Autenticação baseada em ondas eletromagnéticas

Ivo Carlson, Max Gonzaga, Leonardo B. Oliveira, Heitor S. Ramos

[1]Departamento de Ciência da Computação – Universidade Federal de Minas Gerais (UFMG)
CEP 31.270-901 – Belo Horizonte – MG – Brasil

{ivo.carlson, leob, ramosh}@dcc.ufmg.br, maxgonzaga@ufmg.br

Resumo. *O crescente uso de dispositivos inteligentes traz preocupações quanto a segurança das informações trocadas. Métodos de Autenticação nos permitem verificar a procedência de uma mensagem. Um método utilizado para autenticar uma mensagem é utilizar a assinatura digital do dispositivo. Da mesma forma que os seres humanos possuem características que permitem distinguir uma pessoa de outra, dispositivos eletrônicos possuem diferenças provindas do processo de manufatura. Quando um dispositivo eletrônico é utilizado para gerar algum tipo de mídia, as diferenças podem ser vistas no resultado. Essas características vistas no sinal gerado podem ser utilizadas para extrair uma assinatura digital. Nesse trabalho Arduinos, dispositivos típicos de IoT, foram utilizados para gerar um sinal que foi observado por um rádio e estudado com o intuito de encontrar diferenças que poderiam levar à criação de uma assinatura para cada dispositivo e assim, desenvolver um método de autenticação entre dispositivos IoT e um servidor.*

Abstract. *The rising moment in the use of smart devices brings up concerns about the security of the information exchanged. Authentication methods allow us to verify the provenance of a message. A way of authenticating a message is using the fingerprint of the device. Like humans, smart devices presents differences between them, that comes from the impossibility of manufacture two identicals devices. Once this device is used to generate some kind of media, these differences can be seen on the result. Those features that come along with the generated media can be used to fingerprint the device. In this paper we used Arduinos, typical IoT devices, to generate a signal and a Software Defined Radio to listen to it. The signal was then studied looking for features that could be used to fingerprint each device and then, develop an authentication method between IoT devices and a server.*

1. Introdução

O atual aumento na utilização de dispositivos inteligentes eleva a preocupação quanto à segurança. Um pilar da segurança da informação é a autenticação, onde se verifica a legitimidade das informações trocadas. Um método de autenticar dispositivos é utilizando assinaturas digitais. Um ponto importante sobre dispositivos eletrônicos é a sua divergência a nível de hardware. Mesmo dispositivos que compartilham o mesmo fabricante, modelo e até lote de produção não são idênticos em todos os seus aspectos [Das et al. 2014]. Pequenas variações na resistência intrínseca dos transistores utilizados no processador levam a uma variação da corrente que circula no dispositivo e consequentemente no campo

elétrico e magnético gerados. Essas pequenas variações divergem entre dois ou mais dispositivos mas permanecem inalteradas quando o mesmo dispositivo é analisado múltiplas vezes de tal modo que um sinal gerado é único de cada dispositivo [Yilmaz et al. 2018]. Desse modo, os componentes eletrônicos são únicos e irreplicáveis. Em paralelo com o mundo real onde uma assinatura pode ser utilizada para distinguir um indivíduo essas pequenas variações podem ser utilizadas para diferenciar dispositivos e os sinais gerados pelos mesmos.

Ao utilizar um dispositivo eletrônico para gerar algum sinal, as características do dispositivo irão interferir no produto final. Quando se conhece um dispositivo e os sinais gerados por ele é possível isolar a interferência ou o ruído que ele gera. A partir desse ponto, o problema de autenticar um sinal se torna extrair o ruído adicionado pelo dispositivo e comparar com o que era esperado do mesmo. Ao explorar as diferenças entre microfones e alto-falantes, [Das et al. 2014] conseguiram extrair suas assinaturas e autenticar ou não um dispositivo baseando-se no sinal de áudio gerado. A mesma lógica pode se estender para dispositivos que trabalham com outros tipos de ondas como processadores e rádios, componentes que são facilmente encontrados em sistemas embarcados.

As diferenças entre os dispositivos entretanto vão além dos sinais e arquivos gerados. Tempo de processamento, ruídos e variações nas ondas geradas também são características de cada dispositivo. Em uma área similar, [Zajic and Prvulov 2014] utilizaram a onda eletromagnética gerada pela corrente no circuito para observar o comportamento do software em tempo de execução e encontraram uma periodicidade enquanto o programa executava. Esse trabalho busca medir a eficiência na execução de um software utilizando um hardware separado para que não houvesse paralelismo no processamento. Como um software utilizado para autenticação pode ser corrompido, tratar a autenticação em um nível mais baixo como o nível de hardware, dificulta a ação de adversários, visto que simular o comportamento exato de um hardware com seus mesmos ruídos é uma tarefa complexa.

Justificativa. Abordagens que não demandam muito processamento do dispositivo que se deseja autenticar são interessantes para dispositivos utilizados em IoT. Na área da criptografia, sistemas eficientes são desejados, mas não caso venham a diminuir a segurança. Esse detalhe pode ser um empecilho para o advento da Internet das Coisas, visto que muitos dispositivos que a compõem têm pouco poder computacional, para, por exemplo, executar um algoritmo de autenticação. Uma alternativa para isso é gerar uma assinatura utilizando características do dispositivo. Com essa abordagem, o poder computacional do dispositivo pode ser dimensionado para que ele execute suas tarefas corriqueiras, visto que o processo de gerar a assinatura é independente. Desse modo, um servidor externo observa o comportamento do dispositivo e o modo como ele executa suas funções e o autentica baseado nessas observações.

Objetivos. Nesse trabalho nós temos como objetivo desenvolver uma metodologia para obter, interpretar e tratar ondas eletromagnéticas provenientes do funcionamento de um dispositivo eletrônico e utilizar essas ondas para classificar dispositivos. Mais precisamente, nós pretendemos criar um método para analisar as ondas geradas pelo dispositivo e explorar as características que são passadas para elas, de modo a utilizá-las para autenticar ou não um dispositivo.

Este método utiliza dispositivos exteriores para captar e processar o sinal recebido para não demandar muito poder de processamento por parte do gerador do sinal podendo ser utilizado por dispositivos com baixo poder computacional como é o caso de diversos componentes IoT.

2. Trabalhos relacionados

Canal Secreto e Canal Lateral (Covert/Side-channel) são canais de comunicação produzidos de forma não intencional durante a execução de um programa por um circuito eletrônico. Uma forma de se observar esses canais é pela emanação de ondas Eletromagnéticas (EM) produzidas. No trabalho [Yilmaz et al. 2018] os pesquisadores introduziram uma forma de medir quanta informação é transmitida por esses canais. A metodologia aplicada relacionava matematicamente a energia despendida pelo canal secundário quando executando instruções individuais e a alteração provocada no comportamento desse sinal. Com essa relação, um método proposto era aplicado e utilizado para avaliar a capacidade de obter informações em sistemas reais. Com a execução do trabalho, os autores apresentaram uma nova metodologia para medir quanta informação pode ser obtida e apresentaram observações relevantes sobre o sinal emitido por cada instrução de acordo com a sua complexidade e tempo de execução.

[Sehatbakhsh et al. 2018] apresenta uma metodologia para garantir a segurança de dispositivos médicos IoT. Nas simulações, foi usada uma bomba de infusão – aparelho eletrônico que injeta ou remove fluidos. A aplicação recebe como parâmetro uma direção (injetar ou sugar) e a quantidade de líquido que será injetada ou sugada. Devido a limitações de hardware e software, é inviável implementar uma aplicação anti-malware no dispositivo. Fez-se necessário, então, que o monitoramento se desse externamente. Como é sabido, todo aparelho eletrônico emite involuntariamente ondas eletromagnéticas, as quais possuem um comportamento que depende da atividade que está sendo realizada no processador do eletrônico. A estratégia traçada pode ser resumida em duas fases. A primeira consiste em coletar o sinal eletromagnético da bomba de infusão quando ela não está sob o efeito de uma aplicação maliciosa. A segunda fase consiste em comparar constantemente o sinal eletromagnético e decidir se há alguma atividade anômala no dispositivo.

No trabalho de [Zajic and Prvulov 2014] mostra-se como a obtenção de dados via vazamento de ondas eletromagnéticas pode ser facilmente alcançada. Como isso coloca em risco a segurança dos dados do usuários, faz-se necessário analizar radiação eletromagnética emitida por um dispositivo quando da execução de uma aplicação e, com isso, propor formas de minimizar o vazamento de informações. Com isso em mente, o artigo citado descreveu os experimentos realizados para captar as ondas geradas por três computadores diferentes ao executar um trecho de código. Constatou-se que é possível captar as informações desde centímetros até poucos metros de distância do processador.

O trabalho feito por [Das et al. 2014] propõe um método de autenticar dispositivos utilizando componentes acústicos. Como microfones e auto-falantes são diferentes entre si, ele podem ser uma fonte para a obtenção de assinaturas extrínsecas visto que irão gerar sinais diferentes. Analisando a onda de som salva em um arquivo WAV os autores levantaram uma lista de características que podem ser obtidas pela leitura do arquivo. Utilizando essas características como entrada, um algoritmo de classificação tenta enquadrar

o áudio em uma lista de dispositivos conhecidos obtida no processo de aprendizagem. Nesse trabalho, os autores variaram diversos parâmetros como distância do servidor e ruídos no ambiente e também testaram dispositivos de diversos fabricantes obtendo uma média de sucesso de 98%.

Verificar a autenticidade de uma mensagem pode incluir observar o local e/ou a hora de sua criação. O método apresentado por [Hajj-Ahmad et al. 2015] propõe utilizar a frequência da rede elétrica (ENF) e a interferência gerada por ela para reduzir as possíveis localidades de gravação de um sinal de mídia. Depois de gravar os sinais em diferentes locais os autores observaram que a interferência causada pela rede elétrica fica presente no sinal gerado. Como o comportamento de cada rede é diferente, seja em relação a tensão nominal, seja em relação ao modo como a tensão varia, isso constituía um problema clássico classificação em aprendizado de máquina. Utilizando um algoritmo treinado a partir do comportamento de cada rede, eles desenvolveram um método para identificar o local de origem do sinal e obtiveram uma precisão de 88,4%.

3. Desenvolvimento

Existe um grande número de dispositivos inteligentes no mundo e esse número tende a aumentar. Esses dispositivos podem ser utilizados para garantir acesso a produtos ou serviços, essa funcionalidade traz consigo a necessidade de verificar se o dispositivo que requer o acesso é realmente quem ele diz ser. Confirmar que a mensagem recebida foi gerada pelo cliente ou servidor esperado é chamado autenticar e constitui um dos pilares da segurança de informação. Vários métodos de autenticação utilizados atualmente pelos mais inúmeros dispositivos são baseados em softwares onde um programa segue um sequência de passos para determinar a validade de uma mensagem recebida. Esses métodos convencionais de autenticação, portanto, demandam que o cliente ou serviço a ser autenticado possua certo poder de processamento o que pode inviabilizar a sua aplicação em certos dispositivos.

Um ponto importante sobre dispositivos eletrônicos é a sua divergência a nível de hardware. Mesmo dispositivos que compartilham o mesmo fabricante modelo e até lote de produção não são idênticos em todos os seus aspectos [Das et al. 2014]. Pequenas variações na resistência intrínseca dos transistores utilizados no processador levam a uma variação da corrente que circula no dispositivo e consequentemente no campo elétrico e magnético gerados. Essas pequenas variações divergem entre dois ou mais dispositivos mas permanecem inalteradas quando o mesmo dispositivo é analisado múltiplas vezes de tal modo que um sinal gerado é único de cada dispositivo. Em paralelo com o mundo real onde uma assinatura pode ser utilizada para distinguir um indivíduo essas pequenas variações podem ser utilizadas para diferenciar dispositivos e são chamadas de Assinaturas Digitais.

Ao extrair uma assinatura digital que provém do uso do dispositivo, um servidor pode autenticar um componente sem interferir em seu funcionamento normal por meio da análise do sinal emitido e das variações geradas pelo processador nesse dispositivo.

Ambiente. O ambiente utilizado foi um laboratório funcional, simulando uma aplicação genérica e corriqueira do problema. Contendo ruídos de diversas fontes como lâmpadas fluorescentes, ar-condicionado e sinas de diversos dispositivos computacionais. O problema estudado nesse momento foi medir a eficiência da captação do sinal assu-

mindo várias distâncias da fonte e do receptor e dois modelos de obtenção do sinal. Foram utilizados 7 Arduinos que executavam o mesmo código. Um módulo de rádio definido por software (SDR) ligado a um computador foi utilizado para capturar o sinal gerado por cada Arduino. A distância do Arduino para a antena foi variada para se avaliar a robustez da captura do sinal

Sinal. O sinal gerado pelos Arduinos foi um sinal eletromagnético proveniente de movimento dos elétrons no processador, vazados pelos canais lateral e secundário e capturados por um rádio definido por software. O rádio foi configurado para operar na mesma frequência da onda gerada. O sinal foi armazenado de duas formas: (i) de maneira bruta, onde os bits da entrada eram escritos diretamente no arquivo e (ii) com um filtro passa baixa que foi utilizado para tentar eliminar parte dos ruídos do ambiente.

Algoritmo. Dos sinais obtidos foram extraídas características clássicas de séries temporais. Essas características foram passadas para um algoritmo de aprendizado de máquina que usava um classificador KNN para tentar enquadrar cada um dos sinais de entrada em 7 classes, que correspondiam ao dispositivo que os gerou.

4. Metodologia

O sinal dos dispositivos foi observado utilizando um rádio definido por software (SDR) do modelo RTL2832U, mostrado na Figura 1, um dispositivo similar a um rádio comercial mas com componentes implementados em software. Esse SDR foi ligado a um laptop Dell Vostro e capturou o sinal com o auxílio de dois softwares, o GQRX[1] e o GNURadio[2]. Nesse experimento sete Arduinos Mega 2560, mostrado na Figura 2, simularam os dispositivos a serem autenticados. Esses Arduinos executaram um código que possuia um laço simples, composto por comandos que permitem maior vazamento de informação pelos canais laterais e secundário. O laço foi dimensionado para que o seu período de execução levasse à geração de um sinal cuja frequência estivesse na faixa de operação do SDR. Nesse caso abordado, o Arduino conectado ao computador é tratado como emissor e o SDR conectado ao computador é tratado como receptor.

Figura 1. Rádio SDR

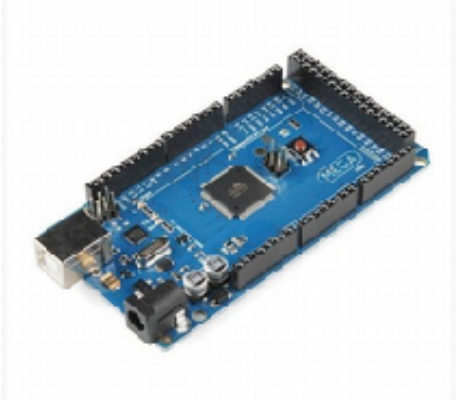

Figura 2. Arduino Mega

Como visto em trabalhos anteriores o período de execução de um laço é similar ao período da onda eletromagnética gerada pelo trânsito de elétrons no processador. Tendo

[1]http://gqrx.dk/
[2]https://www.gnuradio.org/

isso em mente, o tempo médio que o Arduino levava para executar o programa foi medido e o programa foi alterado até que a duração da execução estivesse dentro da faixa de operação do SDR. Conforme o Arduino executava o novo programa, o sinal gerado podia ser visto na forma de um espectrograma no display do GQRX como na Figura 5. Nesse mesmo software, em uma representação da Frequência do Sinal pelo Tempo, podiam ser observados picos de atividade em uma determinada frequência que variava de cada dispositivo vide Figuras 3 e 4. Isso caracterizou diferenças entre os dispositivos no domínio da frequência e no domínio do tempo.

Com o pico de atividade observado no domínio da frequência, Figura 3 e Figura 4, o sinal bruto gerado por cada um dos Arduinos foi salvo utilizando o GNURadio centrado no pico de atividade observado no display do GQRX com um filtro de largura 5MHz. Foram coletadas amostras de todos os 7 dispositivos durante um intervalo de execução de 10 segundos. Pelas Figuras 3 e 4 é possível notar a diferença na frequência de operação de dois dispositivos, essa observação, mais tarde, encorajou a busca de características de sinais no domínio da frequência.

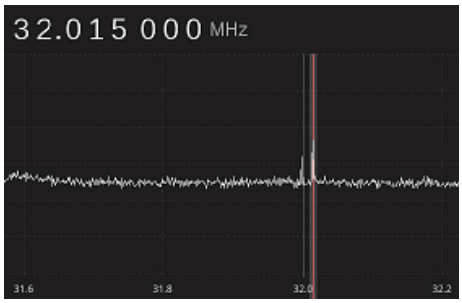

Figura 3. Frequência Arduino 1

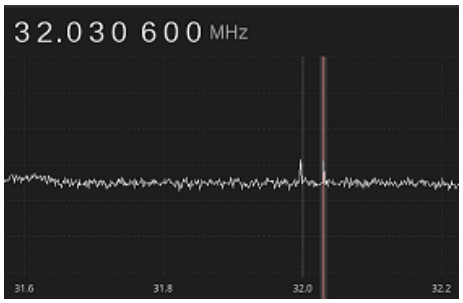

Figura 4. Frequência Arduino 2

O sinal completo visto no display do GQRX pode ser observado na Figura 5 na forma de um espectrograma. Nessa representação, a tonalidade mais escura, num tom mais forte de amarelo, representa uma maior amplitude e informa onde existe um maior fluxo de dados. As cores mais claras como o verde e o azul caracterizam uma amplitude menor e foram tratadas como ruídos do ambiente. O eixo horizontal informa a qual frequência do espectro correspondem essas informações e o eixo vertical caracteriza o tempo. O sinal salvo para o experimento era o sinal bruto obtido pelo módulo USB visto na Figura 1, convertido para o formato de números complexos do GNURadio, um variante do IEEE-754, re-amostrado e com sua largura de banda reduzida para aproximadamente 96KHz. Cada arquivo continha, então, um vetor de números complexos de 64 bits na forma de dois números reais de 32 bits.

Os sinais coletados de cada Arduino apresentavam uma média de 3.000.000 de pontos, uma quantidade muito grande de informações. Esses sinais foram divididos em porções de 1/100 e 1/200 do tamanho total, uma vez que foi observado que o Python não lidava bem com frações menores. Dessas frações de sinal foram extraídas 12 características comuns de sinais, 9 no domínio do tempo e 3 no domínio da frequência.

Características. As características no domínio do tempo foram:

1) Root Mean Square ou RMS, também conhecida como potência do sinal, calculado utilizando:

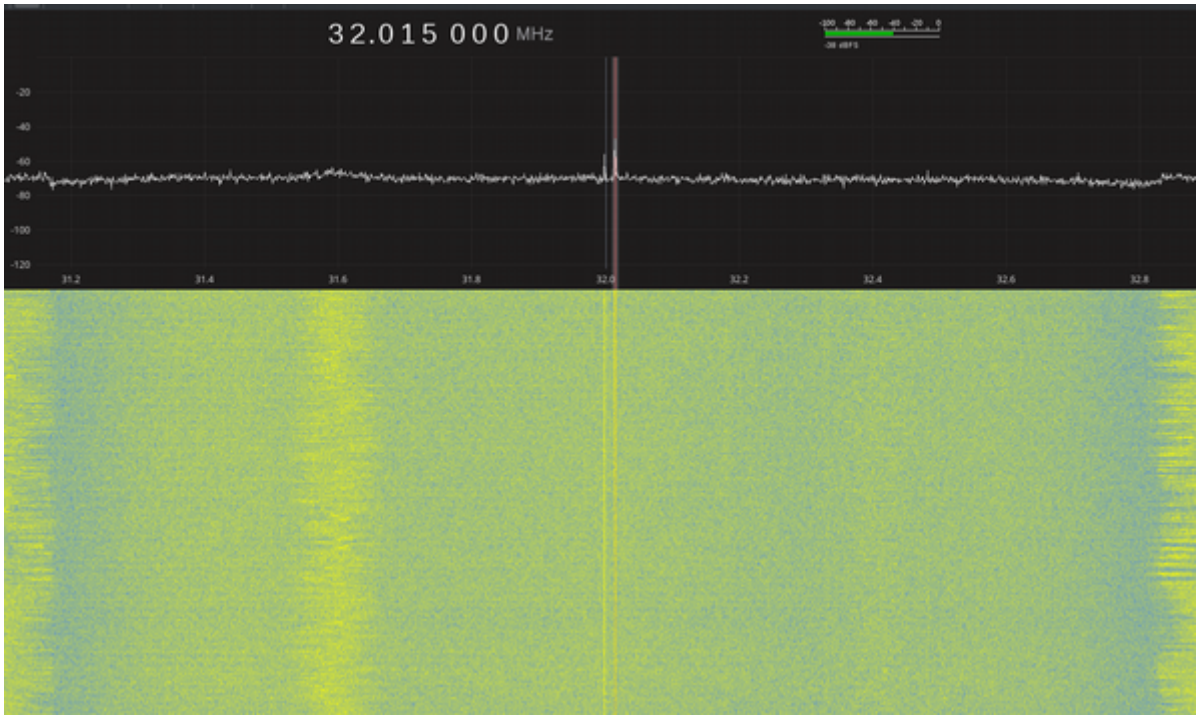

Figura 5. Espectrograma do Sinal

$$RMS = \sqrt{f(x) * \bar{f}(x)} \tag{1}$$

com $\bar{f}(x)$ sendo o conjugado do sinal complexo.

2) Valor máximo do sinal Max,

3) Valor mínimo do sinal Min,

4) Média dos valores Mean,

5) Mediana dos valores, Median

6) Variância do sinal Var,

7) Entropia de Shannon, que representa o grau médio de incerteza intrínseco à fonte, calculado com:

$$H = -\sum_{i=1}^{n} p_i log_e p_i \tag{2}$$

com p_i sendo a probabilidade do *i-ésimo* resultado.

8) Kurtosis, que representa o formato da distribuição em comparação com uma Gaussiana e é calculado com :

$$K = \frac{n \sum_{i=1}^{n} (x_i - \bar{x})^4}{(\sum_{i=1}^{n} (x_i - \bar{x})^2)^2} - 3 \tag{3}$$

com n sendo o número de pontos do sinal e $\bar{x}$ sendo a média da amplitude de todos os pontos.

9) Skewness, que é referida usualmente como uma medida da simetria do sinal ou como o grau de distorção da Gaussiana, é calculado segundo:

$$S = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^3}{ns^3} \tag{4}$$

com s sendo o desvio padrão da amostra.

Como o sinal possuia variações visíveis na frequência quando analisado o seu espectrograma, também foram utilizadas características no domínio da frequência. Para encontrar essas caracterísitcas, primeiro calculamos o vetor m como a transformada de Fourier do sinal original. Nesse vetor, cada componente m_i possui a energia/magnitude do *i-ésimo* componente de frequência do espectro. Com esse vetor calculado, encontramos a

10) Spectral Centroid, que representa o centro de massa do sinal e é dada por:

$$\mu = \frac{\sum_{i=1}^{n} f_i m_i}{\sum_{i=1}^{n} m_i} \tag{5}$$

onde m_i representa a magnitude do *i-ésimo* componente e f_i representa a frequência daquela amostra.

11) Spectral Entropy, que informa o quão puntiformidade é o sinal, pode ser encontrado transformando-se o sinal no espectro da frequência em uma função de probabilidade de massa (PMF) por meio de normalização do espectro:

$$w_i = \frac{m_i}{\sum_{i=1}^{n} m_i} \tag{6}$$

Com essa distribuição de massa, encontramos a Spectral Entropy utilizando:

$$\phi = \sum_{i=1}^{n} w_i log_2 w_i \tag{7}$$

a ideia central dessa característica é capturar as frequências do espectro nas quais existe um pico.

Por fim, o 12) Spectral Spread quantifica o afastamento do sinal do seu centro e é calculado por:

$$\sigma = \sqrt{\sum_{i=1}^{n} [f_i - \mu]^2 * w_i} \tag{8}$$

Com essas características extraídas os dados obtidos foram utilizados como entrada de um algoritmo de aprendizado de máquina implementado em Python utilizando ferramentas do scikit-learn[3]. O algoritmo utilizou um classificador KNN para criar grupos e por meio de cross validation, separou os dados de entrada em grupos que foram utilizados alternadamente para treinar e testar a taxa de acerto do procedimento.

[3]https://scikit-learn.org/stable/

5. Resultados

A acurácia de cada uma das 5 execuções que compuseram o cross-validation foi utilizada para se calcular a média de sucesso de classificação de cada um dos dispositivos em cada uma das distâncias de afastamento entre o emissor e o receptor. Esses dados são mostrados na Tabela 1.

Tabela 1. Média das 5 execuções do cross-validation

Distância	1/100 Bruto	1/100 Filtrado	1/200 Bruto	1/200 Filtrado
10 cm	0.4571	0.4286	0.3571	0.5286
20 cm	0.5000	0.3286	0.3714	0.3571
30 cm	0.3286	0.2857	0.3857	0.1143
40 cm	0.3714	0.2286	0.2000	0.2286
50 cm	0.5143	0.2286	0.3000	0.2429
Média	0.4343	0.3000	0.3229	0.2943

Como podemos analisar pelos dados coletados, os sinais que passaram por um filtro passa baixa apresentaram um pior desempenho em geral, o que leva a crer que parte da informação capturada pelo SDR como constituinte do sinal apresenta frequências elevadas. Também pela análise da tabela, é possível notar, em valores percentuais, que quando um pedaço menor do sinal é utilizado a taxa de sucesso é menor, o que já era esperado.

Uma observação interessante é que, ao contrário do esperado, distâncias maiores entre o receptor e o emissor dos sinais não tiveram grande interferência no resultado final, apesar de que os resultados obtidos com as distâncias de 10 cm e 20 cm mantém uma constância maior.

O melhor resultado foi encontrado ao se utilizar uma fração de 1/100 do sinal bruto, o que mostra que intervalos maiores e menos filtrados contém mais informações do dispositivo. Uma sequência possível para esse trabalho é utilizar porções maiores do sinal aliados a uma apuração das características mais significativas.

6. Conclusão

Este trabalho explora efeitos secundários de todos os circuitos elétricos e os utiliza como fonte de informação para um método de separação em classes e propõe uma sequência de operação que pode ser utilizada para construir um método de obtenção de sinais e reconhecimento de dispositivos.

A taxa de sucesso de 43% obtida com o sinal bruto e com 1/100 da amostra coletada em 10 segundos mostra que é possível separar os sinais utilizando essa técnica, visto que a taxa de acerto ficou acima do aleatório.

Como trabalhos futuros, pretendemos refinar o método. Como o escalamento temporal é proporcional a divisão de amostras, o sinal utilizado provinha de 0,1 segundo de coleta de dados, que é um intervalo bem reduzido. Utilizar uma amostra de sinal maior, fazer um refinamento das características utilizadas para encontrar as mais relevantes e expandir o número de Arduinos para aumentar a eficiência e a robustez do método.

Referências

Das, A., Borisov, N., and Caesar, M. (2014). Do you hear what i hear?: Fingerprinting smart devices through embedded acoustic components. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, CCS '14, pages 441–452, New York, NY, USA. ACM.

Hajj-Ahmad, A., Garg, R., and Wu, M. (2015). Enf-based region-of-recording identification for media signals. *IEEE Transactions on Information Forensics and Security*, 10(6):1125–1136.

Sehatbakhsh, N., Hong, H., Lazar, B., Johnson-Smith, B., Yilmaz, O., Alam, M., Nazari, A., Zajic, A., and Prvulovic, M. (2018). Syndrome: Spectral analysis for anomaly detection on medical iot and embedded devices-experimental demonstration.

Yilmaz, B. B., Callan, R. L., Prvulovic, M., and Zajić, A. (2018). Capacity of the em covert/side-channel created by the execution of instructions in a processor. *IEEE Transactions on Information Forensics and Security*, 13(3):605–620.

Zajic, A. and Prvulov, M. (2014). Experimental demonstration of electromagnetic information leakage from modern processor-memory systems. *IEEE Transactions on Electromagnetic Compatibility*, 56(4):885–893.