2022年度　卒業論文 2023/01/26 02:37

MACアドレスランダム化の検証

大阪産業大学 デザイン工学部 情報システム学科
情報教育システム研究室

19H009 板持秀輝

MACアドレスランダム化の検証

19H009 板持秀輝

1　はじめに

スマートフォンなどの無線LAN搭載端末 (以下、モバイルデバイスという) はWi-Fi通信を行う際、様々な情報をアクセスポイントとやり取りしている。その情報の中の1つにMACアドレスがあり、原則として世界中に一つのユニークな値に設定されているため、連続で計測することで個人の行動を特定することが可能である [1]。そのため、ユーザーからプライバシーの観点で疑問の声が上がり、メーカーはMACアドレスの値のランダム化を搭載させるようになった。MACアドレスがランダム化されても、個人と結びつけることができないか検証した。

2　目的

本研究の目的は、MACアドレスをランダム化することで同一の端末かどうかの特定が不可能になり、セキュリティ面で向上しているのかを解析することである。分析する方法は、スマートフォン、携帯用ゲーム機などのWi-Fi対応の無線LAN搭載機器が、アクセスポイントと接続する際の通信を計測する。計測した通信からMACアドレスを取り出し、MACアドレスがランダム化されている通信の中に、同一のモバイルデバイスから発信されたものがあるかどうかプログラムで分析する。

3　Probe Request

モバイルデバイスはアクセスポイントと接続する際にProbe Requestを発信する。Probe Requestには送信元のMACアドレスやシーケンスナンバーなどが含まれている。モバイルデバイスのOSや機種によって発信するMACアドレスやシーケンスナンバーの値のパターンや特徴に違いがあり、MACアドレスをいくつかに分類した。

4　同一推定の方法について

Probe Requestの計測機器を本学内の11箇所に設置し、一ヶ月もの期間計測を行った。計測した中からランダム型のMACアドレスのみを取り出し、シーケンスナンバーなどから他のMACアドレスと同一のものがないかを探し、数える。

5　結果と考察

測定したMACアドレスをユニーク型とランダム型に分類した結果、ユニーク型の方が多いことがわかった。ランダム型のMACアドレスの中で一日平均800個程のMACアドレスが他のMACアドレスと同一の端末から発信されたものであると推定できた。計測範囲を改善するか、プログラムの精度を上げることでより多くのランダム型のMACアドレスの同一推定が可能になると考えられる。

6　まとめ

本研究の目的は、MACアドレスをランダム化することで同一の端末かどうかの特定が不可能になり、セキュリティ面で向上しているのかを解析することである。MACアドレスがランダム化されていても、同一推定を行うことは成功したが、課題点がある。計測範囲が狭く、モバイルデバイスが少しでも計測機器から離れると次計測する際のシーケンスナンバーに開きができてしまい、同一かどうかの判定がっできなくなってしまう。今後の課題としては、計測面積の拡大、他の情報も用いてより細かく解析する、と考える。

参考文献

[1] 望月祐洋, 上善恒雄, 西田純二, 中野秀男, 西尾信彦 ほか. Wi-fi パケットセンサを利用した匿名人流解析システムの構築. 研究報告ユビキタスコンピューティングシステム (UBI).

目次

1　はじめに

近年、無線 LAN（Local Area Network）の一つである Wi-Fi を利用できるスマートフォンが急激に普及している。スマートフォンが普及するにつれ、無線 LAN サービスの利用者も増加しており、国内で 2018 年時点で 5746 万人 [1] となっている。公衆無線 LAN 利用者数の推移を**図 1** に示す。利便性を高めるためにも、行政では無線 LAN を増強する取り組みを行っており、今後無線 LAN の利用者はさらに増加する見込みである。

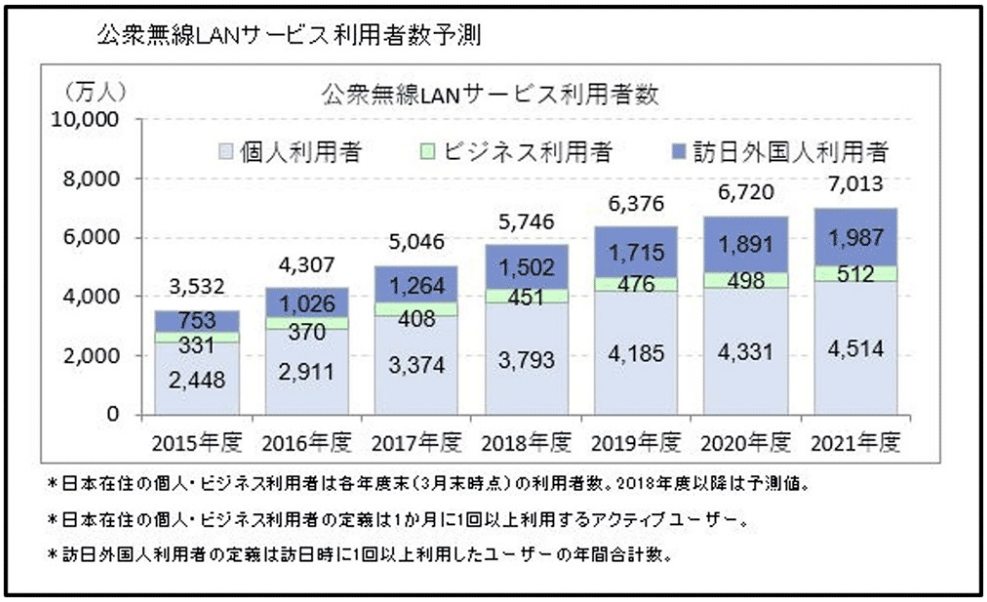

図 1　ICT 総研の調べによる国内での公衆無線 LAN 利用者数と予測。縦軸は利用者数 (万人)、横軸は年度。グラフの青色の箇所は訪日外国人、薄緑色の箇所はビジネス利用者、空色の箇所は個人利用者の公衆無線 LAN 利用者数を表している。いずれにおいても公衆無線 LAN 利用者数は上昇する傾向にある。個人・ビジネス利用者は各年度末 (三月末時点) の利用者数で、個人・ビジネス利用者の定義は 1 か月に 1 回以上利用するアクティブユーザーである。訪日外国人利用者の定義は訪日時に 1 回以上利用したユーザーの年間合計数。2018 年度以降は予測値である。

スマートフォンなどの無線 LAN 搭載端末 (以下、モバイルデバイスとする) は、Wi-Fi ネットワークを検索する場合にアクセスポイント[*1]に対して問い合わせをするために Probe Request[*2]という信号を飛ばす。Probe Request の中に含まれる MAC アドレス[*3]は本来は意味を持たない値ではあるが、他の情報を紐づけることで利用者の個人の行動を調べることも可能で駅構内や遊園地、人気飲食店などの人が集中する場所での利用者の混雑度や流動をリアルタイムで分析することが可能ある。そのため、MAC アドレスを収集することでユーザーの行動を監視するようなシステム [2] も作られ、ユーザーからプライバシーの観点で疑問の声が上がっていた。そこで、Apple が 2012 年に自社の OS を用いたスマートフォンで、アクセスポイントと通信する際に MAC アドレスのランダム化を行うようになった。MAC アドレスをランダム化することで MAC アドレスとデバイスの結びつけが不可能になったため、プライバシーの保護を達成できると思われたが、MAC アドレスのランダム化による弊害として MAC アドレスを用いたフィルタリングや認証や MAC アドレスをベースとしたシステムが旧来と同じように機能しなくなってしまった。また、Probe Request は MAC アドレス以外にも様々な情報を持っているため、発

[*1] アクセスポイントとは、スマートフォンなどをネットワークに接続する機器である。
[*2] Probe Request については 3 章で解説する。
[*3] MAC アドレスについては第 3.1 章で解説する。

信元のモバイルデバイスを結び付けられるのではないかと考えた。

第 2 章では本研究の目的について述べる。第 3 章では Probe Request について述べる。第 4 章ではランダム型の MAC アドレスの同一推定方法について述べる。第 5 章ではプログラムで実行した結果と考察について述べる。第 6 章では研究の成果とともに今後の課題についてまとめる。

2　目的

本研究の目的は、MAC アドレスをランダム化することで同一の端末かどうかの特定が不可能になり、セキュリティ面で向上しているのかを解析することである。分析する方法は、スマートフォン、携帯用ゲーム機などの Wi-Fi 対応の無線 LAN 搭載機器（以下、モバイルデバイスという）が、アクセスポイントと接続する際の通信を計測する。計測した通信から MAC アドレスを取り出し、MAC アドレスがランダム化されている通信の中に、同一のモバイルデバイスから発信されたものがあるかどうかプログラムで分析する。

3 Probe Request

Probe Request はプローブ要求とも呼び、モバイルデバイスが周囲のアクセスポイントと接続する際に（または接続している状態でも） Wi-Fi 通信の標準規格である IEEE802.11 で定義された Probe Request を送信する。SSID で指定された特定のアクセスポイントから情報を要求するクライアントステーションによって送信される特別なフレームである。モバイルデバイスは Wi-Fi 機能を有効にしている場合、数秒から数分間隔で自動的に送信される。送信間隔は、Windows,iOS,Android といった OS や機種のメーカーでそれぞれ異なり、Wi-Fi 接続時には間隔が大きくなり Wi-Fi 未接続の状態だと小さくなる。

Probe Request には送信端末の MAC アドレス、シーケンスナンバー[*4]などが含まれる。無線 LAN のフレームフォーマットは、PHY[*5]フレームの中の MAC フレームのヘッダ内にフレームコントロールフィールドがある。フレームコントロールフィールドでは 2byte のデータが格納されており、その組み合わせで MAC フレームの種類が変わる。無線 LAN のフレームコントロールフィールドにはマネージメント（管理）フレーム、コントロール（制御）フレーム、データフレームの三種類が存在する。その中の 1 つであるマネージメントフレームの中に Probe Request は存在する。フレームコントロール内の Frame Type の値が [00][*6]かつ、Frame Subtype の値が [0100][*7]であれば、MAC フレームはマネージメントフレームの PR になる。無線 LAN のフレームフォーマットの簡易版を**図 2** に示す。

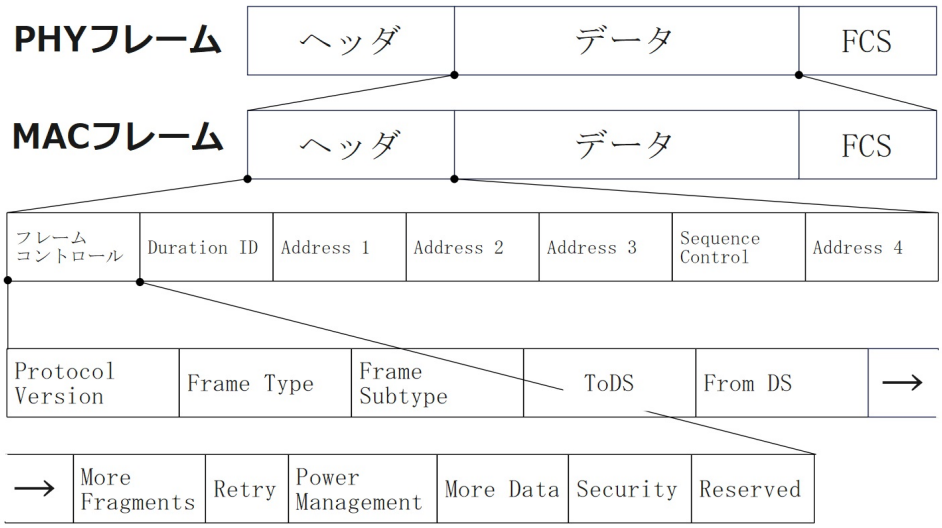

図 2　MAC フレームのヘッダ内部における各フィールドの名称及びその内訳を表している。MAC フレームのヘッダ内にあるフレームコントロールフィールド。このフィールド内の Frame Type が [00] かつ Frame Subtype が [0100] であれば、PR になる。

[*4] パケットに付与される通し番号。
[*5] Physical の略である。
[*6] 00: マネージメント、01: コントロール、10: データの値が入る。
[*7] このフィールドには 4bit の値が入る。

3.1 MAC アドレスの種類

MAC アドレスはパソコンやルータなどのネットワーク機器に付いている 12 桁の 16 進数の番号で、全世界で MAC アドレスが重複することがない値である。本研究では、ネットワークカードに割り当てられ世界で一意であるユニークな MAC アドレスを用いた「ユニーク型」と、アクセスポイントと通信する毎に値が変わるランダムな MAC アドレスを用いた「ランダム型」の二種類に分類した。また、ランダム型をさらに「ベンダー固定ランダム型」「MAC アドレスランダム型」「完全ランダム型」の三種類の型に分類する。本研究では、MAC アドレスのベンダーコード[*8]は 24bit の 16 進数で表すのではなく、「Google」や「Apple」などのベンダー名で表記される。

3.1.1 ユニーク型

ベンダーからユニークな値が割り当てられた「ユニーク型」。全てのネットワーク機器に割り当てられている番号で、原則として重複することがない。この値は各端末のデバイス情報画面で確認できる。しかし、MAC アドレスのランダム化によって、本来の MAC アドレスとして用いられることが無くなってきている。ユニーク型の Probe Request を送信する機種は全て Android であり、iOS の機種では扱われることがない。シーケンスナンバーは通信した数と共に増加していく。ユニーク型のモバイルデバイスから実際に発信された Probe Request の例を**表 1** に示す。

表 1　ユニーク型

Time	MAC	シーケンス番号
8:00:00	RuckusWi_3a:90:a8	2167
8:03:00	RuckusWi_3a:90:a8	2184
8:03:00	RuckusWi_3a:90:a8	2185
8:04:00	RuckusWi_3a:90:a8	2209
8:06:00	RuckusWi_3a:90:a8	2220

[*8] MAC アドレスの上位 24bit のことを言い、値から製造元のメーカーを特定できる。

3.1.2　ランダム型

Probe Request が条件によってアクセスポイントと通信する毎に変化する。ランダム型を Probe Request のパケットの内容からさらに分けて、「ベンダー固定ランダム型」「MAC アドレスランダム型」「完全ランダム型」の三種類の型に分類する。MAC アドレスランダム型と完全ランダム型では、ベンダーコード部分もランダム化されており、第一オクテットの 7 8bit 目が 1 0 となり、MAC アドレスの二文字目の値は必然的に 2,6,A,E になる。ベンダーコードでは、第一オクテットが下 2bit が 0 0 となっているため、ベンダーコード部分がランダム化されても特定のベンダーコードと被る心配はない。そのユニーク型とランダム型の第一オクテットの違いと MAC アドレスの二文字目の値を**表 2** で示す。

表 2　ランダム型の性質
表の一番右の列は MAC アドレスを 12 桁の 16 進数で見た際の二文字目である。

種類	第一オクテット	二文字目
ユニーク型	xxxxxx00	0,4,8,C
ランダム型	xxxxxx10	2,6,A,E
アクセスポイント	xxxxxx01	1,5,9,D
テザリング	xxxxxx11	3,7,11,F

3.1.3　ベンダー固定ランダム型

上 24bit がベンダコードで下 24bit がアクセスポイントと通信するごとに値が変化する「ベンダー固定ランダム型」。ベンダー固定ランダム型のモバイルデバイスから実際に発信された Probe Request の例を**表 3** に示す。

表 3　ベンダー固定ランダム型

Time	MAC	シーケンス番号
10:14:53	Google_8c:d0:bb	130
10:14:54	Google_8c:3a:25	133
10:14:54	Google_ab:26:da	135
10:14:54	Google_55:d0:84	136
10:14:54	Google_67:8f:90	137

3.1.4　MAC アドレスランダム型

MAC アドレスの 48bit 全てがアクセスポイントと通信するごとに値が変化する「MAC アドレスランダム型」。MAC アドレスランダム型のモバイルデバイスから実際に発信された Probe Request の例を**表 4** に示す。

表 4　MAC アドレスランダム型

Time	MAC	シーケンス番号
11:28:48	5a:b3:3e:01:80:b4	808
11:28:53	ca:c0:26:64:36:3c	844
11:29:01	66:fa:4b:64:fd:0e	900
11:29:09	c6:8c:ff:1c:fa:dd	967
11:29:13	ea:62:60:66:eb:e6	1035

3.1.5 完全ランダム型

MAC アドレス、シーケンスナンバー、送信間隔がアクセスポイントと通信するごとに値が変化する「完全ランダム型」。2016 年以降のスマートフォンで用いられており、iPhone7 以降の全ての iOS 端末で導入されている。Android 端末では確認できなかった。完全ランダム型のモバイルデバイスから実際に発信された Probe Request の例を**図 5** に示す。

表 5　完全ランダム型

Time	MAC	シーケンス番号
8:00:00	25:df:e4:54:68:43	20
8:00:30	45:8a:5b:19:30:5b	298
8:01:00	8d:78:93:bc:c5:77	908
8:01:30	23:46:7b:0f:8d:32	58
8:02:00	ac:82:67:8f:90	1900

3.2 OS 毎のランダム化の仕様

表 6 OS によるランダム化実装時期の違い

OS	実装年	バージョン
Android	2017 年	Android 8.0
iOS	2014 年	iOS 8
Windows	2015 年	Windows 10

Android では 8.0 から Probe Request 時の MAC アドレスのランダム化が実装された。このランダム化では、probe Request 時の MAC アドレスのランダム化であった。Android 8.0 9 ではランダム化機能はデフォルトでオフになっており、開発者モードから変更する必要があった。2019 年にリリースされた Android 10 からランダム化機能がデフォルトでオンになった。Android 10 では接続する SSID によって MAC アドレスが変化し、基地局に接続後は MAC アドレスが変化しない仕様になっている。端末を再起動したり、一度プロファイルを削除してユーザ名を変更しても同じ SSID なら同じ MAC アドレスになる。ランダム化される際はベンダーコードもランダム化される。Android 12 以降で一部のネットワークに対して使用される非永続的ランダム化では、ネットワークから接続が切れて DHCP リース期間が満了後 4 時間を超える場合か、ネットワークプロファイルに対してランダム化された MAC アドレスが生成されてから 24 時間を超えている場合にも再度 MAC アドレスがランダム化される。

iOS では、iOS 8 で Probe Request 時のみ MAC アドレスがランダム化する初期のランダム化が取り入れられた。iOS ではインタフェイスごとに一意のアドレスが個別に生成される。2016 年の iOS 10 搭載の iphone 7 から MAC アドレスだけでなくシーケンスナンバーもランダム化される仕様になった。iOS では 2020 年リリースの iOS 14 から SSID によって MAC アドレスが変化し、基地局に接続後は MAC アドレスが変化しない仕様になった。端末を再起動したり、一度プロファイルを削除してユーザ名を変更しても同じ SSID なら同じ MAC アドレスになる。ランダム化される際はベンダーコードもランダム化される。ネットワークプロファイルに対してランダム化された MAC アドレスが生成されてから 24 時間を超えている場合にも再度 MAC アドレスがランダム化される。

Windows では、Windows 10 から Probe Request 時のみ MAC アドレスがランダム化する初期のランダム化が取り入れられた。また、MAC アドレスのランダム化はデフォルトでオンになっている。

本研究で用いた Probe Request のデータは 2018 年のデータ [3] で、iOS の最新バージョンは iOS 12 が用いられており主な最新機種は iPhone 8,X が挙げられる。Android の最新バージョンは Android 9.0 が用いられており、Windows の最新バージョンは Windows 10 である。

4　同一推定の方法について

ここでは MAC アドレスの収集方法について解説したのち、MAC アドレスの分類方法、ランダム型の MAC アドレスの同一推定方法について解説する。

4.1　MAC アドレスの収集方法について

本研究では本学内における学生の移動を把握するために、モバイルデバイスが発する Probe Request を計測し、学生の移動を把握する手法で 2018 年に集めたデータ [3] を用いる。Probe Request には MAC アドレスが含まれている。一人につきモバイルデバイスを一台持っていると仮定した場合、MAC アドレス=モバイルデバイス=個人が成り立つ。本研究では、MAC アドレスがランダム化しているモバイルデバイスから発信された Probe Request から同一端末の推定をしていくものである。本研究では、PR を計測するための機器を作成し、計測期間は 4 週間で時間は 8:00〜18:00 の間で、本学内の 11 箇所で人通りが多いと予想した場所に設置した。機器の計測範囲は廊下で障害物がない場合は水平方向に約 38 メートル計測できる。2 枚のガラス扉が間にある場合は水平方向にに屋外に設置したモバイルデバイスを約 18 メートルで計測できた。

4.2　設置場所

設置場所と名前を**表 7** と**図 3** にそれぞれ示す。

表 7　計測機器設置場所の名前の一覧である。本学内の敷地は東高野街道で区切られている。本学はその西側の敷地を中央キャンパス、東側の敷地を東キャンパスとそれぞれ呼称している。

	設置場所の名前
中央キャンパス	5 号館
	7 号館
	9 号館
	16 号館
	図書館
	本館
東キャンパス	3 号館
	4 号館
	8 号館
	15 号館
	クリスタルテラス

基本的に設置場所は生徒の手が届きづらく、また設置場所で業務を行う人の監視の目が届きやすい学科事務室に設置した。学科事務室以外では、図書館内の受付カウンターの内側などに設置した。設置した機器から通行人の Probe Request を計測し、Probe Request の中に含まれる MAC アドレスを収集する。

図 3　Probe Request を計測した機器を設置した場所がどこなのかを示した簡易マップである。図中で用いられる RP と数字は設置場所を表しており、11 機設置した。RP の付いていない数字は本学の建物の番号を示している。図中右上にある学食はクリスタルテラスという名前である。P はパーキング（駐輪場）である。Wellness は本学の総合フィットネス健康施設である。図中の東高野街道から左側が中央キャンパスである。中央キャンパスから用水路を跨いで下側の Wellness があるところが、南キャンパスである。東高野街道から右側が東キャンパスである。

4.3　計測機器

計測機器の仕様を**表 8** で示す。実際の計測機器の写真を**図 4** で示す。

表 8　計測機器仕様一覧

コマンドライン	tshark ver 2.2.6
ハードウェア	Raspberry Pi 3 Model B
無線 LAN アダプタ	WLI-UC-AG300N
OS	Ubuntu MATE 16.04.4 LTS(Xenial Xerus)
SD カード容量	64GByte
その他付属パーツ	RTC

図 4　計測機器の写真。中央に写っているプラスチックケースは Raspberry Pi である。黒いコンセントに差し込める形状のものが Raspberry Pi の AC アダプタである。黒い細長い形状のものが USB で接続できる無線 LAN アダプタである。ケーブルは USB から映像の出力ができ、HDMI から映像の入力ができるケーブルである。

4.4 Probe Request から取り出し

pcap[*9]データを tshark コマンドを用いて txt データで出力した。その txt ファイルからタイムスタンプ[*10]、MAC アドレス、シーケンスナンバーの部分のみを取り出す。フレームからの取り出しについて**図 5** で示す。

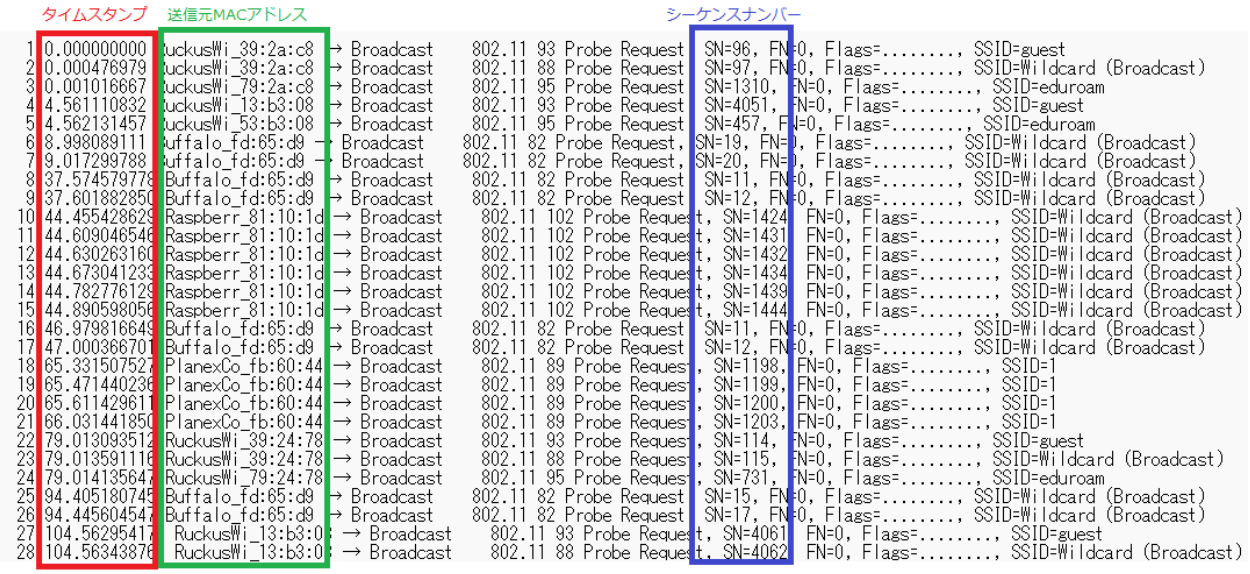

図 5　フレーム内で取り出す部分について。色付きの枠で括った部分を取り出す。赤枠が計測を開始してから何秒経ったかを示すタイムスタンプ。緑枠がモバイルデバイスが Probe Request を送信する際に使用した MAC アドレス。青枠がアクセスポイントとやり取りする際に用いるシーケンスナンバーで「SN=」の後の部分である。

4.5 MAC アドレスの分類方法

ユニーク型は二回以上同じ MAC アドレスを計測して、第一オクテット目の 7,8bit 目が 0 であるものを選んだ。本研究で用いた MAC アドレスでは、ベンダーコードが本来の 16 進数で表されずベンダー名で表される MAC アドレスもあるので、二回以上計測していればユニーク型として判断した。

ランダム型は、他と被りがなく一度だけの計測で、第一オクテット目の 7bit 目が 1 で 8bit 目が 0 であるも MAC アドレスを選んだ。第一オクテット目の 7bit 目が 1 で 8bit 目が 0 である MAC アドレスとベンダーコードの部分がベンダー名で表される MAC アドレスでも、一度のみの計測ならランダム型として判断した。ユニーク型とランダム型の MAC アドレスを分けるためのプログラムのフローチャートを**図 6** で示す。

なお、一部のモバイルデバイスでは、ルーターと認証していない Probe Request の値がランダム型と判断できる MAC アドレス（第一オクテット目の 7bit 目が 1 で 8bit 目が 0 である MAC アドレス）でも同じ MAC アドレスの値のまま通信しており、ルーターと認証して初めて値がランダム化することが確認できた。この MAC アドレスは値が変化しないため、ランダム型とは判断せずユニーク型として判断した。

[*9] pcap とはネットワーク上に流れるパケットを、キャプチャするためのファイル形式である。
[*10] タイムスタンプとは時刻情報のことである

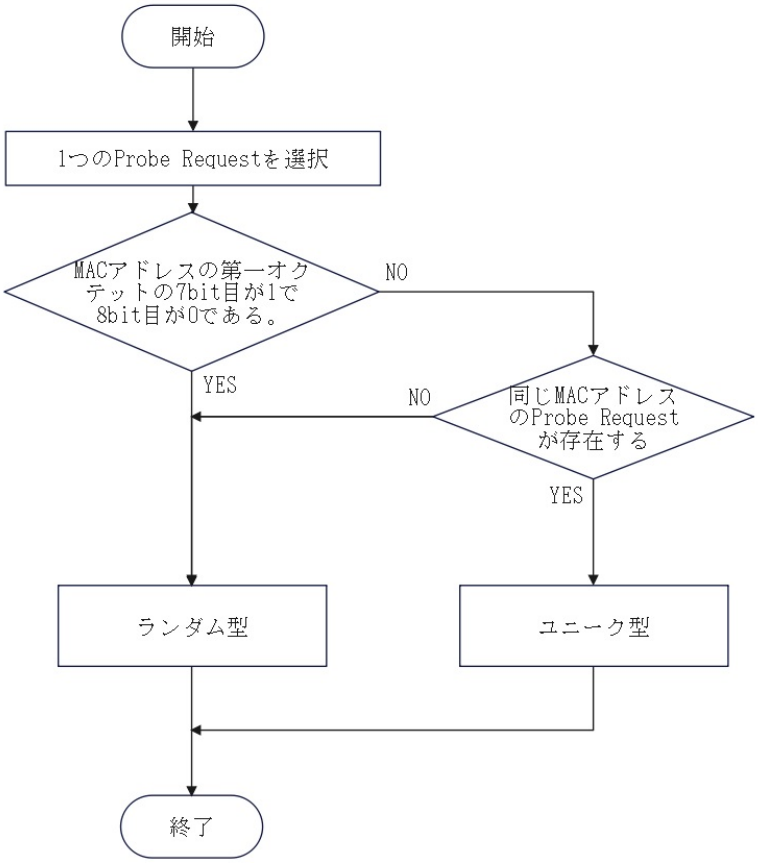

図６　ランダム型とユニーク型に分類のフローチャート

4.6　同一推定方法

Probe Request を発信した端末の同一判定をシーケンスナンバーやベンダーコードからプログラムで分析する。全体の MAC アドレスをユニーク型・ランダム型に分けたのち、ランダム型の MAC アドレスのみを用いて同一推定を行う。ランダム型の MAC アドレスを一つ選びその MAC アドレスのタイムスタンプから 180 秒以内の MAC アドレスの中でシーケンスナンバーの値が 100 以内の MAC アドレスを候補としてリスト化する。シーケンスナンバーの値を 100 以内と大きい値に設定した理由は、モバイルデバイスが計測範囲から少しでも離れると次のシーケンスナンバーの値に開きができてしまうからである。最初に選んだ MAC アドレスがベンダー固定ランダム型なら同じベンダーコードのものを候補内から選ぶ。最初に選んだ MAC アドレスが MAC アドレスランダム型なら候補内から計測場所が同じで上位 24bit もランダムで生成されているものを選ぶ。候補内から複数見つかった場合はよりシーケンスナンバーがより近いものを選ぶ。シーケンスナンバーが同じだった場合はタイムスタンプがより近いものを選ぶ。同一推定ができた MAC アドレスの数はカウントしておく。もし候補が 1 つもない、もしくは候補内から見つからなければ完全ランダム型とする。完全ランダム型は MAC アドレスだけでなくシーケンスナンバーもランダム化されているので同一推定は難しいため、本研究では取り扱わない。このプログラムのフローチャートを**図 7** で示す。

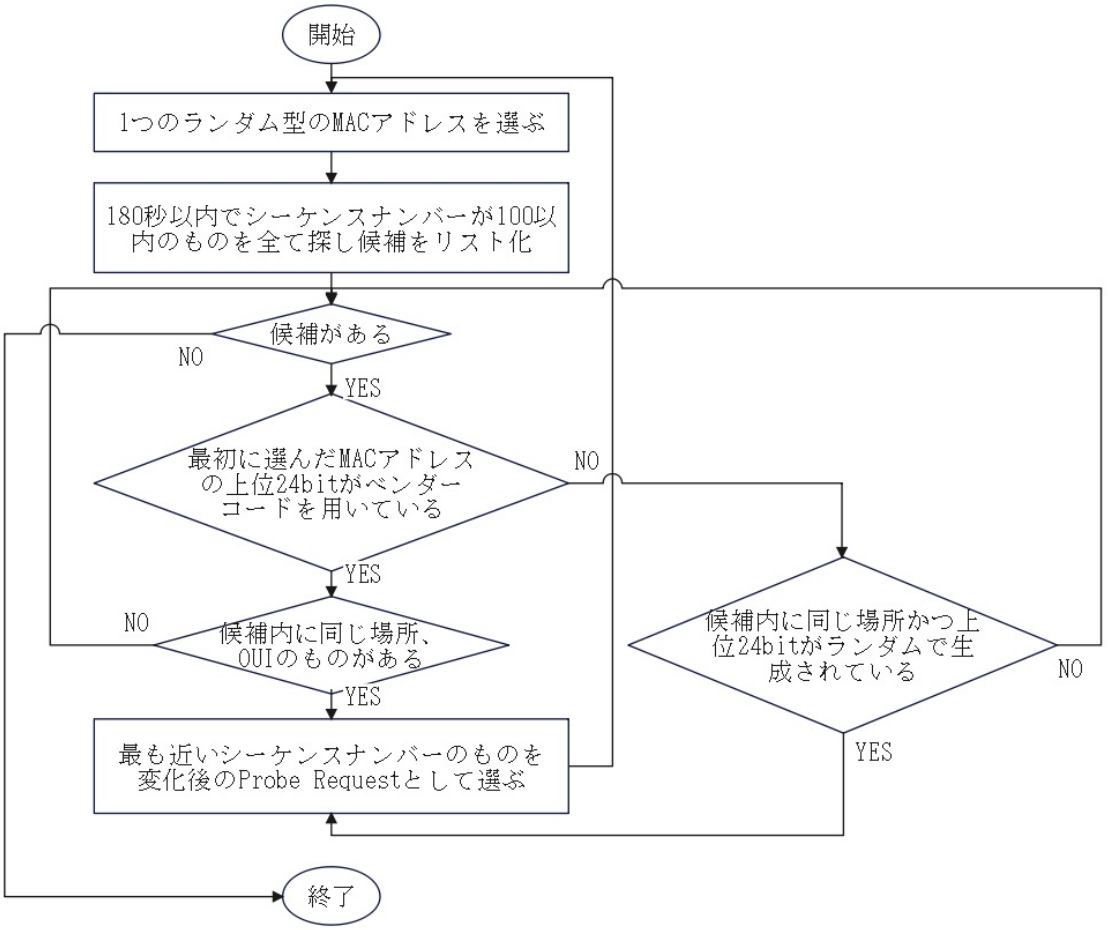

図 7　MAC アドレスの同一推定フローチャート

5　結果と考察

5.1　MAC アドレス全体の内訳

4 週間分のデータを 4.4 章で説明した方法でユニーク型、ランダム型に分類した。また、ランダム型の中からベンダー固定ランダム型、MAC アドレスランダム型に分類した。11 月 29 日のデータではランダム型の MAC アドレスの計測が無かったため、ランダム型に関する値はいずれも 0 になっている。全ての MAC アドレスの中のユニーク型とランダム型に分けて日付ごとに示したグラフを**図 8** に示し、場所ごとに示したグラフを**図 9** に示す。全てのランダム型 MAC アドレスの中のベンダー固定ランダム型と MAC アドレスランダム型・完全ランダム型の内訳を日付ごとに示したグラフを**図 10** に示し、場所ごとに示したグラフを**図 12** に示す。

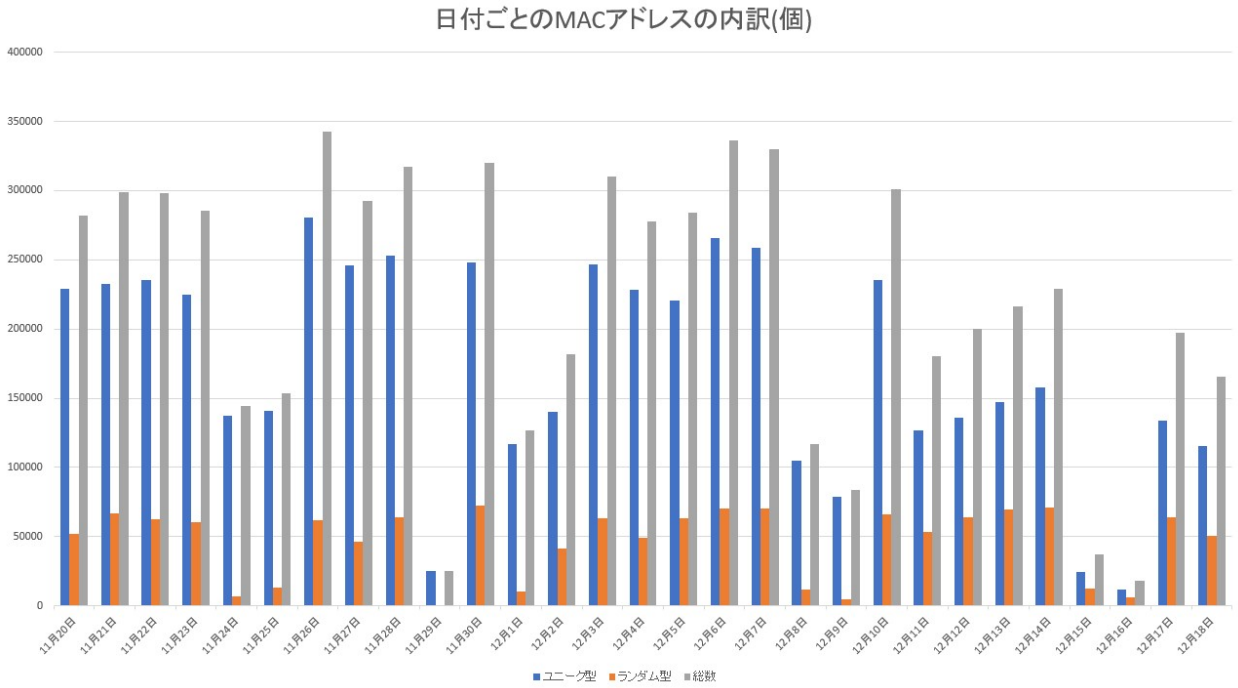

図 8　MAC アドレスの総数と内訳
縦軸は計測した MAC アドレスの総数 (個) で、横軸では計測した日付を示している。MAC アドレスは計測した回数であり、同じモバイルデバイスから発信された MAC アドレスでも重複している。灰色の線グラフが全ての計測した MAC アドレスの総数、青色の線グラフがユニーク型、橙色の線グラフがランダム型を示している。

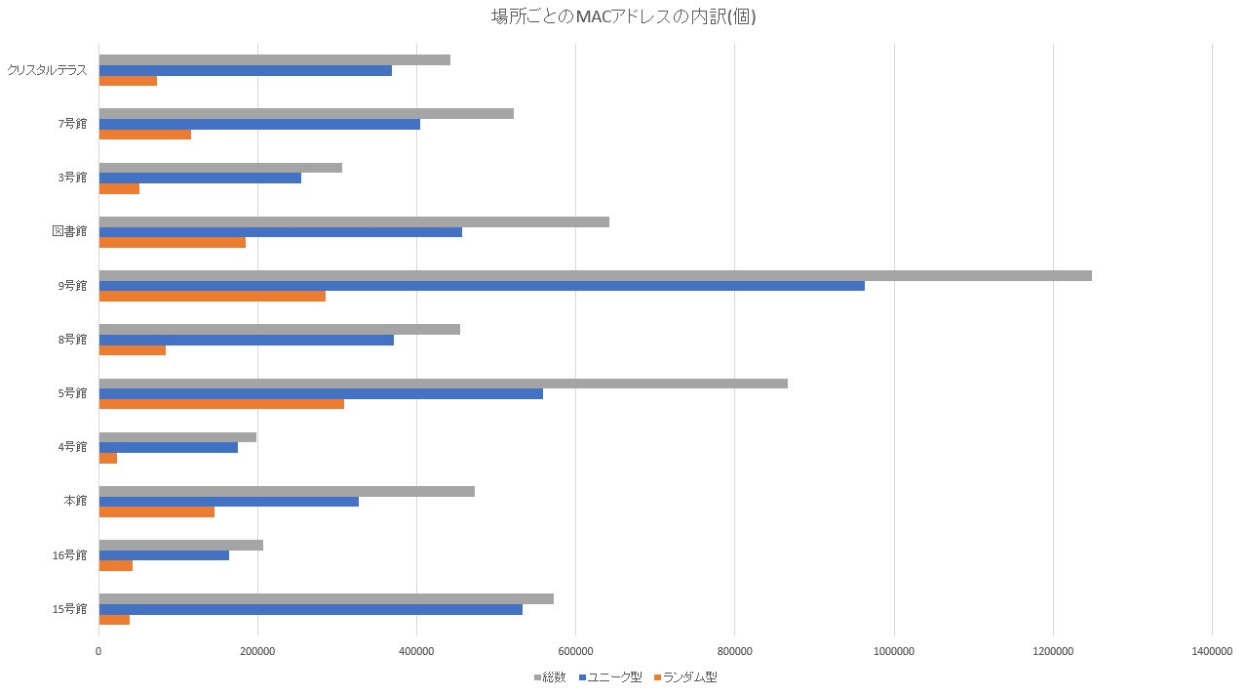

図 9　場所ごとの MAC アドレスの内訳
縦軸は計測した場所を示しており、横軸は計測した MAC アドレスの総数を示しており単位は個である。MAC アドレスは計測した回数であり、同じモバイルデバイスから発信された MAC アドレスでも重複している。灰色の棒グラフが全ての計測した MAC アドレスの総数、橙色の棒グラフがユニーク型、青色の棒グラフがランダム型を示している。

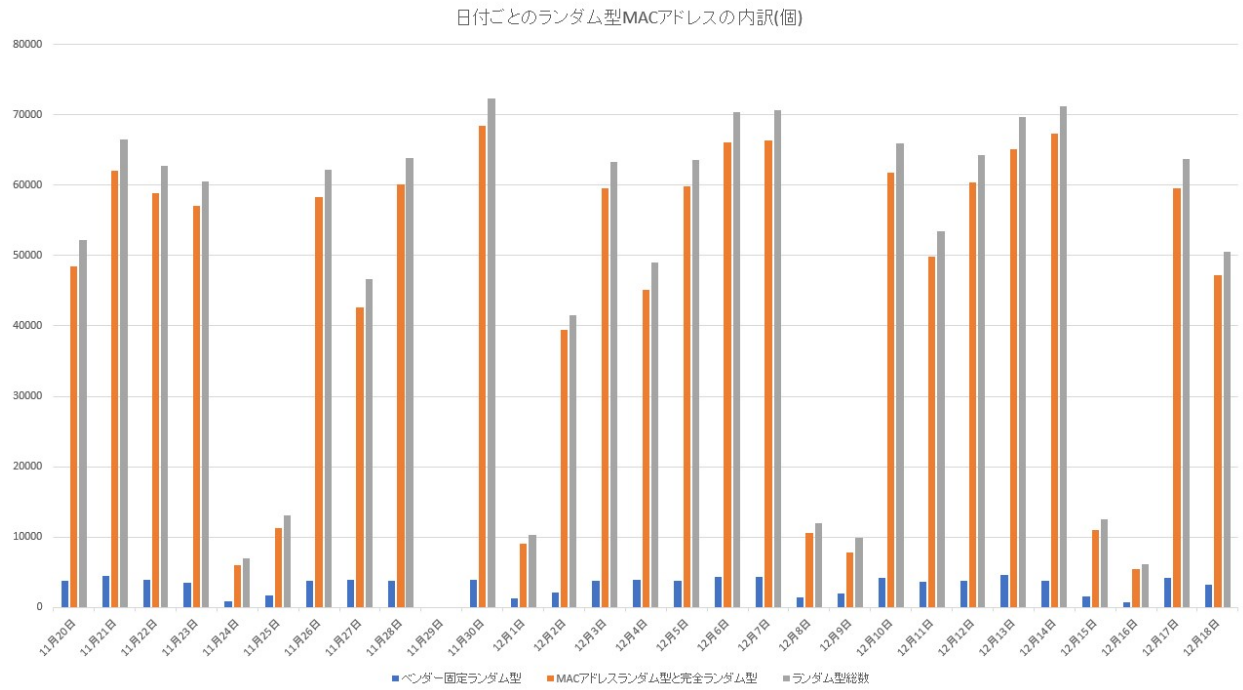

図 10　ランダム型 MAC アドレスの内訳
縦軸は計測した MAC アドレスの総数 (個) で、横軸では計測した日付を示している。MAC アドレスは計測した回数であり、同じモバイルデバイスから発信された MAC アドレスでも重複している。灰色の線グラフが全ての計測したランダム型 MAC アドレスの総数、青色の線グラフがベンダー固定ランダム型、橙色の線グラフが MAC アドレスランダム型と完全ランダム型の合計数を示している。MAC アドレスランダム型のと完全ランダム型は MAC アドレスの値のみでは判別が付かなかった。

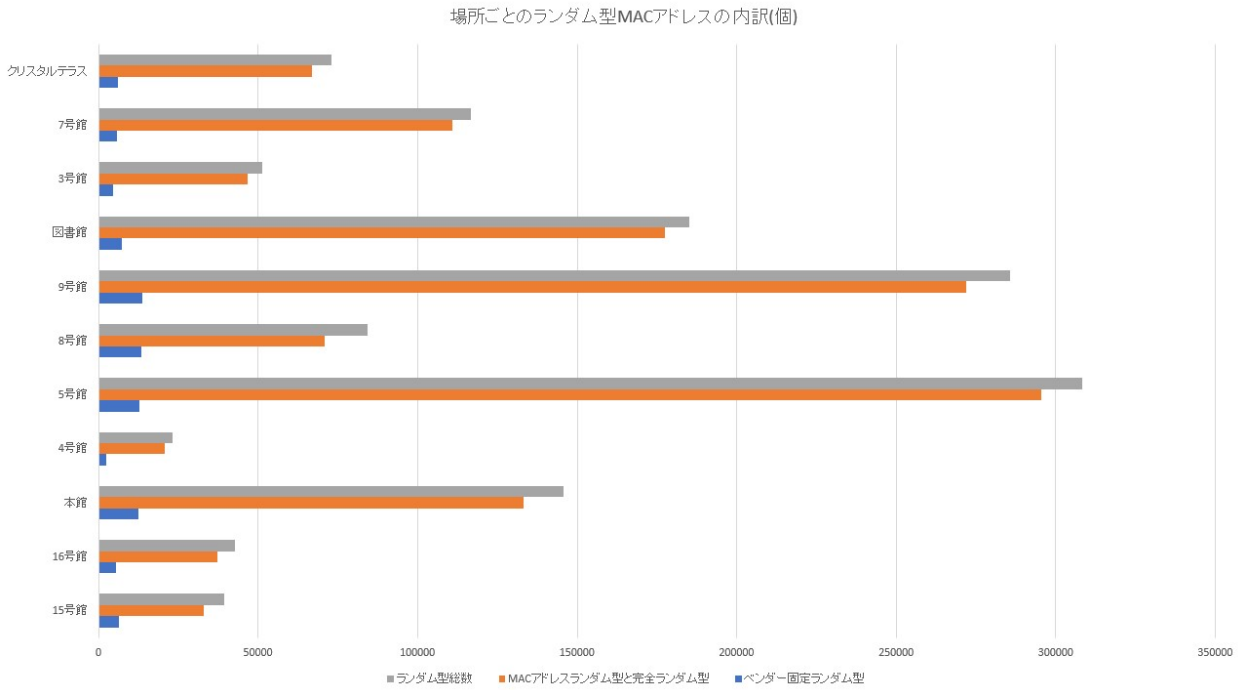

図 11　場所ごとのランダム MAC アドレスの内訳
縦軸は計測した場所を示しており、横軸は計測した MAC アドレスの総数を示しており単位は個である。MAC アドレスは計測した回数であり、同じモバイルデバイスから発信された MAC アドレスでも重複している。灰色の棒グラフが全ての計測したランダム型 MAC アドレスの総数、青色の棒グラフがベンダー固定ランダム型、橙色の棒グラフが MAC アドレスランダム型と完全ランダム型の合計数を示している。MAC アドレスランダム型のと完全ランダム型は MAC アドレスの値のみでは判別が付かなかった。

5.2　同一推定した結果

4 週間分の Probe Request のデータをランダム型のみに分類したのち、同一推定プログラムで実行した。11 月 29 日のデータではランダム型の MAC アドレスの計測が無かったため、いずれの値も 0 になっている。他にランダム型 MAC アドレスと同一のものがあると推定できた MAC アドレスの数を計測したものを**図 12** で示す。全体に比べると MAC アドレスランダム型の方が多く同一推定ができた。ランダム型 MAC アドレスの連続での計測時間の平均を**図 13** で示す。

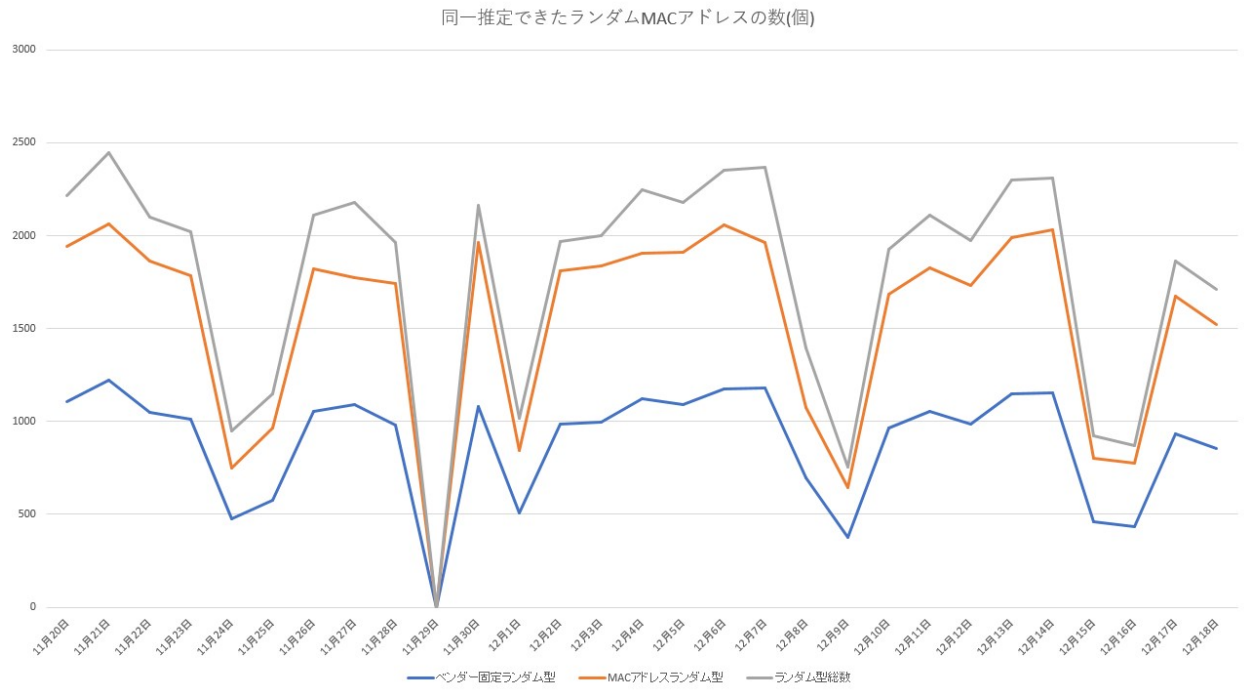

図 12　同一推定できたランダム型 MAC アドレスの数
縦軸は同一推定できた MAC アドレスの数を示しており、横軸は日付を示している。この数値は他の MAC アドレスと同一と推定できた数である。

図 13　ランダム型 MAC アドレスの平均計測時間
縦軸はランダム型 MAC アドレスを連続で計測した平均時間を示しており、横軸は日付を示している。

5.3 考察

ここでは MAC アドレスの分類や同一推定によって得られた MAC アドレスの数について予想及び考察について記述する。

5.3.1 予想

ランダム型とユニーク型の MAC アドレスのそれぞれの合計数はランダム型の方が多いと予想した。日本のスマートフォンの OS のシェア率 [4] は 2018 年の時点で iOS の利用者が全体で 46.7% であり、Android の利用者が 53.3% である。10 代 20 代での iOS シェア率は 10 代女性が 72.5%、20 代女性が 62.7%、10 代男性が 63.5%、20 代男性が 49.1% となっている。全体でみると Android の方がシェア率は多いが、本学で大半を占めている 10 代,20 代では iOS のシェア率が多かった。iOS は MAC アドレスのランダム化を 2014 年の iOS 8.0 では取り入れており、2018 年時点では MAC アドレスのランダム化のデフォルトで搭載していた。そのため、iOS 端末で用いる MAC アドレスの多くはランダム型であると予想した。

Android では 2017 年にリリースされた 8.0 以降のバージョンでランダム化がサポートされているが、2019 年に Android 10.0 までは MAC アドレスのランダム化がデフォルトで無効化されている。したがって、2018 年時点のデータでは Android 端末で用いる MAC アドレスはランダム型が少なくなると予想した。したがって、学内では iOS 端末の利用者の方が多いと考えられるので、MAC アドレスはランダム型の方が多いと予想した。

ランダム型の中のベンダー固定ランダム型、MAC アドレスランダム型、完全ランダム型の割合で一番多いのは完全ランダム型と予想した。計測時期が 2018 年 11 月 12 月なので、Android の最新のバージョンは 9.0 で iOS の最新のバージョンは 12 である。完全ランダム型を用いた端末は Android では Pixcel シリーズや Nexus シリーズなど、iOS では iPhone8,X 以降の端末で確認できた。前述の通り本学内では iOS 利用者の方が多いと考えたので iPhone8,X 以降の端末の利用者も本学内で多いと予想した。

同一の端末の推定ではベンダー固定ランダム型は上位 24bit が固定なので同一推定しやすいと考えたのでベンダー固定型の MAC アドレスの方が多くなると予想した。

5.3.2 結果の考察

4 週間分の Probe Request のデータを全体でみるとユニーク型の MAC アドレスの割合が高かったため予想と外れた。これは、学内には全域にわたって教職員や生徒が自由に接続できるアクセスポイントがある。これにスマートフォンで接続する場合、ランダム型の MAC アドレスを用いる機種であっても、ユニーク型の MAC アドレスかランダム型でも同じ値の MAC アドレスを含む Probe Request を送信しているからだと考えられる。5 号館と図書館ではランダム型の MAC アドレスの計測の割合が比較的高かった。これはアクセスポイントに接続している端末が少ないため、ランダム型として計測したからだと考えられる。

ランダム型の MAC アドレスの中では大半が MAC アドレスランダム型か完全ランダム型であった。MAC アドレスランダム型のと完全ランダム型は MAC アドレスの値のみでは正確に判別ができないため、予想が当たったかはわからない。ベンダー固定ランダム型は Android の端末でしか確認できていないし、Android では MAC アドレスのランダム化がまだ進んでいなかったため少なかったと考えられる。

同一端末の推定では、MAC アドレスランダム型よりベンダー固定ランダム型の方が母数に比べると多く推定でき、予想通りであった。ベンダー固定ランダム型は上位 24bit が固定なので同一の推定がしやすかったためだと考える。計測範囲を改善するか、プログラムの精度を上げることでより多くのランダム型の MAC アドレスの同一推定が可能だったと考えられる。

ランダム型 MAC アドレスの計測時間は平均で約 5 秒程だった。計測結果からプログラムの条件や精度を上げれば、もっと長く計測できただろうと考えた。

6 結論

本研究ではランダム化された MAC アドレスを、シーケンスナンバー、タイムスタンプ等の情報から同一のモバイルデバイスを推定するプログラムを作成した。同一推定は可能であったため、MAC アドレスをランダム化してもモバイルデバイスの動きを追うことは成功した。

6.1 今後の課題

MAC アドレスをランダム化しても、シーケンスナンバーやベンダーコードなどから同一端末であることを推定することに成功した。機器の数を増やして計測範囲を広げるか、計測範囲を 1 つの建物に絞り、計測機器を集中することでモバイルデバイスの Probe Request をより正確に追う必要があると考えた。また、シーケンスナンバーの値だけでなく、今回扱うことがなかった RSSI の値やパケット長も同一のランダム型 MAC アドレスを判別する上で重要だと考えた。たまたま同じ場所で近いシーケンスナンバーのモバイルデバイスがあったときに、RSSI の値を比べてより値が近いモバイルデバイスの方が信憑性が上がる。また、RSSI の強度によっては無線 LAN とモバイルデバイスの距離も推定できるだろう。パケット長はモバイルデバイスの機種によって数値が決まっているので、モバイルデバイスの種類の推定に役立つためである。

本研究の課題及び改善点を以下にまとめる。

- 計測面積を拡大する。
- Probe Request 内の RSSI やパケット長といったデータも用いる。
- 機種や OS によって異なる仕様を見つけて詳細を場合分けをする。

謝辞

本論文執筆及び研究作業等、大垣斉准教授からご指導及びご協力を頂きました。また、本研究を進める上でご協力頂いた情報教育システム研究室の皆様に深く感謝いたします。

参考文献

[1] ICT 総研. 2018 年公衆無線 lan サービス利用者動向調査. https://www.ntt-west.co.jp/news/1904/190425b.html.

[2] 望月祐洋, 上善恒雄, 西田純二, 中野秀男, 西尾信彦ほか. Wi-fi パケットセンサを利用した匿名人流解析システムの構築. 研究報告ユビキタスコンピューティングシステム (UBI).

[3] 角屋卓哉. ”wi-fi を利用した行動解析”のためのシステム開発. 2018 年度 デザイン工学部 情報システム学科 卒業論文, 2018.

[4] mmd 研究所. 2018 年 8 月モバイルデバイスシェア調査. https://mmdlabo.jp/investigation/detail$_1$737.*html*.