Randomização de endereço MAC como técnica para prover Privacidade a Usuários de redes WiFi

Bruno S. Alves[1], **Yagor S. Duarte**[2], **Bolívar M. Silva**[1]

[1]Centro de Tecnologia – Universidade Federal de Santa Maria (UFSM)
Caixa Postal 1000 – 97105-900 – Santa Maria – RS – Brazil

[2]Colégio Técnico Industrial – Universidade Federal de Santa Maria (UFSM)

bdalves@inf.ufsm.br,{bolivar, yagor}@redes.ufsm.br,

Abstract. *MAC (Media Access Control) address randomization is a technique commonly used to raise the level of privacy of devices, implemented in the vast majority of current operating systems. This hardware address remodeling acts to restrain third parties from analyzing the information traffic and linking them to a device, and even using them for tracking, that is, obtaining the physical location of the device and consequently the user. The way in which such randomization will be performed is exclusively at the discretion of the device manufacturers. The main idea of this article is to expose reasons why privacy measures are still necessary and why only MAC address randomization is necessary butnot enough to guarantee the expected protection and anonymity.*

Resumo. *MAC (Media Access Control) address randomization é uma técnica comumente utilizada para elevar o nível de privacidade dos dispositivos, implementada na grande maioria dos sistemas operacionais atuais. Essa remodelação do endereço de hardware atua com intuito de coibir terceiros de analisar o tráfego de informações e vinculá-las a um dispositivo, e de até mesmo utilizá-los para rastreio, ou seja, obter a localização física do dispositivo e consequentemente do usuário. A maneira como tal randomização será realizada é de decisão exclusiva dos fabricantes dos dispositivos. A ideia principal deste artigo é expor motivos pelos quais ainda se fazem necessárias medidas de privacidade e por que somente à randomização do endereço MAC é necessária, porém não suficiente para garantir a proteção e anonimato esperados.*

1. Introdução

Com o crescimento do uso de smartphones, cada vez mais capazes de realizar com eficiência tarefas diárias, e até mesmo substituir os computadores pessoais em alguns casos, surgem novas possibilidades de inovações tecnológicas nas mais diversas áreas. De acordo com uma pesquisa realizada pela Folha de São Paulo [de Sao Paulo 2017] e levantamentos do IBGE (Instituto Brasileiro de Geografia e Estatística), o número de *Smartphones* no Brasil supera o número de habitantes. Praticamente todas as pessoas carregam consigo no mínimo um desses dispositivos. Tais fatos demonstram que a sociedade está cada vez mais ligada a dispositivos conectados em rede. Como consequência, o nível de privacidade do usuário está cada vez menor. Cada um desses dispositivos está equipado com diversos sensores (GPS, acelerômetros, sensor de proximidade, microfone, sensor de luminosidade, giroscópio, etc), que capturam e compartilham informações do usuário e

do ambiente, para diversos fins e aplicações [Ali and Khusro 2016]. Entre esses, está a interface *Wireless* que possibilita conexões WiFi (padrão IEEE 802.11). Essa interface é identificada por um endereço denominado MAC (*Media Access Control*).

O endereço MAC é um endereço físico composto por 12 dígitos hexadecimais (48 bits) utilizado para identificar as interfaces de rede de um dispositivo [Eastlake 3rd and Abley 2013]. Por se tratar de um endereço globalmente único, para rastreá-lo, basta que o interessado capture informações de rastros digitais deixados por esse dispositivo, para saber sua localização. Uma vez conhecida a localização do dispositivo, consequentemente, a localização do portador do mesmo também será conhecida. Essa técnica, passou a ser referida na literatura como *tracking* [Ali and Khusro 2016] [Patil and Kokil 2015].

O rastreio de pessoas em ambientes públicos e privados representa uma ameaça a segurança e privacidade dos usuários. Uma vez que essas informações podem ser utilizadas para diversos fins, desde a identificação de mobilidade urbana em cidades, até a identificação de clientes em centros comerciais. De posse das informações a respeito das preferências dos clientes, o *marketing* das lojas, podem ser direcionados para estes usuários. Outro problema de privacidade que pode vir a surgir, está relacionado com o anonimamento incorreto dos dados capturados. Por esses e outros motivos, empresas de tecnologia passaram a implementar a randomização do endereço MAC [1], com objetivo de dificultar o rastreamento de um dispositivo. Essa técnica consiste em gerar endereços MAC para as interfaces do dispositivo, de forma aleatória e periódica. Infelizmente, essas técnicas apresentam falhas, e na maioria das vezes, não são efetivas, conforme demonstrado em [Jeremy Martin 2017][Vanhoef et al. 2016] [Martin et al. 2017].

Nesse contexto, o objetivo desse artigo é realizar um estudo sobre as técnicas de randomização adotadas pelos principais sistemas operacionais utilizados em dispositivos que oferecem suporte a conexão WiFi, focando principalmente nos sistema utilizados em *Smartphones*. Para isso, será montado um cenário, com um *Access Point* e posteriormente testado, sondagem e conexão de alguns dispositivos de sistemas operacionais e fabricantes diferentes. Além disso, será apresentado alguns trabalhos relacionados, a fim de apresentar as pesquisas que estão sendo realizadas nessa área.

2. Trabalhos Relacionados

Diversos pesquisadores têm realizado estudos a respeito das vulnerabilidades em técnicas de randomização existentes. Além disso, algumas das pesquisas realizadas são voltadas para a exploração e utilização desses dados para uso nas mais variadas áreas de aplicação.

No trabalho de [Mathy Vanhoef 2016], são apresentadas formas de contornar a randomização dos endereços MAC, explorando dispositivos com suporte a função WPS (*Wi-Fi Protected Setup*). No trabalho é observado que durante a utilização da função WPS [2], algumas das informações fornecidas, representam uma falha no processo de mascara-

[1] A randomização do endereço MAC é uma técnica que visa impossibilitar o rastreamento do dispositivos, através dos rastros deixados pelas requisições de sondagem. Uma vez que o dispositivo utiliza endereços randômicos durante o processo de sondagem, em teoria, o mesmo não pode ser rastreado através do endereço MAC anunciado em *broadcast* [Jeremy Martin 2017].

[2] O WPS permite que um dispositivo realize uma conexão com um ponto de acesso, sem a necessidade de inserção das credenciais necessárias.

mento de endereço. O uso do WPS, se baseia na utilização de um endereço UUID-E (*Universally Unique IDentifier-Enrollee*), que é derivado do endereço MAC do dispositivo. Dessa forma, é possível obter o endereço real, a partir do identificador UUID-E. Uma vez feito isso, a randomização de endereços MAC torna-se ineficaz.

Em [Jeremy Martin 2017], os autores apresentam as técnicas de randomização utilizadas por alguns dos Sistemas Operacionais mais populares da atualidade, bem como as falhas encontradas em cada técnica. Nesse trabalho, são apresentadas duas contribuições principais. A primeira contribuição, está relacionada a identificação de uma nova vulnerabilidade no processo de randomização, onde é explorada uma falha em alguns *chipsets wireless*. A segunda contribuição consiste na utilização conjunta de algumas técnicas já conhecidas, para possibilitar o rastreio de dispositivos apesar da randomização. Os autores concluem que, apensar dos dispositivos utilizarem técnicas de randomização, muitas vezes, não são totalmente efetivas, no que diz respeito a impedir o rastreamento dos dispositivos.

[Mathieu Cunche 2016], explora as vulnerabilidades deixadas com base nos outros campos do *frame* de *probe request*, tanto no cabeçalho, quanto no conteúdo do mesmo. Uma está relacionada ao número de sequência dos *probes request* [3] emitidos, que não são reinicializados a cada randomização. Outra forma de rastreamento, seria através da obtenção de impressões digitais dos dispositivos, utilizando algumas informações como o IEs. Estes são blocos de dados que identificam os recursos suportados por uma estação. Além da análise de informações dos *frames*, é apresentado uma análise com base na frequência e no tempo com que as requisições são enviadas. Por fim, os autores concluem que apenas a randomização de endereços MAC não é capaz de garantir totalmente o anonimato pretendido. Além disso, citam a possibilidade da adoção futura de uma nova camada para a pilha de protocolos.

Tanto os trabalhos relacionados apresentados, como os analisados, apesar de apresentarem um foco ligeiramente diferente, têm em comum a opinião dos autores de que a randomização de endereços MAC não é efetiva em diversos casos, por motivos diversos. Nesse trabalho, pretende-se identificar, através de alguns testes, se os dispositivos Wi-Fi, cujos os SOs oferecem suporte, estão de fato implementando a randomização de endereços.

2.1. Métodos de Descoberta do endereço MAC verdadeiro

Utilizando determinados métodos, é possível identificar o dispositivo que se deseja coletar informações, descobrir seu endereço MAC original, e rastreá-lo. Um dos métodos, consiste em observar o *handoff* entre APs. Outro fator que pode ser levado em consideração, é a frequência de envio dos *probe requests*. Observados os padrões de requisições, é possível traçar perfis de comunicação dos dispositivos, e isolar o dispositivo alvo [Célestin Matte 2016].

Muitos dos APs atuais, possuem um protocolo chamado WPS (*WiFi Protected*

[3] *Probes request* são requisições enviadas quando um dispositivo encontra-se com a interface de rede *Wi-Fi* ativa, o mesmo envia mensagens em *broadcast* para qualquer outro dispositivo que esteja escutando. Estas mensagens de *probes request* têm como objetivo principal situar o dispositivo requisitante sobre os pontos de acessos ao seu alcance. Estas solicitações contém, entre outras coisas, o endereço MAC do mesmo.

Setup), que permite que dispositivos passem do estado de dissociado para associado sem que o usuário tenha que informar as credenciais. Para permitir isso, alguns campos extras são adicionados aos *probe requests*. Esses campos, além de conter o fabricante e o modelo do dispositivo, contém um identificador único utilizado para realizar uma conexão WPS. Esse identificador é derivado do endereço MAC, e a partir do conhecimento do identificador é possível reaver o MAC [Martin et al. 2017].

Outro método conhecido para a obtenção do endereço MAC real, é um ataque denominado *Karma* [Vanhoef et al. 2016] [Martin et al. 2017]. Ele baseia-se em configurar um roteador com o SSID *(Service Set IDentifier)* idêntico a um já conhecido pelo dispositivo. É comum que os dispositivos se conectem automaticamente em redes nas quais já se conectaram anteriormente. Isso faz com que o dispositivo se conecte no AP, e anuncie seu endereço MAC verdadeiro.

O algoritmo utilizado em [Célestin Matte 2016] se aproveita do fato de que os dispositivos costumam enviar *frames* de mensagens com a mesma frequência de tempo. Dessa maneira, é possível agrupar *frames* de mensagens de um dispositivo, através da análise da diferença de tempo entre mensagens. Assim, mesmo se ao longo do tempo, o dispositivo tenha seu MAC randomizado, ainda assim é possível descobrir sua identidade. Esse algoritmo obteve, em média, 77,2% de sucesso, na identificação de *frames* provenientes do mesmo dispositivo.

3. Metodologia

Com intuito de testar se os métodos de randomização de endereço MAC efetivamente omitem o endereço MAC real das interfaces WiFi, foram realizados testes capturando os *probes request* emitidos por alguns dispositivos WiFi de Sistema Operacionais diferentes. Os testes realizados foram pensados partindo do princípio de que os dispositivos, estando com sua interface de rede Wifi ativa, enviam requisições em busca de pontos de acesso próximos para uma possível conexão. Essas requisições não são destinadas a dispositivos específicos, mas sim em *broadcast*. Dessa forma, qualquer dispositivo ao alcance, pode interceptá-las. Ao interceptar os *probes requests*, foram comparados os endereços MAC obtidos, com os endereços reais dos dispositivo.

Os Sistemas Operacionais (SO) abordados nos testes apresentam randomização do endereço MAC. Cada um utiliza uma abordagem própria, conforme descrito a seguir.

- *Microsoft Windows*: dispositivos com sistema operacional Windows, a partir da versão 10, dão suporte à randomização do endereço MAC. Porém, para que a randomização seja aplicada, o dispositivo deve também possuir um hardware com suporte para tal procedimento. Até o momento, o Windows 10, é o único SO que oferece randomização para dispositivos associados a uma rede WiFi [Mathy Vanhoef 2016].
- Android: o Sistema Operacional Android oferece suporte à randomização de endereços MACs, para requisições de sondagens, a partir de sua versão 6.0. Para utilizar a randomização, é importante que o dispositivo ofereça suporte de hardware e drivers necessários. Algumas aplicações com suporte a alteração manual ou randômica de endereços MACs estão disponíveis para Android, porém é necessário privilégio de aceso de root ao dispositivo [Mathy Vanhoef 2016].

- iOS: este foi o primeiro SO a oferecer a randomização de endereço MAC para ser utilizado durante o processo de busca por APs ao alcance (sondagem). A partir da versão 8, os dispositivos iOS da *Apple* passaram a utilizar a randomização. Assim como no SO Android, abordado anteriormente, os endereços randômicos nesse SO, são utilizados apenas quando o dispositivo está fazendo a sondagem por redes sem fio ao alcance. Após a conexão com o AP, passa a ser utilizado o endereço real do dispositivo [Technologies 2015] [Jeremy Martin 2017].

3.1. Cenário de Testes

Os testes foram realizados na UFSM (Universidade Federal de Santa Maria), no prédio do CTISM (Colégio Técnico Industrial), em um ambiente sem qualquer modificação, ou seja, sujeito a interferências de sinais provindos tanto de outros dispositivos, como de outros pontos de acesso próximos.

Utilizou-se um roteador com o *Firmware OpenWRT* instalado, configurado em modo monitor, onde foram implementados *scripts* para obtenção dos *probes requests* emitidos pelos dispositivos. O *Script* utilizado na captura dos *probes* foi desenvolvido na linguagem *Shell Script* e utiliza as funcionalidades do *software TCPdump* para isso. Dentre os dados capturados estão o horário de captura, a intensidade do sinal e o endereço MAC do dispositivo.

Foram selecionados quatro dispositivos para a realização dos testes. Dentre eles um dispositivo com SO Windows versão 10, dois com o SO Android versão 5 e 7, e por fim um dispositivo com SO iOS versão 10. Inicialmente, foram identificados os endereços MACs originais de cada um dos dispositivos, disponíveis nas informações de configuração do sistema. Os primeiros testes foram realizados, com os dispositivos já conectados em um AP da rede do CTISM.

Foi definido um intervalo de cinco minutos para análise do comportamento do dispositivo, captando e analisando os *probe requests*. Após a coleta dos dados com dispositivo conectado, na segunda etapa, foram realizados testes com os dispositivos desconectados de qualquer rede WiFi, porém ainda com a interface ativa.

4. Resultados e Conclusão

A partir da coleta de dados, foi possível observar que o dispositivo com o Android 5, conforme esperado (por não possuir suporte à randomização), não apresentou uma randomização do seu endereço em questão. Foi possível identificá-lo através do endereço MAC, tanto conectado, quanto desconectado. Vale comparar a coleta de dados tanto de dispositivos que, em teoria, fornecem a possibilidade de randomização, com dispositivos que não possuem suporte. Dessa forma, pode-se observar as diferenças que existem em ambos, na prática.

O dispositivo com Android 7 e o dispositivo com Windows 10, apesar de possuírem suporte a randomização, não apresentaram nenhum tipo de alteração em seus endereços MACs reais. O mesmo endereço foi distribuído durante as requisições quando conectado ao AP e durante a realização de sondagem. Para que a randomização ocorra, é necessário estar implementada no software e no hardware do dispositivo. Embora ambos os SOs possuam suporte à randomização, possivelmente, o hardware de ambos não seja compatível com a funcionalidade.

Dos sistemas operacionais testados, o único que se mostrou capaz de realizar a randomização, conforme prometido pelo fabricante, foi o iOS 10. De acordo com os testes, apenas quando o dispositivo estava conectado, que o seu MAC verdadeiro foi divulgado.

Com a análise dos resultados dos testes obtidos, foi possível perceber que o dispositivo com iOS 10 não pode ser identificado por meio do mesmo endereço que distribuía quando desconectado. Já o de Android 5 apresenta um grande número de divulgações do MAC verdadeiro, o que aumenta a precisão da localização de um dispositivo, utilizando a potência do sinal das requisições. Apesar de apresentar uma diferença significativa do número de divulgações, a versão do Android 7, ainda apresenta falhas quanto à randomização, ou seja, continuou a divulgar seu endereço verdadeiro.

A privacidade do usuário é um fator que deve receber atenção das empresas desenvolvedoras de Sistema Operacional. Apesar dos esforços para prover maior privacidade aos usuários, os resultados mostram que a randomização do endereço MAC, aparentemente efetiva em teoria, ainda não é adotada de forma plena por todos os SOs, na prática.

Referências

Ali, S. and Khusro, S. (2016). Mobile phone sensing: A new application paradigm. *Indian Journal of Science and Technology*, 9(19).

Célestin Matte, Mathieu Cunche, F. R. M. V. (2016). Defeating mac address randomization through timing attacks.

de Sao Paulo, F. (2017). Número de smartphones em uso no brasil chega a 168 milhões. http://www1.folha.uol.com.br/mercado/2016/04/1761310-numero-de-smartphones-em-uso-no- brasil-chega-a-168-milhoes-diz-estudo.shtml. May 24, 2017.

Eastlake 3rd, D. and Abley, J. (2013). Iana considerations and ietf protocol and documentation usage for ieee 802 parameters. Technical report.

Jeremy Martin, Travis Mayberry, C. D. L. F. L. B. C. R. E. C. R. D. B. (2017). A study of mac address randomization in mobile devices and when it fails.

Martin, J., Mayberry, T., Donahue, C., Foppe, L., Brown, L., Riggins, C., Rye, E. C., and Brown, D. (2017). A study of mac address randomization in mobile devices and when it fails. *arXiv preprint arXiv:1703.02874*.

Mathieu Cunche, C. M. (2016). On wi-fi tracking and the pitfalls of mac address randomization.

Mathy Vanhoef, Célestin Matte, M. C. L. S. C. F. P. (2016). Why mac address randomization is not enough: An analysis of wi-fi network discovery mechanisms.

Patil, P. and Kokil, A. (2015). Wifipi-tracking at mass events. In *Pervasive Computing (ICPC), 2015 International Conference on*, pages 1–4. IEEE.

Technologies, Z. (2015). Analysis of ios 8 mac randomization on locationing.

Vanhoef, M., Matte, C., Cunche, M., Cardoso, L. S., and Piessens, F. (2016). Why mac address randomization is not enough: An analysis of wi-fi network discovery mechanisms. In *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*, pages 413–424. ACM.

