

# DNS クエリログを活用した国籍判定手法による 多言語デジタルサイネージシステムの提案

江口 直輝<sup>1</sup> 崔 赫秦<sup>2</sup> 中村 優吾<sup>2</sup> 福嶋 政期<sup>2</sup> 荒川 豊<sup>2</sup>

**概要：**観光地において訪日観光客が求める情報に合わせて観光案内をするなど、訪日観光客の属性を判別して効果的に情報を発信できるサービスの実現が期待されている。特に、デジタルサイネージなどの時間に応じて表示コンテンツを変えられる媒体では、イベントや観光情報を周知する手段である言語情報をその利用者に応じて変えることは重要であると考えられる。これにより多様な地域から集まる観光地において、訪日観光客の国籍ごとに言語を変えることができ、販売促進やイベント周知を効果的に行えるようになる。本研究では、空港や駅などに設置してある無料 Wi-Fi が提供する DNS サーバから取得できる DNS クエリログを用いてリアルタイムに国籍を判定する手法を提案する。具体的には、DNS クエリログに含まれる問い合わせドメインから、各国を識別するためのドメインである ccTLD や、RIR に基づいて国コードを算出し、それらを特徴量として分析することで実現した。日本人 6 人、中国人 6 人、その他国籍 4 人に対して評価実験を行った結果、15 秒間のクエリログを使用して 61% の精度で国籍判定ができることを明らかにした。

## Study of Multilingual Signage System by Nationality Determination Method Using DNS Query Log

NAOKI EGUCHI<sup>1</sup> HYUCKJIN CHOI<sup>2</sup> YUGO NAKAMURA<sup>2</sup> SHOGO FUKUSHIMA<sup>2</sup>  
YUTAKA ARAKAWA<sup>2</sup>

### 1. はじめに

新型コロナウイルス収束後のインバウンドを見据え、訪日観光客が求める情報に合わせて観光地案内の方法を工夫することが求められている。観光客の国籍・性別などのユーザ属性に合わせて公共機関や飲食店に設置してあるデジタルサイネージ広告を個人に最適化することで、観光客の回遊支援や観光地の効果的な発信が実現できるようになる。特に、外国籍の観光客に対する情報提供において、それを伝える手段である言語情報を多言語化することは最も重要である。観光庁においてもホームページや駅の券売機の多言語対応する取り組み [7] が行われていることから、ユーザ属性の中でも国籍情報を判別できる手法の検討が必要であると考えられる。

観光客の国籍を判定する手法として、カメラを用いてリアルタイムで画像分析を行い、国籍を判定することが考えられる。しかし、マスクをしている場合や遮蔽物の影響により、国籍の判定精度が低くなることが懸念される。さらに、インタラクティブ型のデジタルサイネージにおいて、ディスプレイを通じて観光客が言語を選択することも可能であるが Interactive Blindness[2] により、タッチ操作できることに気づかないことや、観光地においてスタンドアロン型のデジタルサイネージも多くあるため、多くのデジタルサイネージを活用できるようにすることが望ましい。

そこで本研究では、公共施設などで利用できる無料 Wi-Fi から取得できる Domain Name System(DNS) クエリログを活用した国籍判定を行う手法の実現を目指す。その判定結果を用いることで、無料 Wi-Fi 付近に設置しているデジタルサイネージの言語情報を変えることで効果的な情報発信が行うことができると考える。ただし、訪日観光客が Wi-Fi に接続してからしばらく同じ場所に滞在してい

<sup>1</sup> 九州大学大学院 システム情報科学府

〒 819-0395 福岡市西区元岡 744

<sup>2</sup> 九州大学大学院 システム情報科学研究院

〒 819-0395 福岡市西区元岡 744

るとは限らないため、リアルタイムで国籍を判定を行い、ディスプレイの言語情報を変更しなければならない。つまり、短時間で蓄積される少ない DNS クエリログから国籍を判定する手法を検討する必要がある。これまでの取り組みとして、1ヶ月分の DNS クエリログに含まれるクエリドメインの Country Code Top Level Domain(ccTLD) の数をユーザ毎に集計することで、国籍の判定を試みている [8]。しかし、問い合わせドメインには ccTLD だけではなく Generic Top Level Domain(gTLD) も多く含まれることから、Wi-Fi に接続してから短時間のデータで国籍を判定できるだけの ccTLD が含まれているか否か明らかになっていない。そこで、DNS クエリログに含まれる問い合わせドメインを IP アドレスに変換した後、Regional Internet Registry(RIR) に基づいて国コードを算出し、国籍判定を行う手法を提案する。RIR とはアドレス空間を管理している組織であり、地域毎に IP アドレスブロックが割り当てられている。これを用いることで全てのクエリドメインから国コードを算出することができるため、国籍判定精度を高めることが期待できる。

本研究により、短時間で国籍判定を行うことで表示言語をリアルタイムで切り替えられるデジタルサイネージが実現できることや、デジタルサイネージの前に複数の国籍の観光客がいる場合、言語ごとに表示時間を割り当てることが可能となる。現在、無料 Wi-Fi は空港や飲食店のみならず、観光地と自治体が連携することで設置箇所が拡大しており、それらから得られるデータを利活用することで、さらなるマーケティング戦略の立案や観光業の発展に寄与できると考える。

本稿では、2章において本研究における関連研究について記述し、3章では今回提案する多言語デジタルサイネージシステムの想定環境やシステム要件について記述し、4章では事前実験について記述し、5章では DNS クエリログから国籍判定する手法について記述し、6章において実験結果及び考察について記述し、最後に7章でまとめる。

## 2. 関連研究

通信インフラを活用したサービスの提供や、そこから得られるデータを利活用して国籍判定を行う取り組みは多数報告されている。

福岡県福岡市に設置している Fukuoka City Wi-Fi は来街者の利便性向上のためだけでなく、情報発信力の強化や災害時の活用にも用いており、利便性だけでなく、それを活用したマーケティング戦略や災害時のインフラの提供などにも活用している。具体的にはバナー機能を用いて観光スポットやイベントの周知で利便性を図ることや、Wi-Fi から取得できるデータを利活用することで、観光ルート造成の活用に充てている。このように通信インフラを活用したサービスの提供は既に実用化されており、観

光業において今後も期待されている。

文献 [5] ではディスプレイに近づいたユーザに対して、デジタルサイネージ上のコンテンツを個人に最適化できる手法を紹介している。独自のスマートフォンアプリをインストール後に自身の興味のあるコンテンツを登録して、ディスプレイに近づくことでリクエストを行う手法であるが、ユーザがスマートフォンにアプリをインストールしないと見えないため、観光客のように一度きりしか使わないディスプレイには適していないことが課題である。

文献 [8] では Domain Name System(DNS) クエリログのうち ccTLD(Country Code Top Level Domain) に着目した国籍判定手法を提案している。学内 Wi-Fi 経由で蓄積された1ヶ月分の DNS クエリログをユーザ毎に分類し、ccTLD のアクセス回数を集計した結果を数が多いものから降順で並べ替え、Doc2Vec[4] を用いて得られる特徴量からユーザ間の類似性を推定している。しかし、観光客が観光地に設置している Wi-Fi に接続するのは短時間であることが想定されるため、国籍判定を行うために十分なログの量が蓄積されていない可能性がある。したがって、今回想定する多言語デジタルサイネージシステムは先行研究のように大量のログをバッチ処理し国籍判定するのではなく、少ないログを分析することで国籍判定を行う必要がある、リアルタイム性が高い判定手法を検討する必要がある。

## 3. 多言語デジタルサイネージシステム

本章では提案する多言語デジタルサイネージシステムの想定環境とシステム要件について記述する。

### 3.1 想定環境

本節では今回提案する多言語デジタルサイネージの想定環境について説明する。ICT 総研によると近年、SNS の普及やスマートフォンのパケット利用料を低価格にするために、空港やバス停などで無料 Wi-Fi を利用する人は増加している [3]。またその利用者の属性を見ると、新型コロナ流行前の 2019 年度の全利用者数が 6342 万人であるのに対し、訪日外国人利用者数は 1687 万人であり全体の約 25% を占めている。このことから、無料 Wi-Fi は訪日観光客において簡単にインターネットに利用でき、利便性を向上させることから重要な役割を果たしている。そこで本研究では、駅や空港に設置された Wi-Fi が提供する DNS サーバに蓄積された DNS クエリログを活用し、その Wi-Fi に接続しているユーザの国籍をリアルタイムで判定するシステムを提案する。文献 [1] によると、デジタルサイネージの効果を高めるにはある程度のユーザが滞在する場所である必要であるが、駅や飲食店に設置された無料 Wi-Fi も同様の条件で利用されるため、広告効果を損なわずサービスを提供できる。DNS クエリログには問い合わせ IP アドレス、タイムスタンプ、問い合わせドメインなどが含まれる

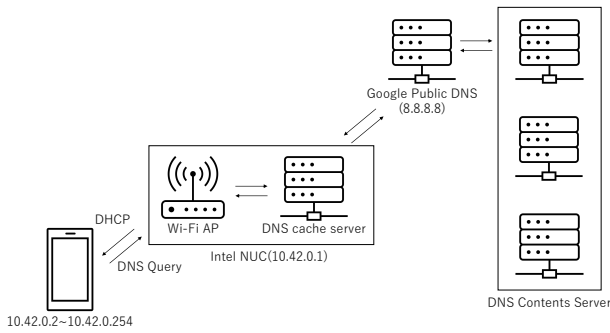


図 1: システム構成

ため、この情報を基に判定することができると考えた。そして、Wi-Fi 付近に設置してあるデジタルサイネージの言語情報を変えることで、観光情報やイベントを周知する。このように、DNS クエリログを利活用することで訪日観光客に対して、最適なサービスを還元することができるようになる。

### 3.2 システム要件

本節ではシステム要件について記述する。本システムの目標は、ユーザが無料 Wi-Fi に接続すると、ユーザの国籍をリアルタイムで判定して、ディスプレイの言語情報を変えることである。そこで以下 2 つの要件を定義する。

- **リアルタイム性**

ユーザは駅や空港に設置している無料 Wi-Fi をしばらく同じ場所に滞在しながら接続するとは限らず、すぐに別の場所に移動する可能性がある。そのため、ユーザが Wi-Fi に接続すると即座に付近に設置しているデジタルサイネージの言語情報が変わるようにする必要がある。よって、本研究の目標として 15 秒という短時間の DNS クエリログで国籍判定を行うことを目指す。

- **多言語対応**

訪日観光客で多い国籍は台湾、韓国、中国、香港であり [6]、東アジアからの訪日が多いことから中国語での表記は必要である。また、世界で最も広く利用されている英語での表記も行うことで、それ以外の国からの訪日観光客の言語要件を満たす。そのため、本研究では日本人（日本語）、中国人（中国語）、その他（英語）の国籍を判定し、3カ国の言語を切り替えることができるようにする。

## 4. 事前実験

本章では、DNS クエリログを収集するために構築したシステム構成及び既存研究での課題について記述する。

### 4.1 システム構成

DNS クエリログを収集するためのシステム構成を図 1 に示す。本システムは DNS クエリログの収集及びそれを用

いた国籍判定精度の検証を目的としているため、デジタルサイネージのアプリケーションに判定結果を提供する API サーバは構築していない。Wi-Fi アクセスポイント兼 DNS キャッシュサーバとして、Intel NUC Mini PC Kit を使用した。Wi-Fi アクセスポイントは Ubuntu OS が提供しているインターネット共有機能を使用した。DNS サーバの IP アドレス (10.42.0.1) を配布するための DHCP サーバとして、CLI でインストールできるパッケージ isc-dhcp-serve を使用した。さらに本システムでは、Wi-Fi の通信速度を向上させるために、Intel NUC 内蔵の無線アンテナを使用するのではなく、USB で接続できる外付けアンテナとして tplink 社製の Archer T3U Plus を使用したため、DHCP サーバの設定において無線インターフェースを本アンテナに設定した。DNS キャッシュサーバは Internet Systems Consortium が開発している Bind9 を使用し、クライアントからの DNS クエリをファイルに保存するようにした。その際に保存されるクエリには問い合わせ IP アドレス、タイムスタンプ、問い合わせドメインが含まれる。今回想定しているシステムは、処理負荷をできる限り分散するために、エッジノードで分析するような構成で構築した。また、パブリック DNS サーバとして Google 社が提供する Google Public DNS (8.8.8.8) を使用し、構築した DNS キャッシュサーバからフォワーディングするようにした。このようにユーザは構築した Wi-Fi に接続することで DNS クエリログを記録しながらインターネットを利用できる。

### 4.2 既存手法の検証

DNS クエリログから国籍判定できるか否かを検証するために事前実験を行った。文献 [8] ではクエリドメインのうち ccTLD を用いて国籍判定を試みているが、1 ヶ月という多量のデータをもとに分析しており、ユーザ毎の ccTLD の出現数が国籍判定を行うには十分であったと考えられる。しかし今回想定する多言語デジタルサイネージシステムではユーザが Wi-Fi に接続してから 15 秒程度の短時間で国籍判定を行うことを目標としており、その時間内に ccTLD がどれほど出現するかは未知数である。さらに、ccTLD 以外にも地域毎に IP アドレスブロックが割り当てられていることから、RIR が管理している国コードに基づいて国籍判定する要素になると考えられる。これは、日本語のウェブサイトは日本に割り当てられた IP アドレスであり、中国語のウェブサイトは中国に割り振られた IP アドレスである傾向があることが考えられるためである。そこで ccTLD と RIR に基づく国コード算出に関して以下 2 つの検証を行う。

- **検証 1: 全 DNS クエリログのうち ccTLD はどれくらいの数が含まれるのか。**
- **検証 2: 全 DNS クエリログに対し RIR に基づく国コード算出を行った場合、ユーザ毎の各国コードの割**



図 2: ccTLD と RIR のデータ処理方法

### 合はどの程度か。

上記 2 つを検証するために、研究室の日本人 (JP) 7 人と留学生 8 人 (中国 (CN): 5 人, 韓国 (KR): 1 人, インドネシア (ID): 1 人, マダガスカル (MG): 1 人) を対象に DNS クエリログを収集する実験を行った。Wi-Fi 接続中はスマートフォンの操作を制限することなく、10 分間自由にウェブサーフィンや SNS を利用するように指示をした。その結果、実験対象者全員分の DNS クエリログを収集できていることを確認した。また、各検証において収集した DNS クエリログの処理方法を図 2 に示す。検証 1 においては、問い合わせドメインのうち TLD の部分のみ抽出し、これを全クエリに対し適用することでユーザごとの各 TLD の数を数える。検証 2 においては (1) 問い合わせドメインを python のライブラリで IP アドレスに変換した後、(2) Asia Pacific Network Information Centre (APNIC) が公開しているデータを基に国コードを算出した。また、APNIC のデータに存在しない IP アドレスが DNS クエリログに含まれていた場合、その問い合わせドメインに対しての国コードを None とした。

### 4.3 検証結果

全クエリログに対する Top Level Domain (TLD) の割合とクエリドメインを IP アドレスに変換した後、RIR に基づく国コードに変換した結果を図 3, 図 4, 図 5 に示す。縦軸は観測した ccTLD や国コードの数であり、横軸は被験者 (15 人) である。

- **検証 1:** 図 3 より、全クエリログに対して、com や net などの gTLD の割合がほとんどであり、jp や cn などの ccTLD の出現頻度が少ないこと明らかになった。さらに今回観測できた ccTLD である jp・cn・fr・ru・kr・uk・id・de のみを抽出しユーザごとにその割合を図 4 に示す。中国人のデータには中国の ccTLD である cn が出現し、日本人のデータには日本の ccTLD である jp が他の ccTLD よりも多く出現していることから、長時間 DNS クエリログがあった場合、判定要素に含めることで判定精度が向上できる見込みがあると考えられる。しかし、10 分間の全クエリログのうち ccTLD は

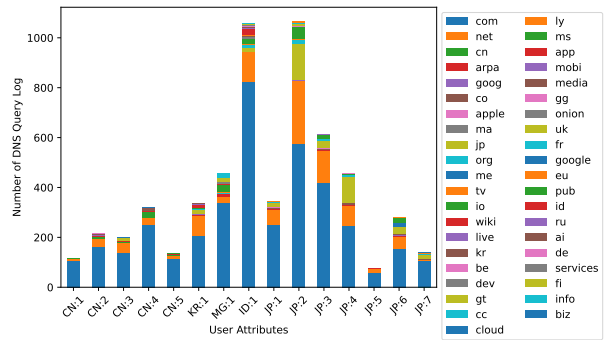


図 3: ユーザ毎の各 TLD の割合

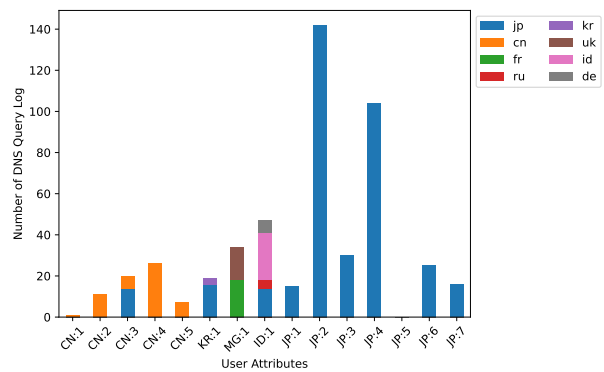


図 4: ユーザ毎の ccTLD の割合

約 12% しか含まれず、出現数・出現頻度も被験者ごとに偏りがあるため、常に安定した判定精度でリアルタイム性の高いシステムの実現は困難であると考えられる。

- **検証 2:** 図 5 より、全ユーザにおいて US の出現頻度が多いものの、中国人のログには CN が出現し、日本人のログには JP が出現する傾向があることが明らかになった。さらにユーザの国籍に応じて、特定の国の RIR にアクセスする傾向がある可能性もあり、少ないクエリログから国籍判定をするにはより国の情報が多い、RIR を用いる方が高い精度で国籍を判定できると考える。

以上の結果を踏まえると、ccTLD に関しては多量のデータが存在していた場合、その被験者の国籍の ccTLD が多く検出できる傾向があることから、多量なデータがあった場合、正確性は優位だが短時間となるとその精度は低下する可能性がある。一方で RIR に関しては全ての問い合わせドメインを国コードに変換することができるため、様々な地域の国コードが観測でき、少量のデータから特徴を抽出できる可能性があると考えられる。

## 5. 提案手法

本節では提案手法および実験方法について述べる。

### 5.1 提案手法: RIR の国コードを用いた判定

事前実験の結果から、リアルタイムで国籍を判定する際

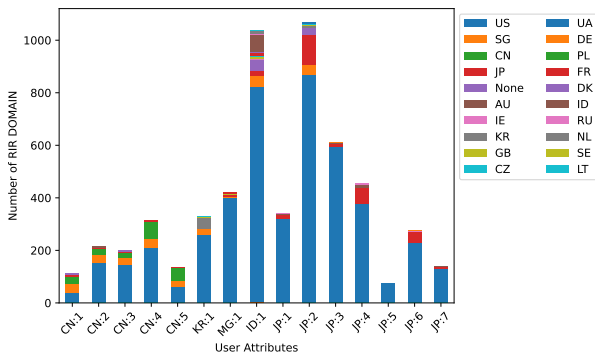


図 5: RIR に基づく国籍割合

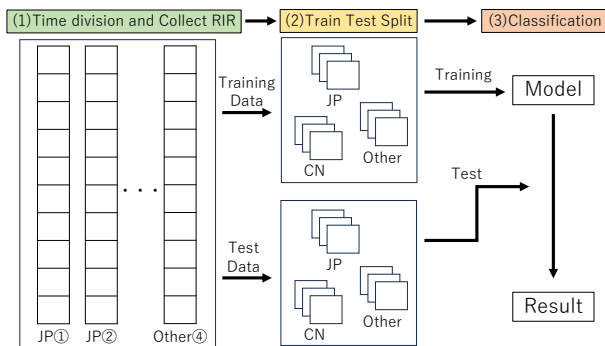


図 6: データ処理の流れ

表 1: 15 秒分の ccTLD と RIR を集計した例

	ccTLD			RIR				Label
	id	cn	jp	US	JP	CN	ID	
JP:1	0	0	2	3	10	0	0	JP
JP:1	0	0	5	4	11	0	0	JP

には RIR に基づいて国コードに変換することが有効である可能性があることがわかった。そこでデータの処理に関して以下 2 つの工夫を行った。

#### 工夫 1: ccTLD と RIR を特徴量とした多クラス分類

データの処理の流れを図 6 に示す。まず、被験者ごとの DNS クエリログを検証 1 や検証 2 の処理で行った図 2 と同様に ccTLD や RIR に基づく国コードに変換する。次に、各ユーザで時間幅  $W$  ごとにそれらの種類ごとの数を集計し、その集計結果に対し正解ラベル (JP・CN・Other) を割り振る。その例を表 1 に示す。また、集計する時間幅  $W$  を 15 秒から 80 秒までの範囲で、5 秒ごとにした。時間幅  $W$  における ccTLD や RIR に基づく国コードを集計した結果を正解ラベルごとに 7:3 でトレーニングデータとテストデータに分割した。またこの時間幅  $W$  に DNS クエリログが含まれないときは、国籍判定が不可能なことから分析するデータに含めなかった。そしてトレーニングデータを用いて LightGBM を用いてモデルを作成し、テストデータで推論を行う。

#### 工夫 2: 時系列を考慮した多数決による国籍判定

本手法は、手法 1 のように 15 秒分の DNS クエリログを

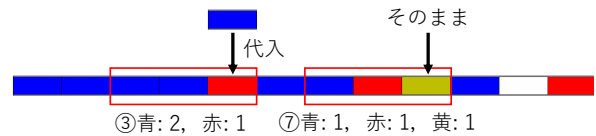


図 7: スライディングウィンドウを用いた多数決

分析した後、その予測結果を時系列で並べ、隣接する予測結果を基に多数決を行う手法である。実際のデジタルサイネージ運用上は判定結果が頻繁に変わってしまうとユーザに対し不快感を与えてしまう可能性がある。そのため、DNS クエリログが記録されていない時間帯や、予測結果が頻繁に変わってしまうものに対し、工夫を行う必要がある。そこで、予測した結果を先頭にウィンドウ幅を決め、スライディングウィンドウ方式のように一つずつずらして多数決をとる方式を提案する。

多数決を行うときの条件を以下に示す。また、このような条件で多数決を行った時の具体例を図 7 に示す。

- 予測ラベル (JP・CN・Other) に対してそのラベルを先頭に過去 3 つの予測データの中から多数決を行い、最も数が多い予測ラベルを先頭データに代入する (図 7 の 3 スライド目)。
- DNS クエリログが記録できなかった 15 秒 (None) は、それを含まない過去の 3 つの予測結果を用いて多数決を行う。
- 多数決の結果、同点の場合は、最新の予測データを代入する (図 7 の 7 スライド目)。

このようにユーザから取得できた直近のデータのみではなく、過去のデータも判定要素として含めることで精度が向上すると考えた。また分析の際は本手法では Wi-Fi に接続し始めてから 4 分間のデータをテストデータにし、それ以外のデータをトレーニングデータにした。本研究では Wi-Fi に接続し始めて 15 秒程度の短時間で国籍判定を行うことを目指しているため、本手法では 15 秒分の DNS クエリログを用いて予測を行い、その精度について評価する。

## 5.2 実験方法

研究室内の日本人 6 人、留学生 10 人 (中国人: 6 人, エジプト人: 2 人, 韓国人: 1 人, マダガスカル人: 1 人) を対象とし、構築したシステムの Wi-Fi に接続した。また、事前実験では Wi-Fi 接続中は自由にスマートフォンを操作してもらっていたが、今回の実験では実際の観光を再現するために、以下のような条件で行った。

#### 実験条件

- 現在地: 京都駅
- 時刻: 正午 12 時
- 課題: 午後 17 時までの観光スケジュールを決定する。
- 制限: Google Map の使用も可能だが、それだけを使

うのは禁止し、ブラウザでも調べる。

- 言語：母語もしくは母国語
- 時間：10 分程度

上記条件で観光調査を行ってもらい、被験者の DNS クエリログを収集した。また、実験終了後に作成したスケジュールや参考にしたウェブサイト名をアンケートで収集した。

## 6. 実験結果及び考察

本システムの有効性を示すために収集した DNS クエリログの分析を行った。以下、実験結果及び考察について記述する。

### 6.1 ccTLD と RIR を特徴量とした多クラス分類の結果

RIR に基づいた国籍判定の結果を図 8 に示す。横軸は何秒毎にユーザの DNS クエリログを分割し ccTLD と RIR に基づく国コードを集計したかを示し、縦軸は判定精度を示す。

まず、特徴量を ccTLD にした場合は、用いるデータの時間幅が最も短い 15 秒の精度は 44.0%であった。これは 15 秒分の DNS クエリログでは ccTLD の出現頻度が少ないため、精度が低くなったと考える。つまり ccTLD を特徴量とした場合、15 秒のような短時間で国籍判定を行うのは困難であることが明らかになった。しかし、用いる DNS クエリログのデータが多くなるほど、ccTLD の観測数も多くなるため判定精度が高くなる傾向になることが明らかになった。

RIR に基づく国コードの数を特徴量とした場合、短時間の DNS クエリログを用いた時の精度は ccTLD を特徴量とした時より精度が高くなる傾向にあることが明らかになった。しかし、55 秒分の DNS クエリログを用いた時を境に、用いる DNS クエリログを増やしても精度が横ばいになることがわかった。また、80 秒分の DNS クエリログを用いた場合は ccTLD の数が十分であったため、判定精度が RIR のみの特徴量としていた場合と比較して上回る精度となった。

RIR と ccTLD を特徴量とした場合は RIR のみの場合と同様に、短時間の DNS クエリログを用いた時の精度から 55 秒間の DNS クエリログを用いた時まで精度は増加傾向になることがわかった。RIR のみの特徴量とした場合と RIR と ccTLD を特徴量とした場合は、精度に大きく差はないことがわかる。

このような結果から、短時間で国籍判定を行うには、RIR に基づく国コードと ccTLD の数のどちらも特徴量にすることが望ましいことが明らかになった。しかし、ccTLD のみの特徴量とした場合は長時間の DNS クエリログを用いることで精度が向上していることから、今後それらの特徴量を組み合わせて精度向上を見込むことができると考える。

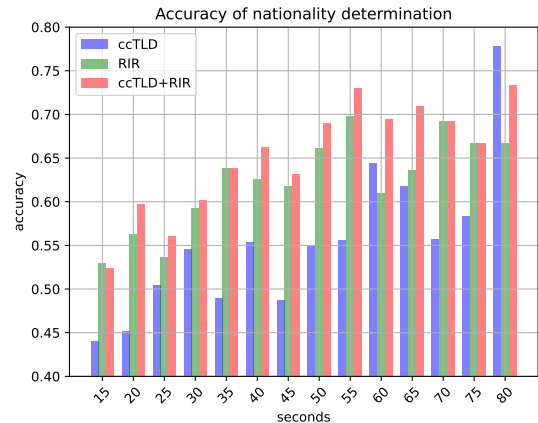


図 8: 時間幅  $W$  における精度比較

### 6.2 時系列を考慮した多数決による国籍判定の結果

判定結果の時系列変化の分析を行った。15 秒分の ccTLD と RIR の集計データを用いて最初の 4 分間分をテストデータとして、それ以外のデータをトレーニングデータとして、モデルを作成した後、国籍の予測を行った。

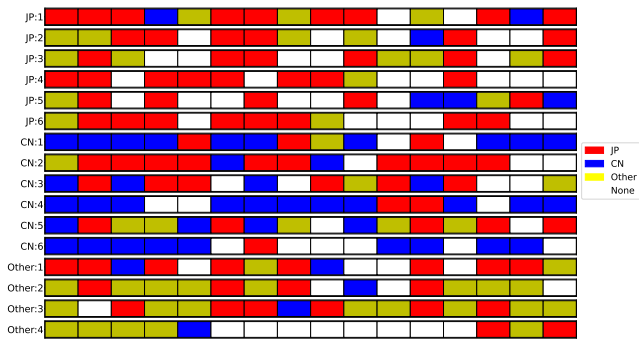
判定結果を時系列に並べたものを図 9a に、DNS クエリログが含まれていない時間 (None) を一つ前の予測結果にした時の結果を図 9b に示す。縦軸は被験者の国籍であり、それぞれ 15 秒の集計データを 4 分間分 (16 個) 並べている。図 9b より、判定結果が連続的に変化するため、実際にデジタルサイネージを運用する上では、ユーザ体験の低下が懸念されると考えられる。また、図 9b の結果を用いて、被験者の国籍に対してディスプレイに表示される言語が正しい時間の割合を算出すると、4 分間のうち 54% で正しい言語が表示される結果となった。

最後に、近接する 3 つの判定結果で多数決を取った結果を図 9c に示す。図 9b と比較して、被験者の国籍が正しく表示されるような傾向であり、かつ連続的にディスプレイに同じ言語が表示されるようになっていくことがわかる。また、アンケートの結果から、被験者 CN:2 において多言語対応の日本のサイトを利用して観光地調査を行っていたことがわかった。そのため中国語のウェブサイトを開覧しても、そのアクセス先が日本にある多言語サイトであった場合、判定精度が低下することが懸念される。図 9c の結果を用いて、ユーザの国籍に対してディスプレイに表示される言語が正しい時間の割合を算出すると、4 分間のうち 61% で正しい言語が表示される結果となった。

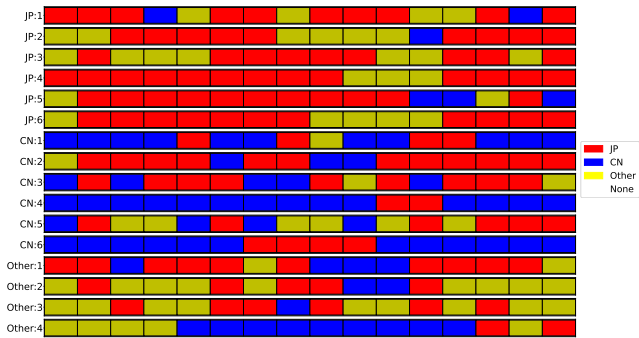
以上より、判定結果を単純に時系列で並べる方式より、近接する過去の判定結果を用いて多数決を取る方式の方が判定精度が高く、連続して同じ言語が表示されるような結果となった。

## 7. まとめ

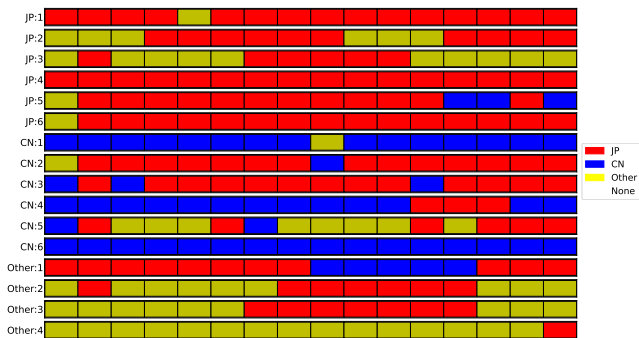
本稿では、端末の DNS クエリログを用いて ccTLD や



(a) None を含む表示言語の時間変化



(b) None に代入した時の表示言語の時間変化



(c) 提案手法の表示言語の時間変化

図 9: 表示言語の時系列変化

RIR に基づく国コードから国籍判定を行う手法を提案した。その結果、15 秒分のデータを時系列に並べ、過去の予測結果を加味することで判定精度が向上することが明らかになった。今後の課題として、判定結果を時系列に並べ多数決を取る際にそのウィンドウサイズを可変にすることや、多数決を行う条件をさらに工夫することで精度の向上が期待できると考える。また、国籍を判定するだけではなく、Wi-Fi 利用者が見ているウェブサイトから利用者の興味を推測しディスプレイ表示を変更するなどの工夫を行うことで、より効果的なデジタルサイネージシステムが実現できると考える。

## 参考文献

[1] Florian Alt and Julia Vehns. Opportunistic deployments: challenges and opportunities of conducting public display

research at an airport. *In Proceedings of the 5th ACM International Symposium on Pervasive Displays*. ACM, 106–117, 2016.

- [2] Steven Houben and Christian Weichel. Overcoming interaction blindness through curiosity objects. page 1539–1544, 2013.
- [3] ICT 総研. 2023 年度公衆無線 lan. <https://ictr.co.jp/report/20230411.html/>.
- [4] Quoc Le and Tomas Mikolov. Distributed representations of sentences and documents. In *International conference on machine learning*, pages 1188–1196. PMLR, 2014.
- [5] Callum Parker, Martin Tomitsch, Nigel Davies, Nina Valkanova, and Judy Kay. Foundations for designing public interactive displays that provide value to users. page 1–12, 2020.
- [6] 観光庁. 訪日外国人旅行者（観光・レジャー目的）の訪日回数と消費動向の関係について, 2019.
- [7] 総務省. 多言語対応改善・強化のための観光庁の取り組み. [https://www.soumu.go.jp/main\\_content/000677784.pdf](https://www.soumu.go.jp/main_content/000677784.pdf), 2020.
- [8] 矢野安希子, 野林大起, 塚本和也, and 池永全志. Dns クエリログを用いたアクセス先ドメインの類似性に基づくユーザ属性分析. *電子情報通信学会技術研究報告; 信学技報*, 121(433):43–46, 2022.