ABHD
Revista Brasileira de Humanidades Digitais RBHD

Os Dispositivos Móveis na Internet das Coisas e a Privacidade

Mobile Devices in the Internet of Things and Privacy

Gabriel Filippo C. de Andrade, Marcel William Rocha da Silva, Ubiratam Carvalho de Paula Júnior e Marcelo Panaro de Moraes Zamith

Resumo: *Com a popularização das tecnologias de comunicação sem fio, cada vez mais pessoas carregam equipamentos que ficam sempre conectados à Internet. Este cenário tende a ser acentuado com a evolução de novas aplicações na internet das coisas e ambientes inteligentes, onde parte do princípio de que tudo está conectado. Nesse sentido, um dos problemas que surgem nesse contexto é a privacidade, uma vez que a infraestrutura de redes Wi-Fi requer o uso dos identificadores únicos dos dispositivos, e em especial os smartphones, sendo possível identificar seus usuários e até mesmo monitorá-los. No sentido de minimizar esse problema, os fabricantes de dispositivos móveis desenvolveram uma estratégia de esconder o identificador, contudo, não o fazem de forma eficiente, como será apresentado e discutido neste artigo.*

Palavras-chave: *Dispositivos moveis, Internet das coisas, Privacidade.*

Abstract: *With the popularization of wireless communication technologies, more and more people are using devices that are always connected to the Internet. This scenario tends to be accentuated with new applications on the internet of things (IOT) and intelligent environments which assume everything is connected. With this in mind, one of the problems that arises is privacy, since a Wi-Fi network infrastructure requires the use of the device's unique identifiers and especially smartphones, being possible to identify its users and even track them. In order to minimize this problem, device manufacturers have developed a strategy to hide the identifier, however, they do not do it efficiently, as will be presented and discussed in this article*

Keywords: *Mobile devices, Internet of things, Privacy.*

Introdução

A popularização das tecnologias de comunicação viabilizou a conectividade de pessoas e equipamentos, com um foco especial a comunicação sem fio. Dentre as tecnologias envolvidas nesse sistema, a rede sem fio tem grande protagonismo, uma vez que viabiliza a conectividade a qualquer momento e em qualquer lugar de pessoas e objetos ou coisas, fazendo surgir o termo Internet das Coisas (*Internet of Things* - IOT) [Zanella et al. 2014].

Esse cenário traz uma evolução nas aplicações tornando equipamentos e ambientes inteligentes, capazes de perceber os ambientes e atuar sobre eles. Esta atuação pode ser feita de forma direta, através de comandos autônomos, ou de forma indireta, através do fornecimento de informações para que pessoas possam tomar decisões.

Em nosso cotidiano utilizamos, as vezes sem entender bem os riscos envolvidos, sistemas precursores do IOT para nossas decisões. É possível citar, por exemplo, o caso do trânsito, onde diariamente nos grandes centros urbanos motoristas consultam aplicativos para verificar a qualidade do transito e, em alguns casos, o próprio aplicativo sugere a melhor opção [Google 2020,Waze 2020].

Ainda nesse contexto, outras aplicações da mesma família são conhecidas como ambientes inteligentes. Ambientes inteligentes envolvem sistemas capazes de perceberem o ambiente no qual as pessoas estão e atuam de alguma forma para melhorar o próprio ambiente para seus usuários. Por exemplo, luzes e persianas são acionadas automaticamente conforme a proximidade dos usuários [Oxman 2017]. Também há a aplicação que ajusta o aquecimento, ventilação e a refrigeração de ambientes de acordo com a quantidade de pessoas presentes [Akkaya et al. 2015,Zou et al. 2018]. E, atualmente, no contexto da pandemia de COVID-19, seria possível até mesmo detectar aglomerações em ambientes fechados que pudessem causar episódios de propagação viral.

Os sensores são fundamentais na tecnologia IOT, é através deles que os sistemas baseados em IOT percebem o mundo. Uma das propostas é equipar o ambiente, seja ele uma casa, prédio ou rua, com sensores, que são equipamentos capazes de capturar

informação. Câmera de vídeo, interfaces de rede *bluetooth* ou WiFi são alguns exemplos de equipamentos que podem ser equipados nestes sensores [Zanella et al. 2014]. Os dados capturados por estes sensores são enviados através de conexão com a Internet para servidores que processam essas informações para tomar as decisões ou apenas extrair informação [Zanella et al. 2014].

Esse monitoramento sem a percepção do usuário traz consigo um problema de privacidade dos usuários [Cunche 2014]. É um problema mais evidente quando consideramos os *smartphones*, porque esses dispositivos móveis são verdadeiros sensores portáteis e com especial destaque para a interface de rede WiFi. Um *smartphone* pode enviar a sua identificação única da interface de rede WiFi ainda que não esteja conectado a qualquer rede. Essa identificação, denominada endereço MAC, é um código de 48 bits gravado pelo fabricante em cada interface de rede WiFi produzida e é transmitido pelo equipamento em praticamente todas as transmissões realizadas. Assim sendo, a partir da identificação de um endereço MAC por um sensor WiFi é possível identificar o dono do aparelho de forma indireta.

Esse problema em especial ficou evidente na imprensa no caso do Edward Snowden, que vazou diversos documentos restritos no governo americano [Khandelwal 2014]. Desde então, os fabricantes de dispositivos móveis e desenvolvedores de Sistemas Operacionais buscaram uma solução para inibir esse tipo de monitoramento. A abordagem adotada é conhecida na literatura como aleatoriedade do endereço MAC (*RANDOM MAC*) [Martin et al. 2017]. Este artigo visa analisar a efetividade das abordagens adotadas por fabricantes de *smartphones* para proteger os endereços MAC nos dias atuais e no contexto brasileiro. Para isso, foram utilizados dados coletados por um protótipo de rede de sensores WiFi instalado no Campus de Nova Iguaçu da Universidade Federal Rural do Rio de Janeiro.

Esse trabalho está organizado em Seções, onde os trabalhos relacionados são apresentados na Seção 3. A Seção 2 mostra a fundamentação teórica envolvida na análise. A metodologia empregada nos experimentos é descrita na Seção 4. Os resultados são apresentados e discutidos na Seção 5 e, por fim, conclusões e trabalhos futuros são mostrados na Seção 6.

Redes WiFi e os endereços MAC

O comportamento das redes WiFi é baseado nas definições do padrão IEEE 802.11 [IEEE 1998]. Neste padrão, são descritas diversas características do funcionamento das redes WiFi, como os tipos de rede que podem ser formados, o processo de entrada e saída de estações da rede e a coordenação da disputa pelo acesso ao meio. Uma das funcionalidades, que é importante para este trabalho e será detalhada a seguir, é o processo de descoberta de redes WiFi por estações cliente.

O tipo de rede WiFi mais comum é a rede infraestruturada. Nestas redes existe sempre uma estação especial chamada de ponto de acesso (access point - AP) que coordena as atividades da rede. Este AP guarda informações sobre as estações conectadas à ele e também é responsável por divulgar informações sobre a rede para que novas estações possam se associar.

Com certa frequência, o AP transmite um tipo de quadro especial denominado *beacon*, o qual carrega informações, como por exemplo, o nome da rede WiFi (*service set identifier* - SSID), o canal em que opera e os tipos de protocolos de criptografia suportados. Essas informações são muito relevantes para que novas estações que estejam nas proximidades daquele AP possam perceber sua presença e, eventualmente, iniciar um processo de associação.

O *beacon* é um quadro periódico enviado em *broadcast* pelo AP no seu canal de operação e basta permanecer ouvindo o canal por um tempo para que para que uma estação receba o quadro e descubra a rede. Este processo de descoberta de redes é também chamado de varredura passiva. Entretanto, o padrão IEEE 802.11 define muitos canais de operação e, caso uma estação queira descobrir todas as possíveis redes que estão no seu alcance, a estratégia de ficar por um período de tempo parado em cada canal à espera dos *beacons* pode ser um processo demorado. Sendo assim, o padrão IEEE 802.11 define um processo de descoberta rápida de redes, chamado varredura ativa, que utiliza dois quadros especiais: o *probe request* e o *probe response*.

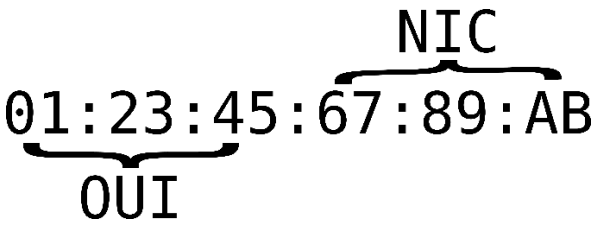
Sempre que uma estação deseja descobrir as redes em seu alcance, realiza-se um processo de varredura dos canais. Este processo consiste em enviar um quadro de *probe request* em cada canal possível e esperar por um tempo curto por respostas dos APs. Todo AP deve estar preparado para, ao receber um quadro de *probe request*, enviar de volta ao transmissor um quadro de *probe response*. As informações contidas no *probe response* são semelhantes àquelas presentes nos quadros de *beacon*, mas o tempo de resposta de um AP no envio do *probe response* é bem menor do que o intervalo médio de tempo entre *beacons*.

Através desse processo de varredura com os quadros de *probe request* uma estação pode descobrir de forma rápida as redes que estão próximas. Por ser mais vantajoso, quase todos os equipamentos dão preferência pela varredura ativa. Entretanto, existe uma característica intrínseca a este processo que pode representar um problema de privacidade.

As estações que realizam varreduras ativas transmitem no *probe request* o seu identificador único, também conhecido como endereço MAC. Como as estações realizam varreduras mesmo quando não estão associadas em nenhuma rede e como o ar é um meio de transmissão *broadcast*, sempre que uma estação realiza uma varredura ativa há um risco iminente de estar expondo a sua presença naquele local onde realizou a varredura. Qualquer outro dispositivo malicioso possuidor de uma interface WiFi que esteja próximo receberá os quadros de *probe request* e identificará que aquele dispositivo detentor daquele endereço MAC está nas proximidades.

Os endereços MAC são projetados para serem únicos e globais em função da própria infraestrutura de rede. A fim de assegurar o endereçamento único de cada dispositivo, o *Institute of Electrical and Electronics Engineers* (IEEE) atribuí blocos de endereços aos fabricantes dos dispositivos. Esses blocos são os três primeiros bytes do endereço MAC ou os primeiros doze bits[Newman 2020] e são chamados na literatura como OUI (*Organizationally Unique Identifier*), enquanto que os demais bits representam a identificação da interface de rede - NIC (*Network Interface Card*) e são atribuídos pelos fabricantes, conforme a Figura 1 ilustra.

Figura 1: Exemplo de endereço MAC



Desta forma, os dispositivos que se comunicam via WiFi incluem algumas informações na comunicação que é possível identificar os fabricantes, os donos dos dispositivos moveis ou ainda monitorar pessoas pela infraestrutura de rede WiFi simplesmente identificando os endereços MAC contidos nos pacotes de comunicação.

Com objetivo de evitar esses problemas de identificação, o IEEE permite que os fabricantes utilizem um OUI privado ou local. Assim é utilizado o bit menos significativo do primeiro byte do endereço MAC para indicar se o endereço é global (bit 0) ou local (bit 1). Quando o bit local está ativado, indica que o endereço MAC é local e, portanto, um endereço aleatório. Caso contrário, endereço é único e global.

Essa abordagem visa resolver os problemas de identificação, privacidade de rastreamento que a infraestrutura do WiFi acaba permitindo. É a abordagem do endereço MAC aleatório. Embora essa a política vise combater rastreamentos, a abordagem do bit local apresenta falhas. Tais falhas comprometem a privacidade uma vez que é possível identificar o endereço real ou global dos dispositivos e, portanto, rastrear ou identificar o seu usuário.

Além do endereço MAC, a informação do SSID cadastrados nos dispositivos móveis também são informados na comunicação WiFi. Serviços como ponto de aceso temporário e comunicação P2P também podem fornecer informações que identifiquem seus usuários.

A política do endereçamento local combinado com a randomização dos endereços MAC, gerando, dessa forma, endereços MAC aleatórios dificulta um pouco o rastreamento e identificação real. Contudo, a questão que surge é se essa política é efetiva e em especial para os dispositivos móveis, onde tais dispositivos podem identificar seus usuários.

Trabalhos relacionados

Alguns trabalhos presentes na literatura científica estudaram o problema da quebra de privacidade causada pela transmissão do endereço MAC nos pacotes de *probe request*. Além disso outros trabalhos avaliaram a eficácia das técnicas de randomização dos endereços MAC adotadas pelos fabricantes.

Franklin et al. [Franklin et al. 2006] propuseram um dos primeiros trabalhos nesse campo de identificação a partir de aparelhos com dispositivos WiFi. Eles desenvolveram um método para criar uma impressão digital dos dispositivos que se comunicam pela rede WiFi, ou seja, uma espécie de assinatura única desses dispositivos. Nesse sentido, os autores desenvolveram uma técnica de aprendizagem que máquina que considera o intervalo de tempo entre os pacotes de *probe request* dos dispositivos, criando uma assinatura a partir do padrão gerado pelo próprio dispositivo na busca por redes WiFi. Os autores ainda discutem algumas limitações da abordagem dos endereços MAC aleatórios.

Demir et al. [Demir, Cunche e Lauradoux 2014] revisão a política de privacidade de 15 grandes empresas de rastreamento de WiFi. A questão que os autores trazem para discussão é sobre a garantia do anonimato ou privacidade dos usuários que foram monitorados sem consentimento, uma vez que essas empresas utilizam o monitoramento para as mais diversas aplicações de IOT. Os autores colocam que a privacidade dos endereços MACs são garantidos pela empresa porque são gravados em suas bases de dados utilizando uma função *hash*. Contudo, os autores demostram que facilmente esses endereços podem ser recuperados e ainda propõem um encadeamento de funções *hash* a fim de dificultar ou evitar que tais endereços sejam recuperados. O principal aspecto deste artigo é que fica bem claro que há empresas que monitoram as redes WiFi sem o conhecimento dos seus usuários.

Cunche, M. [Cunche 2014] discute a questão da privacidade do endereço MAC, colocando esse problema como uma vulnerabilidade. O autor explora duas abordagens para associar o endereço MAC a indivíduos. Na primeira abordagem, o autor explora a

própria estrutura do protocolo 802.11 do WiFi, onde o pacote de *probe request* pode fornecer uma lista de SSIDs preferenciais do dispositivo móvel, para fazer com que o dispositivo móvel revele seu endereço MAC verdadeiro. Para isso, é apenas necessário ter um computador para responder como se fosse um AP, utilizando um dos nomes da SSIDs da lista fornecida pelo *smartphone*. A segunda abordagem é bem complicada de ser utilizada, pois consistem em garantir que não haverá outros dispositivos transmitindo pacotes de *proble request* entre sensor e o *smartphone*, desta forma, os únicos pacotes capturados são do dispositivo móvel, é uma técnica conhecida como *Stalker Attack*.

Martin, J. [Martin et al. 2017] apresentam em seu artigo uma ampla discussão sobre a questão de privacidade e os endereços MAC. Para isso, eles analisam os diferentes fabricantes de smartphones, seus modelos e os sistemas operacionais, que incluem a política de aleatoriedade do endereço MAC pelos computadores pessoais. Eles mostram as falhas das implementações propostas pelas empresas a fim de garantirem a privacidade. E, discutem também técnicas computacionais e algoritmos que tentam identificar os MAC reais a partir dos aleatórios gerados. Ao final discutem sobre a efetividade da aleatoriedade dos endereços MACs e a privacidade dos usuários de smartphones.

Redondi e Cesano [Redondi e Cesana 2018] discutem uma visão geral de três aplicações que podem ser empregadas com o monitoramento da rede sem fio: localização do usuário, perfil do usuário e classificação do dispositivo. Para isso, os autores descrever um sistema de baixo custo capaz de monitorar seus usuários de forma passiva, ou seja, sem intervenção do usuário. Em outras palavras, mostram como é barato e simples fazer o monitoramento da rede WiFi sem o consentimento dos seus usuários.

Coleta dos dados

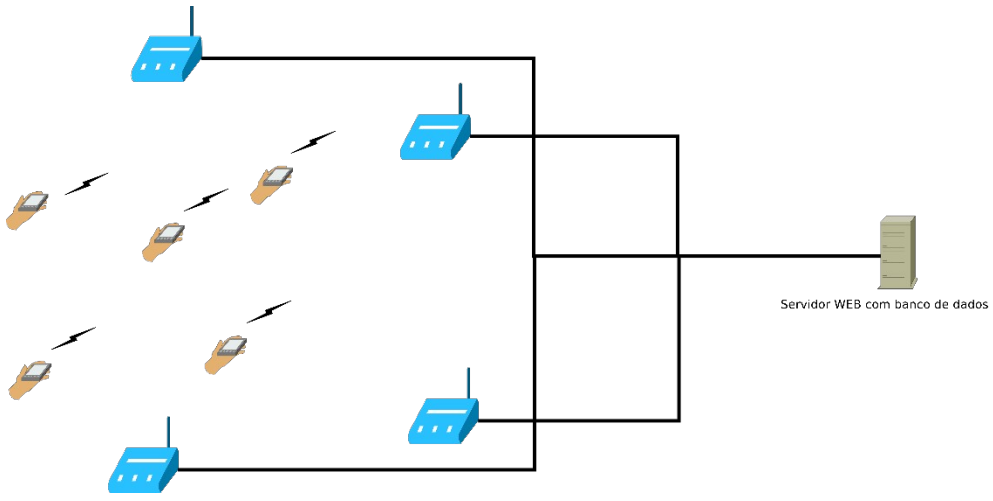
A captura dos dados acontece no ambiente onde os dispositivos com interface WiFi se comunicam em *broadcast* a fim de descobrir possíveis pontos de acesso [Musa e Eriksson 2012]. Para isso, foram instalados sensores, no Campus Nova Iguaçu da

Universidade Federal Rural do Rio de Janeiro, onde fica sediado o Instituto Multidisciplinar.

A coleta de dados WiFi foi realizada por um sistema de sensoriamento formado por cinco computadores SOC (*system on chip*) de baixo custo conhecidos como Raspberry pi que atuam como sensores WiFi e que são equipados com uma placa de rede WiFi e uma placa de rede cabeada Ethernet. O sistema de sensoriamento é formado ainda por um servidor WEB responsável por armazenar numa base de dados as informações recebidas de cada um dos cinco sensores. A comunicação entre os sensores e o servidor é feita por cabo através da rede Ethernet do Campus.

O Campus Nova Iguaçu é composto por cinco prédios, sendo quatro prédios com 3 andares (blocos Administrativo, Multidisciplinar, Informática e Biblioteca) e o quinto prédio anexo com apenas um andar (bloco da pós-graduação). Os sensores foram instalados no segundo andar de cada bloco e no andar térreo bloco da pós-graduação. A Figura 2 ilustra a arquitetura geral do sistema de monitoramento e coleta dos dados. Os sensores captam os pacotes dos aparelhos que estão transmitindo nas proximidades, analisam para extrair informações relevantes e, por fim, enviam pela rede a cabo para um servidor WEB com banco de dados.

Figura 2: Esquema de monitoramento da rede WiFi.



A captura dos pacotes foi realizada com a placa de rede WiFi operando no modo monitor, onde todos os quadros que chegam na placa de rede são entregues para o

sistema operacional. Esses pacotes são analisados por um código próprio desenvolvido em linguagem de programação Python. As informações obtidas dos pacotes são: endereço MAC do transmissor, instante de recepção, tipo do pacote e nível de sinal.

Apesar dos sensores capturarem todos os tipos de pacote dos usuários, para esse estudo da privacidade, serão utilizados apenas os pacotes do tipo *probe request* porque apenas este tipo de pacote randomiza os MAC. Essa diferenciação é possível pois o campo do pacote que identifica o seu tipo também é armazenado no banco de dados pelos sensores. Neste trabalho foram analisados 59 mil pacotes de *probe request*, o que representam todos os pacotes deste tipo coletados em apenas um dia útil de funcionamento do Campus Nova Iguaçu da UFRRJ.

Resultados experimentais

Os resultados são divididos em quatro experimentos, o experimento I visa verificar a efetividade do bit local nos endereços MACs aleatórios; a seguir, no experimento II, é analisado os demais campos dos pacotes de *probe request* com o objetivo de identificar os rastrear pessoas ou aparelhos independentemente do endereço MAC; o experimento III verifica a aleatoriedade dos endereços MACs segundo os fabricantes; e, por fim, o experimento IV buscar recuperar o endereço MAC global a partir do endereço local. Logo, as Subseções a seguir apresentam e discutem cada um dos experimentos realizados para este trabalho.

Experimento I

O primeiro experimento tem por objetivo analisar e avaliar a efetividade do bit local nos MACs aleatórios. Verificar se o padrão definido é seguido pelos fabricantes em relação ao bit local e as implicações que podem acontecer caso o padrão não seja cumprido.

Para isso, foram selecionados os registros da base relacionados aos pacotes de *probe request*, pois os pacotes de dados já expõem o endereço MAC real dos aparelhos.

Em seguida, foram obtidos os endereços MAC dos pacotes a partir do bit indicador de endereço global e local, considerando que os fabricantes adotam o padrão de gerar MAC aleatórios com o bit local ativado.

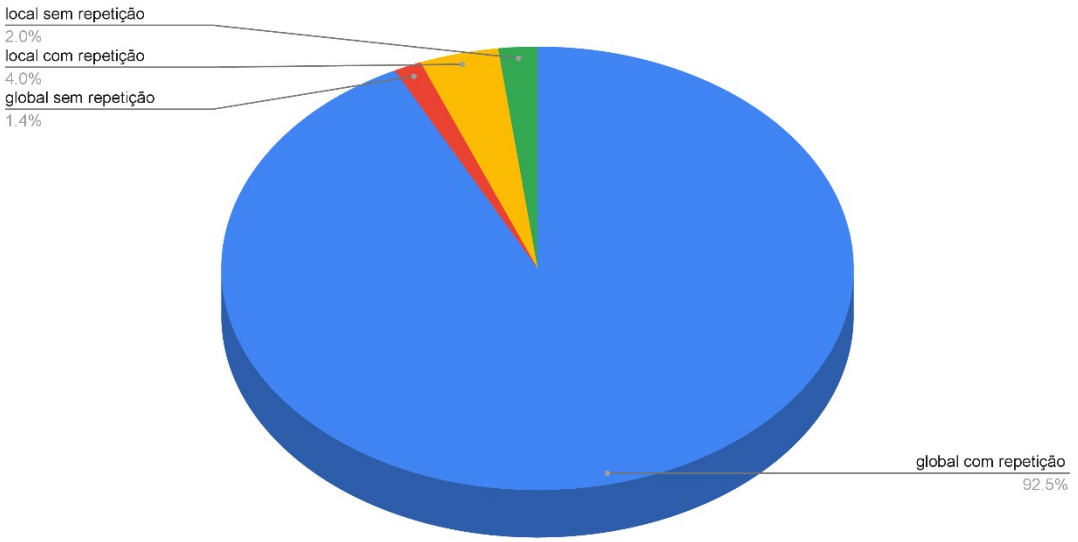
Os endereços MACs foram agrupados em quatro classes: global com repetição, global sem repetição, local com repetição e local sem repetição. Endereços com repetição são aqueles MAC que aparecem na base de dados mais de uma vez, enquanto os MACs sem repetição aparecem apenas uma única vez.

O gráfico da Figura 3 apresenta o resultado da análise da base de dados, onde 92,5% representa endereços com o bit local desativado, endereços globais e que aparecem algumas vezes dentro da base. 1,4% são endereços com o bit local ativado e que aparecem apenas uma única vez. Observando o bit local ativado, 4% aparecem frequentemente, enquanto apenas 2% aparecem uma única vez.

Este resultado sugere que 95% dos pacotes analisados não respeitam o bit local. Logo, é possível concluir que: ou os dispositivos geram os MAC aleatórios utilizando endereços globais, o que eventualmente poderia identificar outro aparelho, ou expõem seus próprios endereços globais. De uma forma ou de outra a privacidade é comprometida, sendo o primeiro caso mais delicado, pois o MAC gerado pode ser de outro aparelho e outra pessoa poderia ser identificada.

Apenas 6% dos aparelhos empregam o MAC aleatório conforme o padrão definido, seja gerando endereços aleatórios com repetição ou sem repetição. No segundo caso, há a possibilidade de que o aparelho fique trocando seu endereço local em um certo intervalo de tempo, ao passo que a repetição indica que o endereço é utilizado por um tempo maior. Em ambos os casos, a aleatoriedade é garantida e cumpri o que foi definido no padrão, dificultando a identificação desses aparelhos e até mesmo seu monitoramento.

Figura 3: MACs com bit local desativado (global) x bit local ativado



Experimento II

Os pacotes de *probe request* possuem outros campos de dados e um desses campos é a informação do ponto de acesso que o aparelho já esteve conectado e que registrou como favorita, ou seja, o SSID das redes favoritas. Este segundo experimento explora esse campo dos pacotes do *probe request* responsável por trazer a informação do SSID das redes favoritas [Zhao et al. 2019]. Logo, a questão que emerge é se a lista de favoritos dos SSID pode comprometer a privacidade do dispositivo?

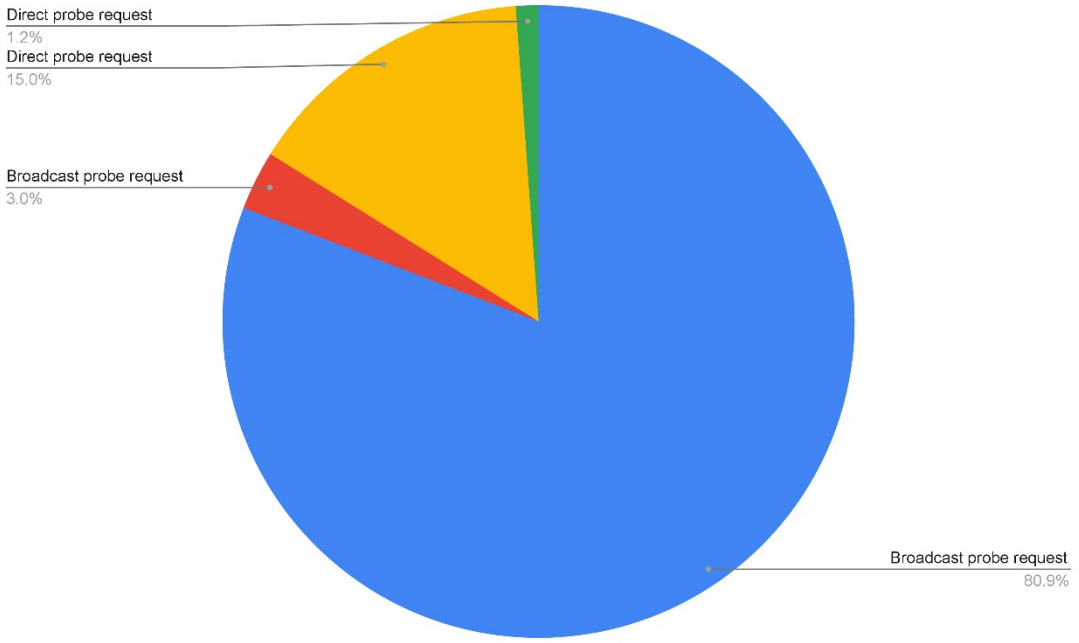
Neste experimento, foi utilizada a mesma base de dados, filtrando os pacotes de *probe request* com endereços MACs globais ou locais. Logo, a Figura 4 ilustra os resultados, onde 16% dos dispositivos enviaram pacotes *probe request* com a lista de SSIDs das redes favoritas, também conhecido como *direct probe request*, sendo que 98% destes não utilizaram o bit local. Nesse resultado foi possível verificar que a privacidade foi comprometida em duas dimensões: a primeira pelo uso do endereço MAC global e na segunda dimensão pela lista do SSID que o endereço MAC usualmente se conecta.

O problema ganha uma outra dimensão, quando combinamos a informação do SSID e o endereço MAC. Desta forma, podemos não apenas identificar o usuário como também

os lugares que costuma frequentar para se conectar à Internet, especialmente quando utilizamos a informação do site https://wigle.net/, que fornece o endereço do ponto de acesso, considerando que o site é atualizado frequentemente.

Embora haja um esquema de ocultar o endereço MAC global dos aparelhos e que se mostrou ineficiente, nada ainda foi proposto para o campo do pacote com informações sobre os SSIDs. Assim, muitos usuários estão deixando um rastro digital por onde passam com seus dispositivos móveis, permitindo o rastreamento de pessoas [Newman 2017].

Figura 4: Pacotes de *probe request* em *broadcast* e *direct*.



Experimento III

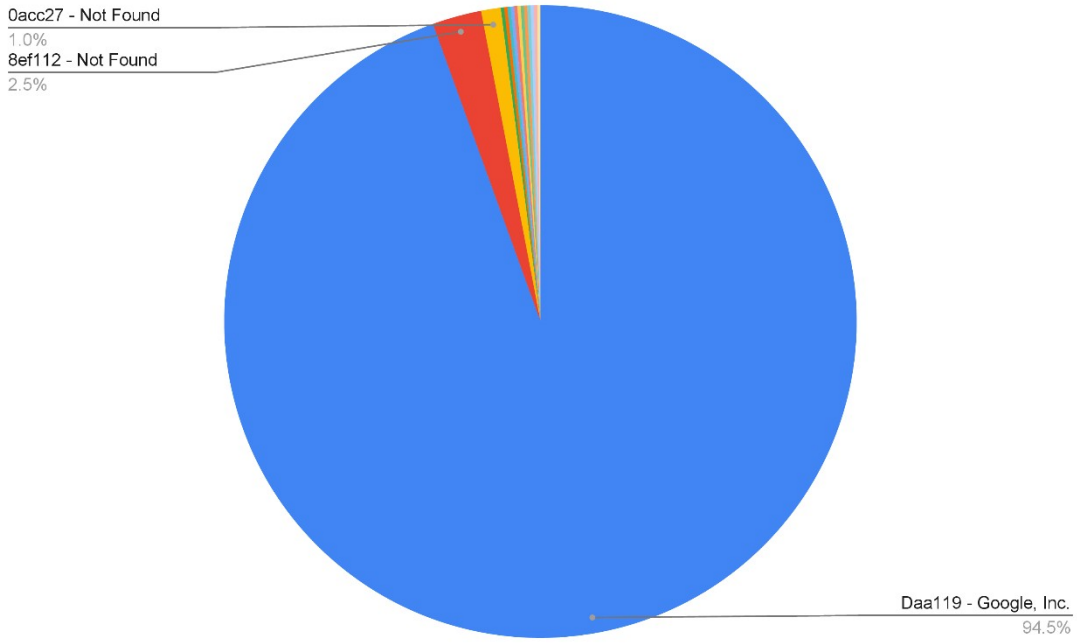
O terceiro experimento visa analisar a política de aleatoriedade dos endereços MACs de acordo com os fabricantes. Para isso, foram selecionados os endereços da base de dados de acordo com os fabricantes.

O endereço do fabricante é identificado de forma direta, utilizando para isso o prefixo do endereço MAC, conhecido como OUI (*Organizationally Unique Identifier*). Ou seja, os três primeiros bytes do endereço MAC.

Assim, foi realizada análise da frequência dos dez fabricantes que mais aparecem na base de dados, considerando tanto o bit local ativado quanto desativado.

A Figura 5 ilustra a distribuição dos endereços conforme seus fabricantes com o bit local ativo. Além disso, o prefixo DA:A1:19 teve grande destaque, pois representa 94% das amostras analisadas. Após consulta realizada no site https://www.macvendorlookup.com/, o prefixo foi identificado como sendo da Google. Os outros nove fabricantes que mais apareceram na base representam apenas 6% dos registros.

Figura 5: Distribuição dos endereços MAC e seus fabricantes - bit local ligado



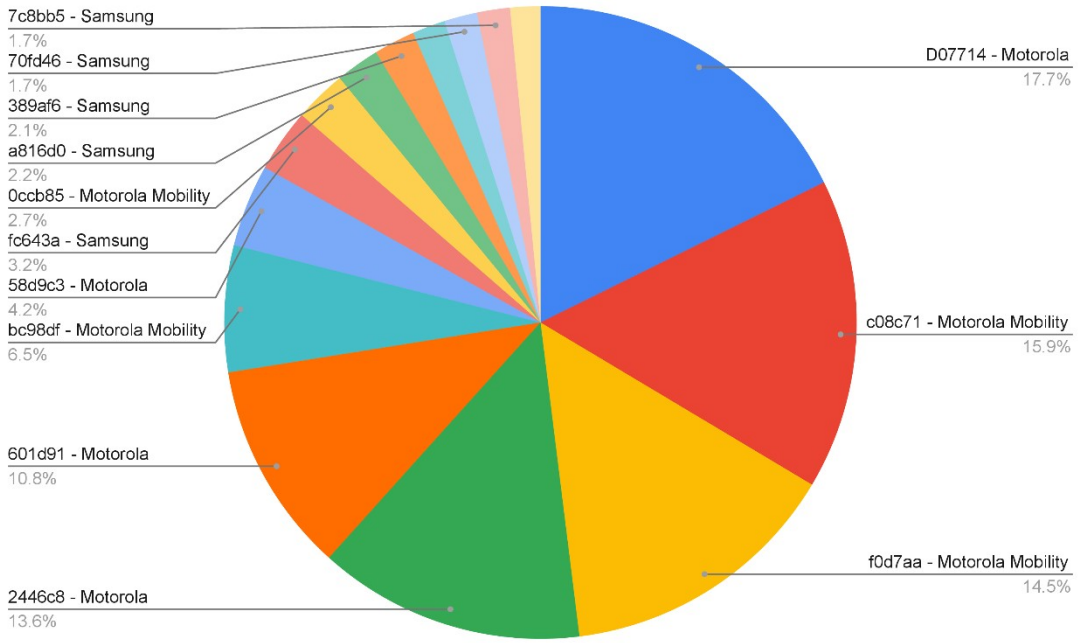
Por outro lado, a Figura 6 mostra a lista dez dos fabricantes com o bit local desativado que apareceram na base de dados. É possível notar que aparecem apenas os aparelhos da Motorola e Samsung. Esse resultado indica ou que esses fabricantes não estão gerando endereços aleatórios ou quando o fazem, fazem com o bit local desativando, o que gera o endereço de um outro aparelho produzido pelo mesmo fabricante.

Há um claro comprometimento da privacidade quando o bit local está desativado, pois gerar pacotes com endereço global compromete a privacidade daquele dispositivo que

gera o pacote, quando gera pacotes com o endereço MAC aleatório sem o bit local ativada pode fazer com que outro dispositivo seja identificado de forma equivocada.

No caso em que o bit local foi considerado desativado, ou não os aparelhos informaram seus verdadeiros endereços ou aleatórios o endereço real de um outro aparelho. Ambos os casos apresentam uma falha na privacidade, mas o segundo caso apresentou um cenário mais grave, uma vez que um outro endereço global gerado implica a identificação de um outro dispositivo e, como consequência, a identificação de uma pessoa errada.

Figura 6: Distribuição dos endereços MAC e seus fabricantes - bit local desligado



Experimento IV

O experimento quatro tem como objetivo verificar a possibilidade de obter o endereço real a partir do endereço local, ou seja, a operação inversa. Para isso, foi invertido o bit local dos endereços identificados como endereços locais da base de dados. Neste sentido, aproximadamente 1% dos endereços analisados foram identificados como endereços globais. Os fabricantes identificados foram Motorola Mobility LLC e a

Lenovo. Este resultado, na prática, identifica apenas um único fabricante, uma vez que a Motorola foi comprada pela Lenovo [Tozetto 2014].

Assim sendo, esses endereços locais ou utilizam randomização para os 24 bits menos significativos do endereço MAC ou, no pior dos casos, não possuem um algoritmo de randomização em si. Esta falha simplifica o rastreamento destes dispositivos, o que infringe diretamente a privacidade.

Conclusão

A cada dia a vida cotidiana fica mais conectada, seja pela comunicação entre as pessoas seja pela comunicação entre dispositivos. Uma grande parcela da população tem acesso à Internet a qualquer hora e a qualquer momento através dos *smartphones* [FGV 2020].

No Brasil, foram registrados em dezembro de 2019 149.057.635 de usuários na Internet e 149.057.635 *smartphones* com acesso à rede mundial de computadores [internetworldstats 2020]. Além disso, foi registrado um crescimento de aproximadamente de 21% dispositivos acessando à Internet e a previsão para 2021 é de mais de 22 bilhões de aparelhos móveis conectados [Baron 2020].

De fato, lojas, shoppings centers, restaurantes e até mais ônibus oferecem acesso à Internet ou gratuita ou por um preço acessível. E esse cenário traz como grande questão o crescimento da conectividade. E, nesse contexto, a privacidade emerge como um problema desconhecido para a grande maioria dos usuários. Possibilidade de rastreamento de pessoas, identificação e construção de padrões comportamentais são alguns os elementos que podem ser obtidos com a identificação dos *smartphones*.

Edward Snowden expôs entre diversos assuntos delicados o rastreamento/monitoramento de pessoas pelos dispositivos móveis [Newman 2017]. Desde então, os desenvolvedores de Sistemas Operacionais e fabricantes de *smartphones* buscam uma solução para o problema da comunicação sem fio.

Assim, esse trabalho buscou explorar a privacidade na comunicação dos dispositivos móveis e a solução proposta pelos envolvidos. Vale ressaltar que o problema da

privacidade na comunicação WiFi acorre com qualquer dispositivo conectado (computadores de mão, tablets, impressoras WiFi, etc), embora fique mais evidente com os *smartphones*.

Com base nos ambientes conectados e na própria infraestrutura da rede WiFi é possível concluir que cada vez estamos expostos. É claro que a identificação de centenas de pessoas num shopping center é um problema complicado de resolver apenas pelo monitoramento da rede WiFi. Por outro lado, é possível rastrear ou monitorar dispositivos desde que se conheça previamente seus os endereços MACs.

Por exemplo, sabendo o endereço MAC dos telefones dos empregados de uma empresa, é possível saber a hora de chegada e saída de cada um deles. É possível também inferir por quais setores eles estiveram ao longo do dia e até mesmo o tempo gasto com os cafezinhos.

Outra variável no problema da privacidade da comunicação sem fio é o conhecimento dos pontos de acesso, os SSID cadastrados. Essa variável viabiliza o conhecimento dos locais que o dispositivo costuma estar conectado. E, eventualmente, identificar os endereços dos SSIDs cadastrados, o que expõe os locais.

Atualmente, não há uma forma efetiva de evitar a exposição dos dados móveis como o endereço MAC ou mesmo o SSID cadastrado nos aparelhos, mesmo com a política de bit local combinado com endereços MACs aleatórios.

Além disso, os resultados analisados sugerem que muitos dos fabricantes de *smartphones* não fazem o que preconizam ou o que foi acordado como padrão. Expondo não apenas seus endereços reais como em alguns casos informando o endereço de outro aparelho, o que pode ser visto como um erro mais grave, uma vez que outra pessoa poderia, em tese, ser identifica ou rastreada de forma equivocada.

Na prática, a análise dos dados mostra que mesmo com todas as deficiências na política de privacidade dos endereços MAC, é bem complicado identificar uma pessoa sem ter acesso físico ao seu aparelho celular, a fim de obter o endereço MAC do aparelho, em função da grande quantidade de aparelhos transmitindo simultaneamente. Por outro

lado, conhecendo previamente o endereço MAC, é fácil rastrear o usuário, pois estando o aparelho com o WiFi ativo, ainda que não conectado, o aparelho vai informar em algum momento seu endereço MAC real, permitindo a identificação do usuário.

Se por um lado a privacidade fica comprometida com a comunicação sem fio, o outro lado da moeda traz as inúmeras aplicações em IOT e em ambientes inteligentes. Tornando a vida mais ágil e confortável. Como por exemplo, o ambiente que ajusta o consumo energético de acordo com a quantidade de pessoas no ambiente. Existe ainda ambientes que percebem a chegada do usuário e respondem de forma automática, acendendo uma luz, por exemplo.

A solução para o problema pode ser dividida em dois momentos. A curto prazo, os fabricantes de dispositivos móveis ou que usam a infraestrutura da rede WiFi poderiam alertar os usuários de um possível comprometimento da privacidade em razão da comunicação WiFi. A longo prazo, desenvolver políticas para sanar de forma eficiente o problema da privacidade na comunicação WiFi, seja por novos protocolos seja por uma nova estrutura de rede sem fio.

Como uma última conclusão, vale mencionar a Lei número 13.709, de 14 de agosto de 2018 (lei de proteção dos dados), que contempla a questão dos dados sensíveis e a sua proteção. Contudo, a questão central deste trabalho é que os dados são transmitidos sem qualquer conhecimento dos seus usuários. E, mais que isso, com um equipamento de baixo custo é possível monitorar o ambiente, permitindo que qualquer pessoa possa construir um ambiente em sua própria residência. Não precisa ser uma grande corporação para monitorar pessoas.

Algumas questões emergiram deste trabalho. Como a necessidade de campanhas, ou até mesmo políticas públicas, para a conscientização das pessoas quanto a sua privacidade no uso dos dispositivos móveis. Que poderia ser a recomendação de manter seus dispositivos com WiFi desligado quando não houver necessidade de uso, pois uma simples atitude de desativar o WiFi do aparelho já é suficiente para solucionar o problema da privacidade de forma eficiente.

Como trabalhos futuros pretendemos utilizar a metodologia empregada nesse trabalho na construção de um ambiente de sensoriamento móvel, permitindo assim capturar os dados de espaços abertos como ruas e praças. E, em seguida, repetir os experimentos deste trabalho para verificar se há a repetição desses resultados.

Além disso pretendemos também propor e desenvolver algoritmos eficientes capazes de minimizar a quantidade de MACs aleatórios. Nessa primeira abordagem, a ideia não é identificar o MAC real, mas eliminar os MAC aleatórios a fim de obter uma precisão maior da quantidade de dispositivos contados. Por fim, propor e testar técnicas para ocultar endereços MACs a fim de garantir um ambiente conectado sem o comprometimento da privacidade.

Referências Bibliográficas

Akkaya et al. 2015

Akkaya, K. et al. Iot-based occupancy monitoring techniques for energy-efficient smart buildings. In: *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. [S.l.: s.n.], 2015. p. 58–63.

Baron 2020

BARON, C. *Statista*. [S.l.], 2020. Disponível em: https://www.statista.com/statistics/802706/world-wlan-connected-device/. Acessado em 29 de novembro de 2020.

Cunche 2014

CUNCHE, M. I know your mac address: Targeted tracking of individual using wi-fi. *Journal of Computer Virology and Hacking Techniques*, Springer, v. 10, n. 4, p. 219–227, 2014.

Demir, Cunche e Lauradoux 2014

DEMIR, L.; CUNCHE, M.; LAURADOUX, C. Analysing the privacy policies of wi-fi trackers. In: *Proceedings of the 2014 workshop on physical analytics*. [S.l.: s.n.], 2014. p. 39–44.

FGV 2020

FGV. *Brasil tem 424 milhões de dispositivos digitais em uso, revela a 31a Pesquisa Anual do FGVcia*. [S.l.], 2020. Disponível em: https://portal.fgv.br/noticias/brasil-tem-424-milhoes-dispositivos-digitais-uso-revela-31a-_pesquisa-anual-fgvcia. Acessado em 29 de novembro de 2020.

Franklin et al. 2006

FRANKLIN, J. et al. Passive data link layer 802.11 wireless device driver fingerprinting. In: *USENIX Security Symposium*. [S.l.: s.n.], 2006. v. 3, p. 16–89.

Google 2020

GOOGLE. *Google Maps*. [S.l.], 2020. Disponível em: https://www.google.com.br/maps/. Acessado em 23 de novembro de 2020.

IEEE 1998

IEEE. Standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications. *ANSI/IEEE Std 802.11, 1999 Edition (R2003)*, p. 1–512, 1998.

internetworldstats 2020

INTERNETWORLDSTATS. *Internet World Stats Usage and Population Statistics*. [S.l.], 2020. Disponível em: https://www.internetworldstats.com/south.htm#br. Acessado em 29 de novembro de 2020.

Khandelwal 2014

KHANDELWAL, S. *Spying agencies tracking your location by capturing MAC address of your devices*. [S.l.], 2014. Disponível em: https://thehackernews.com/2014/01/spying-agencies-tracking-your-location_31.html. Acessado em 12 de outubro de 2020.

Martin et al. 2017

MARTIN, J. et al. A study of mac address randomization in mobile devices and when it fails. *Proceedings on Privacy Enhancing Technologies*, Sciendo, v. 2017, n. 4, p. 365–383, 2017.

Musa e Eriksson 2012

MUSA, A.; ERIKSSON, J. Tracking unmodified smartphones using wi-fi monitors. In: *Proceedings of the 10th ACM conference on embedded network sensor systems*. [S.l.: s.n.], 2012. p. 281–294.

Newman 2017

NEWMAN, L. H. *WikiLeaks Dump Reveals a Creepy CIA Location-Tracking Trick*. [S.l.], 2017. Disponível em: https://www.wired.com/story/wikileaks-cia-wifi-location-tracking/. Acessado em 9 de novembro de 2020.

Newman 2020

NEWMAN, L. H. *Guidelines for Use Organizationally Unique Identifier (OUI) and Company ID (CID)*. [S.l.], 2020. Disponível em: https://www.wired.com/story/wikileaks-cia-wifi-location-tracking/. Acessado em 14 de novembro de 2020.

Oxman 2017

OXMAN. *Internet das Coisas (IoT) Internet of Things*. [S.l.], 2017. https://www.oxman.com.br/post/2018/02/27/internet-das-coisas-iot-internet-of-things. Acessado em 23 de novembro de 2020.

Redondi e Cesana 2018

REDONDI, A. E.; CESANA, M. Building up knowledge through passive wifi probes. *Computer Communications*, Elsevier, v. 117, p. 1–12, 2018.

Tozetto 2014

TOZETTO, C. *Lenovo conclui compra da Motorola e já é terceira maior fabricante de celular*. [S.l.], 2014. Disponível em: https://veja.abril.com.br/tecnologia/lenovo-conclui-compra-da-motorola-e-ja-e-3a-maior-fabricante_-de-celular/. Acessado em 20 de novembro de 2020.

Waze 2020

WAZE. *Instruções de navegação, alertas de trânsito e caronas por Waze*. [S.l.], 2020. Disponível em: https://www.waze.com/pt-BR/. Acessado em 23 de novembro de 2020.

Zanella et al. 2014

Zanella, A. et al. Internet of things for smart cities. *IEEE Internet of Things Journal*, v. 1, n. 1, p. 22–32, 2014.

Zhao et al. 2019

ZHAO, S. et al. Discovering individual life style from anonymized wifi scan lists on smartphones. *IEEE Access*, IEEE, v. 7, p. 22698–22709, 2019.

Zou et al. 2018

ZOU, H. et al. Device-free occupancy detection and crowd counting in smart buildings with wifi-enabled iot. *Energy and Buildings*, v. 174, p. 309 – 322, 2018. ISSN 0378-7788. Disponível em: http://www.sciencedirect.com/science/article/pii/S0378778817339336.