPv4 は 724,662 組，IPv6 は 136 組であった．今回は IPv4 のみを対象とし，724,662 組に対して，大規模ネットワークスキャンツールである zgrab [10] でのアクセスを行った．zgrab は指定した IP アドレスとポート番号に対して HTTP GET リクエストによるアクセスを行い，その応答を出力する．対象の負荷にならないよう考慮し，zgrab を実行する Python 側の並列処理数設定として max_workers=150, zgrab 側で 1 つの IP に対して接続を行う 1 秒当たりのレート制限として server-rate-limit=3 で設定し実験を行った．アクセスに成功し応答を入手できたのは 500,865 件であり，そのうち Basic 認証応答は 188,780 件であった．得られた Basic 認証応答から，Basic realm を抽出する．抽出した Basic realm の種類は 16,796 種類であり，そのうち 2 回以上出現した Basic realm の種類は 4,363 種類であった．これを入力データとして，機器推定を行う．

4.2 実験

機器推定には OpenAI の [11] の gpt-4o モデルを使用し，Responses API を有効にすることで，モデルが応答を生成する前に最新の情報をウェブで検索してから回答するようにした．プロンプトには，「あなたはセキュリティの専門家であること」「あなたの任務は，Basic 認証応答コンテンツに含まれる realm の値から機器情報を特定すること」と記載し，入力として Basic realm の値を与えた時に，出力として「メーカー名」「機器名」「信頼度」およびそれぞれの理由，回答の参考にした URL，が得られるようにした．「メーカー名」は機器の製造会社，「機器名」は機器の名前と型番，「信頼度」は 0 から 10 の範囲で回答させた．temperature=0 で設定した．

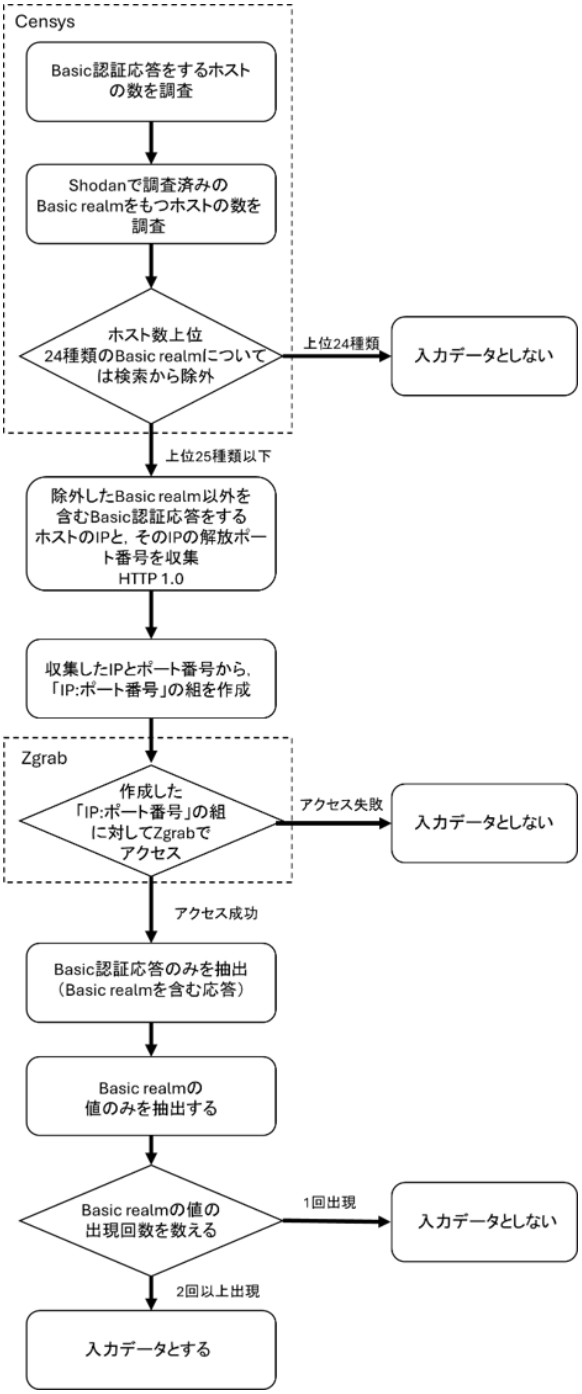

図 2 入力データ作成フロー
Fig. 2 Input data creation flow

定するには，可能な限り既知のデータを除外し，未知のデータを入力に使用する必要があるため．

[*5] Censys の検索クエリの仕様上，入力できるクエリ長に制限があり，24 種類の Basic realm が最長であったため．

4,363 種類の Basic realm を入力として与えた結果，出力に成功したのは 3,763 種類であった．出力に失敗した Basic realm については，gpt-4o が，機器推定するにはより具体的な情報が必要である，と出力し，機器推定結果が得られなかった．機器推定結果が出力された 3,763 種類に対し，信頼度の分布を図 3 に示す．縦軸が各信頼度の判定結果が得られた Basic realm の種類数，横軸が信頼度を示す．1 種類のみ，信頼度を UNKNOWN と出力された Basic realm があり，これは結果から除外した．最も頻度が高いのは信頼度 0 であり，次に頻度が高い回答は信頼度 9 である．3 章の評価実験では正答率 0.88 の機器推定における信頼度は 9.75，分散は 0.15 であり，LLM が出力する信頼度には一定の信ぴょう性がある．信頼度 9 以上かつ，「メーカー名」「機器名」どちらも UNKNOWN と回答されず，機器推定が行われた Basic realm は 1,207 種類であった．

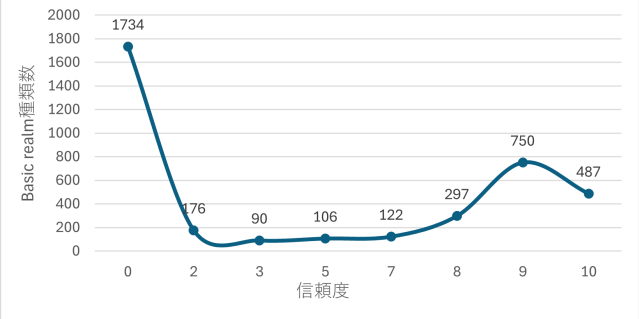

図 3 信頼度の分布
Fig. 3 Distribution of confidence levels

機器推定が行われた 1,207 種類の結果を整理すると，「maker」は 264 種類，「name」は 934 種類確認された．異なる Basic realm から同一の機器が推定された場合があるため，結果として，262 種類のメーカの 934 種類の機器が推定された．推定された機器の中には，一般消費者向けの機器として，ルータ，ゲートウェイ，IP カメラ，IP phone，DVB 衛星受信機などのネットワーク機器，音声認識サーバやウェブサーバなどのサーバ類が確認できた．産業用途の機器では，データロガー，PLC，HMI，コンバータや i/o モジュールなど，インターネットへ晒され攻撃された場合に影響が大きい製品も確認できた．さらに珍しいものでは，魚群探知ソナーも確認できた．

5. 比較実験

2 章に示した先行研究 [9] では，機器推定に ChatGPT search を使用し，入力として Basic 認証応答全体を与えていた．一方，今回の手法では機器推定に gpt-4o モデルを使用し，入力として Basic realm を用いている．そこで先行研究 [9] で使用した，Shodan から得られたデータに対して今回の手法を適用し，Censys から得られた今回の大規模調査の推定結果と比較を行う．先行研究 [9] で得られた HTTP/1.0 の Basic 認証応答 760 種類を対象とし，Basic realm を抽出したところ，391 種類の Basic realm が得られた．本 Basic 認証は，Shodan からデータを取得する際にホスト数が 2 以上あることを確認しているため，そのまま入力データとして扱う．プロンプトおよび temperture は同じ設定で機器推定を行ったところ，出力に成功したのは 387 種類であった．出力に成功した 387 種類の信頼度の分布を図 4 に示す．最も多いのが信頼度 10 であり，図 3 と比較して，全体的に信頼度が高い傾向が分かる．先行研究 [9] で Shodan から取得した Basic 認証応答は，ホスト数が上位であることから普及数が多く，Web 上に情報が存在することが多く，より高い信頼度で推定されていると思われる．さらに，今回の実験では先行研究 [9] で機器推定済みの Basic realm を一部除外しており，推定が難しい応答が相対的に多く含まれることになったと推測される．

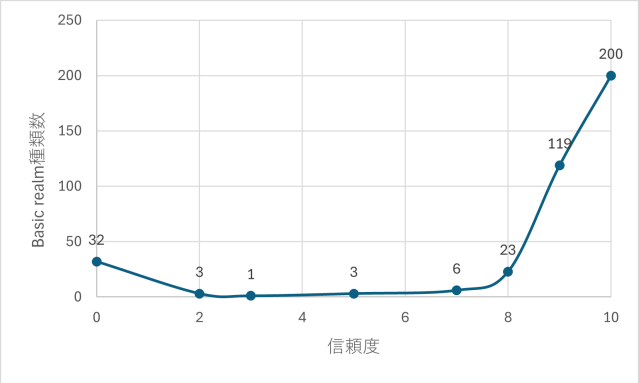

図 4 信頼度の分布
Fig. 4 Distribution of confidence levels

信頼度 9 以上の結果は 319 種類あり，これを整理すると，「メーカー」は 29 種類，「機種名」は 192 種類確認された．Censys による調査との比較を表 3 に示す．先行研究 [9] では，HTTP/1.0 で 239 種類の機器が推定できており，今回の手法を適用することで推定される機器は減少した．これは，Basic 認証応答全体を入力として機器推定を行う場合，Basic realm の値に加えて Server フィールドなどの他のフィールドを参照して推定を行える可能性があるのに対して，今回の手法では Basic realm のみを入力としていることが原因の 1 つと考えられる．一方，server フィールド等に含まれる情報は HTTP サーバ名などであることが多く，機器推定という観点では有効でない可能性もある．今回は zgrab で Basic 認証応答を取得しており，Basic 認証応答を全て復元するには適切な文字コード選択やエスケープ処理が必要である．より簡単に機器推定ができるよう，Basic realm のみを入力としたが，判定に用いる情報の有効性についてはさらに詳細な評価が必要である．

Basic 認証の Basic realm に着目することで，HTTP/1.0 に限定しても，900 種類以上の機器を推定することができ

表 3　比較実験の結果
Table 3　Results of comparative experiments.

	メーカー数	機種数
Censys による調査	262	934
比較実験の結果	29	192

たが，HTTP/1.1 の Basic 認証が動作する機器はさらに数が多いことから，Basic 認証応答に着目した機器推定には更なるポテンシャルがあるといえる．

6. 関連研究

IoT 機器のデバイスフィンガープリントに着目した研究は，本研究以外にも進められている．ネットワークに接続された IoT 機器を識別するために，IoT 機器のネットワークトラフィックに深層学習を適用する方法 [13] や，ネットワークトラフィックに含まれるホスト名やドメイン名などのテキストを特徴量としてラベル付けする方法 [14] がある．これらは IoT 機器を推定する材料としてパケットデータに着目している点は類似しているが，ネットワークトラフィックというパッシブな取得方法を選択している点で、スキャンにより応答を入手するというアクティブな取得方法を選択している本研究とは，情報収集方法の観点で異なる．文献 [15] はスキャンにより情報を取得するというアクティブな取得方法を選択している点で類似しているが，対象が ICS 機器であるため SCADA が一般的に使用するプロトコルを使用しており，対象機器や使用プロトコルが異なる．本研究は他の研究と比較し，gpt-4o モデルを使用したことにより，リアルタイムでのウェブ検索結果を反映させて機器推定をする点が特徴である．また，入力に realm のみを使用し，realm は一般的な接続要求に対する応答に含まれるため，Censys や Shodan などの既存の検索エンジンから入力に必要な情報を得ることができる．遠隔でも容易に機器推定が可能である点も特徴である．

7. 研究倫理に関する考察

本研究は，調査の過程で Basic 認証動作機器やサービスがはたらいていると推定できたホストに対しアクセスを行った．ホストに対する能動的なアクセスは，対象への負荷を与える可能性を完全に否定できない．そのため，アクセス時のタイムアウト設定，並列アクセス数を，負荷が最小限になるよう設定した．また，本研究は十分なセキュリティ対策がとられていない可能性のある Basic 認証動作機器やサービスを調査し推定するものであり，その内容が脆弱な機器への攻撃を助長する可能性を完全に否定できない．そのため，推定した機器の情報は詳細に記述していない．一方，研究者に対しては情報提供依頼に応じて個別に調査結果を提供することで，研究の恩恵の最大化を目指す．これまで明らかでなかった，どのような Basic 認証動作機器やサービスが存在しているかを世界規模で調査した結果を本論文により示すことは，IoT 機器等のセキュリティ向上に貢献するものであり，その恩恵は十分に大きいと考える．

8. 考察

Basic realm の取得方法と文字コード

本研究では．機器推定の入力として Basic realm を用いた．Basic realm については zgrab でアクセスを行い，レスポンスを取得することで入手したが，一部のホストについてはレスポンスを入手することができなかった．Censys が対象ホストに大規模スキャンしたタイミングと，こちらがアクセスを行い調査をするタイミングには，ずれが生じてしまう．そのタイミングの差で，外部からのアクセスを受け付けなくなったことや，対象ポートが閉じられた可能性が考えられる．さらに，zgrab でのアクセスでは，対象ホストに負荷をかけないようタイムアウトを短く設定した．これらが取得できなった原因の 1 つであると考える．また，入手した Basic realm の一部には，文字化けしたものも確認できた．別の文字コードで復元できるものもあり，それらはハングル文字やアラビア文字など，英数字以外の文字列であった．記号が含まれていたり，こちらが試した文字コードでは復元できない Basic realm もあり，それらが機器推定に繋がる情報を含んでいるかの調査は今後の課題である．

Censys の利用について

本研究では，Basic 認証が動作するホストの調査のため Censys を利用した．Censys は API で利用可能で，検索クエリで対象を指定することにより高速かつ大規模にホスト情報を取得できるという利点がある．一方，Censys の API には使用量制限があり，各ホストの応答内容詳細まで取得をすると，使用量制限の範囲内で大規模調査を行うことはできない．そのため本研究では，Censys ではホストの概要情報のみを取得し，詳細については zgrab でアクセスすることにより対応した．

9. まとめ

本研究では，Basic 認証が HTTP/1.0 で稼働している 48 万件以上のホストを対象とし，Basic realm を入力として使用することで，大規模な機器推定を行った．その結果，Censys から入手したデータからは 262 種類のメーカの 934 種類の機器が推定できた．今回は対象が HTTP/1.0 のみであるが，900 種類以上の機器を推定できたため，HTTP/1.1 に拡大するとより多くの Basic 認証稼働機器が推定できると期待できる．

謝辞　本研究は，国立研究開発法人情報通信研究機構 (NICT) の委託研究（JPJ012368C08101）により得られた成果を含む．

参考文献

[1] Reschke, J.: The 'Basic' HTTP Authentication Sch., Internet Engineering Task Force (IETF) (2015).

[2] Shodan: Shodan Search Engine, Shodan (online), available from ⟨https://www.shodan.io/⟩ (accessed 2023-10-23).

[3] Censys: Censys Search, Censys (online), available from ⟨https://search.censys.io/⟩ (accessed 2023-10-26).

[4] R. Fielding, M. Nottingham, J. R.: HTTP Semantics, Internet Engineering Task Force (IETF) (2022).

[5] 大塚瑠莉，吉岡克成：IoT 家電ハニーポットを用いた IoT 家電の Basic 認証に対する攻撃分析，暗号と情報セキュリティシンポジウム（SCIS）(2023).

[6] 大塚瑠莉， 九鬼琉, 吉岡克成：Basic 認証が動作する機器へのサイバー攻撃の観測，情報処理学会論文誌， Vol. 65, No. 9 (2024).

[7] Arwa Abdulkarim Al Alsadi, Kaichi Sameshima, J. B. K. Y. M. L. M. v. E. C. H. G. n.: No Spring Chicken: Quantifying the Lifespan of Exploits in IoT Malware Using Static and Dynamic Analysis, Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security, p. 309–321 (2022).

[8] NOTICE: NOTICE, NOTICE (online), available from ⟨https://notice.go.jp/⟩ (accessed 2024-8-19).

[9] 大塚瑠莉，Pa, Y. M. P.，吉岡克成：Basic 認証要求応答に着目した機器推定への ChatGPT search の適用，暗号と情報セキュリティシンポジウム（SCIS）(2025).

[10] the University of Michigan: zgrab2, , available from ⟨https://github.com/zmap/zgrab2⟩ (accessed 2025-08-15).

[11] OpenAI: OpenAI, OpenAI (online), available from ⟨https://openai.com/⟩ (accessed 2024-11-22).

[12] OpenAI: ChatGPT search, OpenAI (online), available from ⟨https://openai.com/ja-JP/index/introducing-chatgpt-search/⟩ (accessed 2025-5-6).

[13] Kotak, J. and Elovici, Y.: IoT Device Identification Using Deep Learning, 13th International Conference on Computational Intelligence in Security for Information Systems (CISIS 2020), pp. 76–86 (2021).

[14] Bar Meyuhas, Anat Bremler-Barr, T. S.: IoT Device Labeling Using Large Language Models, arXiv:2403.01586 (2024).

[15] Mirian, A., Ma, Z., Adrian, D., Tischer, M., Chuenchujit, T., Yardley, T., Berthier, R., Mason, J., Durumeric, Z., Halderman, J. A. and Bailey, M.: An Internet-wide view of ICS devices, 2016 14th Annual Conference on Privacy, Security and Trust (PST), pp. 96–103 (2016).