Autenticação de Usuários com Wi-Fi CSI: Uma Fonte Estratégica para Inteligência

Eduardo Fabrício Gomes Trindade[1], Felipe Silveira de Almeida[1] e Lourenço Alves Pereira Junior[1]
[1]Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP - Brasil

***Resumo*— O Channel State Information (CSI) do Wi-Fi tem sido amplamente estudado em atividades de sensoriamento. No entanto, implementações práticas voltadas para distinguir usuários com base no gênero são pouco exploradas. Este estudo propõe o uso de dispositivos Raspberry Pi para coleta de dados CSI num ambiente controlado e aplica aprendizado supervisionado para diferenciar usuários homens e mulheres com base em suas características biofísicas e comportamentais. A pesquisa também explora o uso dos filtros de Hampel e Savitzky-Golay (SG) no pré-processamento dos dados e compara o desempenho de algoritmos de classificação, com destaque para o K-Nearest Neighbors *(KNN)*, com acurácia de 99,99% na distinção de usuários, servindo também como fonte de dados para inteligência em atividades de proteção cibernética.**

***Palavras-Chave*— Channel State Information (CSI), autenticação, inteligência.**

I. INTRODUÇÃO

Com o passar dos anos, os sistemas de segurança que utilizam reconhecimento têm avançado substancialmente no que diz respeito à autenticação de usuários e restrição de acesso, especialmente para proteger ambientes e dados sensíveis. No entanto, o crescimento das ameaças cibernéticas tem desafiado os métodos tradicionais de autenticação, como senhas, biometria e reconhecimento facial. Além disso, a perda de dispositivos como tokens, cartões e QR Codes facilita ações fraudulentas, comprometendo a segurança. Diante disso, é imperativo desenvolver novos métodos de autenticação que ofereçam uma abordagem passiva, permitindo monitorar a presença e os gestos dos usuários sem necessidade de interação direta, aumentando assim a conveniência, segurança e confiabilidade.

A [1] enfatiza a necessidade de medidas de autenticação robustas para a segurança cibernética, com o objetivo de proteger infraestruturas críticas e informações sensíveis contra ameaças. Nesse contexto, a utilização de dados do Channel State Information (CSI) do Wi-Fi para sensoriamento tem emergido como uma área promissora de pesquisa, seguindo a linha das novas perspectivas militares globais [2]. Inicialmente, essa tecnologia foi implementada para ajustar o sinal conforme as variações ambientais, resultando em transmissões mais eficientes e confiáveis. No entanto, estudos recentes indicam que os dados CSI podem mapear o ambiente eletromagnético, possibilitando a identificação de anomalias no sinal e a autenticação de usuários, servindo também como fonte de dados para inteligência.

Este estudo apresenta uma nova abordagem utilizando dispositivos Raspberry Pi para coletar dados CSI em um ambiente controlado e aplicando aprendizado supervisionado para distinguir usuários masculinos e femininos com base em suas características biofísicas e comportamentais. Alinhado à [3], esta pesquisa visa fornecer uma camada adicional de segurança e de análise para atividades de inteligência.

Além disso, a [4] estabelece diretrizes focadas na proteção do conhecimento, identificação de vulnerabilidades e suporte à segurança das infraestruturas críticas. Este estudo se insere nesse contexto ao propor um método de autenticação baseado em CSI que pode ser utilizado para monitorar e identificar usuários em ambientes sensíveis, contribuindo para a segurança e defesa cibernética do país.

Assim, este manuscrito investiga a viabilidade do uso de dispositivos portáteis e de baixo consumo energético para a coleta de dados CSI, e propõe uma abordagem baseada em padrões biofísicos dos indivíduos, complementada por técnicas de *machine learning* e algoritmos de classificação. O artigo avança o estado da arte ao sugerir um novo método de controle de acesso com autenticação dos usuários utilizando características individuais extraídas do Wi-Fi, alinhado às determinações do Ministério da Defesa (MD). Até onde sabemos, este é o primeiro estudo a utilizar dados Wi-Fi CSI coletados por Raspberry Pi, baseando-se em *features* biofísicas e especificamente voltado para controle de acesso físico em um contexto de inteligência militar. As contribuições do estudo são:

1) *Dataset* rotulado com 5 atividades distintas, compreendendo silhuetas, gestos e leitura labial;
2) Proposta de um sistema de controle de acesso baseado em características biofísicas, extraídas do Wi-Fi com Raspberry Pi; e
3) Avaliação de desempenho dos algoritmos *Support Vector Machine (SVM)*, *Random Forest (RF)*, *K-Nearest Neighbors (KNN)*, *Árvore de Decisão J48* e *Naive Bayes (NB)* para diferenciação entre homens e mulheres.

O restante deste estudo está estruturado da seguinte forma: a Seção II. discute trabalhos anteriores estabelecendo o contexto para nossa pesquisa. A Seção III. descreve a metodologia adotada, detalhando as abordagens e ferramentas utilizadas. Na Seção IV., apresentamos os experimentos conduzidos, explicando cada passo da execução. A Seção V. é dedicada à análise e discussão dos resultados obtidos, explorando suas principais implicações. Por fim, a Seção VI. conclui o estudo, resumindo os principais achados e propondo trabalhos futuros.

II. TRABALHOS RELACIONADOS

As abordagens de autenticação se concentram principalmente em métodos baseados em padrões, modelos matemáticos e aprendizado profundo, cada um apresentando vantagens e desafios específicos.

[1] {trindade,felipefsa,ljr}@ita.br. Este trabalho foi parcialmente financiado pelo Programa de Pós-graduação em Aplicações Operacionais—PPGAO/ITA, pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) processo #2020/09850-0 e #2022/00741-0, pelo CNPq e pela CAPES.

A. Baseado em Padrões

A autenticação baseada em padrões explora variações no CSI causadas por comportamentos humanos. Em [5] propuseram um sistema de autenticação de dois fatores (2FA) utilizando dados CSI de redes Wi-Fi para verificar a proximidade física entre dispositivos. Contudo, essa abordagem depende da associação do indivíduo ao seu próprio dispositivo wireless ou a um dispositivo próximo, necessitando de equipamento adicional, o que limita sua aplicabilidade em ambientes de segurança militar.

A pesquisa [6] inspira nosso trabalho ao transformarem dados CSI em espectrogramas para traçar um perfil do movimento humano, semelhante aos gerados por radares Doppler. Já [7] propuseram uma abordagem para autenticação de atividades humanas baseada na variação do sinal Wi-Fi. Entretanto, ambos os estudos utilizaram sensores para aprender e caracterizar atividades, explorando a faixa de frequência de 2.4GHz, com menor granularidade na coleta dos dados CSI, diferente deste estudo, que utiliza Raspberry Pi na faixa de 5GHz.

B. Baseado em Modelos

O reconhecimento baseado em modelos utiliza a matemática ou a física para descrever e interpretar as variações do sinal causadas pelo comportamento humano. [8] foram pioneiros na construção de modelos que quantificam a correlação entre a dinâmica dos valores CSI, velocidades de movimento humano e partes do corpo durante atividades específicas.

Os trabalhos [9], [10] e [11] propuseram sistemas de reconhecimento de atividade humana e autenticação multiusuário usando dispositivos Wi-Fi comerciais. Esses sistemas utilizam modelos de detecção baseados em difração, Tempo de Chegada (ToA) e Ângulo de Chegada (AoA) para distinguir usuários. No entanto, são sensíveis às variações ambientais e exigem configurações específicas de hardware, o que pode comprometer sua precisão em ambientes militares e de inteligência.

Os estudos de [12] e [13] exploraram métodos de autenticação em ambientes multiusuários usando normalização para obter AoAs e estimativas de deslocamento Doppler para distinguir atividades humanas. No entanto, ambos enfrentaram desafios: [12] realizaram suas pesquisas em simulações com dados fictícios, ignorando variações e ruídos de ambientes reais, enquanto [13] encontraram dificuldades em diferenciar movimentos com efeitos Doppler quase idênticos, mesmo após retreinamento, comprometendo a confiabilidade na capacidade de autenticação.

C. Baseado em Aprendizado Profundo

O aprendizado profundo pode aprender e extrair automaticamente características significativas dos dados de entrada, eliminando a necessidade de extração manual. [14] revisaram técnicas como redes neurais recorrentes (RNN) e de memória de longo e curto prazo (LSTM), demonstrando melhorias no desempenho.

Aproveitando esse estudo, [15] e [16] propuseram abordagens de segmentação automática e higienização de ruído em dados CSI brutos para reconhecimento preciso da atividade humana. No entanto, essas técnicas enfrentam desafios como complexidade computacional, necessidade de grandes volumes de dados para treinamento eficaz e sensibilidade às variações no ambiente físico, comprometendo a robustez do sistema em cenários reais.

Já [17] e [18] aplicaram aprendizado profundo ao comportamento físico do usuário capturado pelo CSI do Wi-Fi para identificar usuários legítimos e distinguir falsificadores. No entanto, essas pesquisas focaram no comportamento do usuário e no ritmo de digitação, deixando brechas para tentativas de imitação do comportamento físico dos usuários autorizados, tornando essas abordagens inadequadas para aplicações de controle de acesso.

No contexto militar, o emprego de tecnologias inovadoras para proteção de infraestruturas e para a segurança nacional tem grande destaque, conforme a [1] e [4]. Este estudo se destaca pela simplicidade da proposta, utilizando dispositivos Raspberry Pi para coletar dados de Wi-Fi na frequência de 5GHz e empregando *features* biofísicas dos usuários. Diferentemente de abordagens que se baseiam em atividades realizadas, nossa metodologia foca nas características intrínsecas dos indivíduos para autenticação via Wi-Fi. A diferenciação entre homens e mulheres é particularmente relevante em ambientes militares, onde a segurança personalizada, a análise comportamental detalhada e a detecção precisa de ameaças são fundamentais para a proteção de informações sensíveis e infraestruturas críticas.

Por fim, a Tabela 1 confronta os manuscritos relacionados com o atual estudo e ressalta que ainda há espaço para a implementação de um sistema de controle de acesso a partir de dados CSI. Destaca-se que, até onde se tem conhecimento, o modelo proposto neste trabalho é o primeiro a utilizar características combinadas e extraídas do CSI do Wi-Fi por dispositivos com baixo consumo energético no contexto de controle de acesso, apresentando-se como uma proposta complementar de segurança.

III. CHANNEL STATE INFORMATION (CSI)

Esta seção fornece uma visão detalhada sobre as propriedades do Channel State Information (CSI), os fundamentos da autenticação de usuários e a arquitetura do sistema proposto para controle de acesso.

A. Propriedades do CSI

O CSI do Wi-Fi registra como os sinais sem fio se propagam do transmissor ao receptor, detalhando o comportamento das ondas eletromagnéticas em várias frequências. Essas informações incluem a amplitude e a fase do sinal, que podem ser modificadas por reflexões, obstruções e outros elementos no ambiente. Cada componente do CSI reflete o impacto do ambiente na propagação do sinal, criando uma função chamada Resposta em Frequência do Canal (CFR). Em sistemas Wi-Fi com múltiplas antenas, o espectro é dividido em várias subportadoras através da técnica OFDM (Orthogonal Frequency Division Multiplexing). O transmissor emite sinais de treinamento no início da transmissão, permitindo que o receptor estime como o canal Wi-Fi afeta cada subportadora. Isso ajuda o receptor a adaptar a transmissão e recepção dos dados para condições ótimas.

De acordo com [19], a matriz CSI, mostrada na Fig. 1, é um conjunto tridimensional de valores complexos estimados

TABELA 1: COMPARATIVO ENTRE TRABALHOS RELACIONADOS.

Estudo	Abordagem	Dispositivo de Coleta	Frequência utilizada / Largura de banda	Ferramenta de extração CSI	Aplicação em controle de acesso
[5]	Padrões	Notebook	2.4GHz (20MHz)	Não especificado	Não
[6]	Padrões	Notebook	5GHz (80MHz)	Linux 802.11n CSI Tool	Não
[7]	Padrões	Notebook	2.4GHz (20MHz)	Não especificado	Não
[8]	Modelos	Par de Antenas	5.24GHz	Não especificado	Não
[9]	Modelos	Par de Antenas	5.24GHz	Não especificado	Não
[10]	Modelos	Par de Antenas	5.24GHz	Linux 802.11n CSI Tool	Não
[11]	Modelos	Notebook	2.4GHz até 70MHz e 5GHz (até 200MHz)	Atheros CSI-Tool	Não
[12]	Modelos	MATLAB	-	-	Não
[13]	Modelos	Netgear X4S AC2600	5GHz (80MHz)	Nexmon CSI Tool	Não
[15]	Aprendizado Profundo	Notebook	Não especificado	Linux 802.11n CSI Tool	Não
[16]	Aprendizado Profundo	-	-	TP-Link C7A4	Não
[17]	Aprendizado Profundo	Mini PC	2.4GHz (20MHz) e 5GHz (80MHz)	Linux 802.11n CSI Tool	Não
[18]	Aprendizado Profundo	Mini PC	5GHz (80MHz)	Linux 802.11n CSI Tool	Não
Presente estudo	**Padrões**	**Raspberry Pi**	**5GHz (80MHz)**	**Nexmon CSI Tool**	**Sim**

pelo receptor a partir do sinal recebido. Este processo envolve a remoção do prefixo cíclico, desmapeamento e demodulação OFDM. Na prática, o CSI medido é influenciado por canais de multipercurso, processamento nos transmissores e receptores, além de inconsistências de hardware e software. A representação do CSI no domínio *baseband* é uma abstração complexa que inclui esses fatores, como deslocamentos cíclicos e variações nas amostragens de tempo e frequência.

A série temporal das matrizes CSI captura as mudanças no canal MIMO ao longo do tempo, frequência e espaço. Para um canal MIMO-OFDM com (M) antenas transmissoras, (N) antenas receptoras e (K) subportadoras, a matriz CSI forma um cubo de dados, expresso como $(H \in C^{(N \times M \times K \times T)})$. Este cubo registra a atenuação da amplitude e o deslocamento da fase dos sinais ao percorrerem múltiplos caminhos. O CSI oferece uma riqueza de informações superior a outras métricas, como o Indicador de Força do Sinal Recebido (RSSI). No contexto militar, essas propriedades detalhadas do CSI são essenciais para desenvolver sistemas de autenticação robustos e precisos, capazes de diferenciar usuários com base em características biofísicas únicas, alinhando-se às diretrizes de segurança do MD.

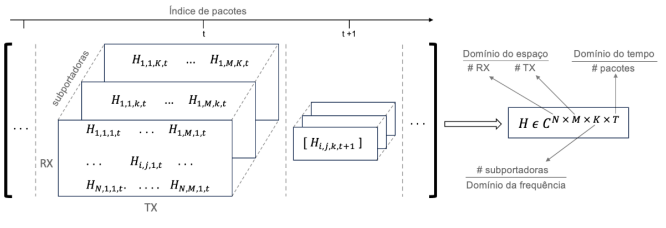

Fig. 1: Matriz CSI adaptada de [19].

Os cálculos sobre os dados brutos do CSI resultam em números complexos que capturam a variação do sinal ao longo do tempo. Esses números são representados como $z = a + bi$, onde a é a parte real, b é a parte imaginária, e i é a unidade imaginária, com a propriedade $i^2 = -1$. A Amplitude (ou módulo) de um número complexo é obtida pela equação $|z| = \sqrt{a^2 + b^2}$, onde a é a parte real e b é a parte imaginária. A Fase (ou ângulo) θ, que é o ângulo do vetor representando o número complexo no plano complexo em relação ao eixo real, é calculada com a fórmula $\theta = \text{atan2}(b, a)$. A função $\text{atan2}(b, a)$ retorna o ângulo cuja tangente é b/a, considerando os sinais de ambos para determinar o quadrante correto. Assim, as variações na amplitude e fase do sinal proporcionam uma compreensão detalhada do comportamento analisado.

[20] relataram que padrões no sinal podem ser identificados e relacionados a características ou comportamentos específicos dos usuários, permitindo, através de aprendizado de máquina, identificar usuários com precisão. Considerando aplicação voltada para inteligência, essas capacidades são essenciais para desenvolver sistemas de autenticação robustos. Este estudo explora métodos de pré-processamento do sinal, extração de características e aplicação de algoritmos de aprendizado de máquina para classificar padrões, alinhando-se às diretrizes de segurança e eficácia operacional estabelecidas pela [1] e pela [4].

B. Autenticação com CSI

De acordo com [13], a utilização do CSI para autenticação humana se baseia na premissa de que diferentes usuários causam variações únicas no CSI ao estarem dentro da cobertura do sinal. Isso permite a identificação de um usuário ao analisar o perfil dinâmico do CSI. Essa abordagem é vital para assegurar a segurança e eficácia dos sistemas de controle de acesso.

Existem dois principais tipos de autenticação de usuários: uma baseada nas flutuações do CSI causadas por movimentos do usuário, como passos, atividades e gestos, e outra que utiliza características de propagação estática do CSI, quando o usuário está parado. A abordagem baseada em ação é amplamente usada devido à capacidade do movimento humano de gerar flutuações evidentes no CSI, que podem ser medidas e processadas por diversos algoritmos. Por outro lado, a autenticação baseada em imobilidade requer a extração de características biofísicas únicas, como silhueta, composição corporal (água, gordura e músculo), ou a combinação de localização do usuário. Tais características são especialmente relevantes em cenários militares, onde a precisão na identificação de indivíduos é essencial para a segurança.

A autenticação do usuário com base no CSI oferece uma maneira promissora de identificar indivíduos com precisão, explorando as nuances das variações do sinal. Além disso, dentro das determinações do MD, essa abordagem pode ser implementada para aumentar a segurança em instalações militares e proteger infraestruturas críticas contra acessos não autorizados. Nesse sentido e para facilitar a identificação de padrões e regularidades nos dados, este estudo segue uma investigação supervisionada, rotulando os dados de entrada e tratando problemas de reconhecimento como classificação.

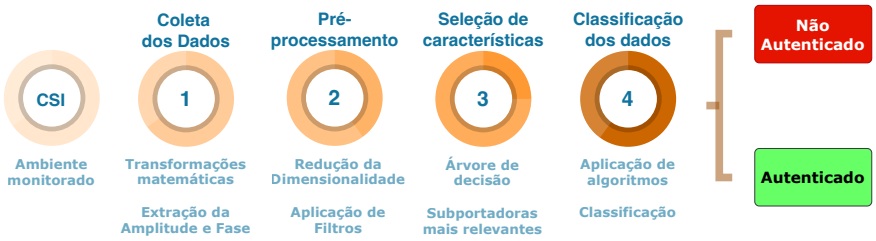

Fig. 2: Modelo proposto.

C. Proposta de Controle de Acesso

Este estudo propõe um sistema de autenticação (Fig. 2) para controle de acesso físico em um ambiente monitorado. Os dados CSI são coletados por dois Raspberry Pi operando em modo monitor. Durante a captura, os dados são transformados do domínio do tempo para o domínio da frequência e reordenados com o *FFT Shift*, permitindo a extração da amplitude e fase do sinal para pré-processamento.

O pré-processamento inclui a normalização dos valores do CFR pela amplitude média dos 242 subcanais monitorados e a aplicação de um algoritmo de sanitização de fase ($\lambda = 10^{-1}$). Os valores do CFR são então combinados em um vetor complexo com 245 componentes. Utilizamos filtros de *Hampel* e *Savitzky-Golay* para clarificar o comportamento do sinal.

Em seguida, aplica-se o algoritmo *Random Forest* para identificar as subportadoras mais relevantes e se desenvolve um modelo de aprendizado de máquina, treinado com validação cruzada de dez etapas, para estimar as características individuais de cada usuário. A autenticação é realizada utilizando o algoritmo *KNN*, reconhecendo usuários autorizados e concedendo acesso conforme as permissões estabelecidas.

A implementação deste sistema baseado em CSI oferece uma camada adicional de segurança para proteger infraestruturas críticas e informações sensíveis. E o uso de dispositivos de baixo consumo energético, como os Raspberry Pi, alinhado com as diretrizes da [1], garante eficiência e sustentabilidade. Além disso, a capacidade de diferenciar homens e mulheres com base em características biofísicas melhora a precisão e a robustez da autenticação, conforme as necessidades de segurança estabelecidas pela [4].

IV. EXPERIMENTOS

Esta seção detalha a realização dos experimentos, a coleta de dados e o protocolo de captura utilizado, permitindo a replicação e futuras melhorias no estudo.

A. Cenário de Captura

Os dados CSI foram coletados em um corredor com acesso a um ambiente controlado, conforme ilustrado na Fig. 3 O corredor, com dimensões de 1,5 m de largura, 5 m de profundidade e 3 m de altura, foi escolhido por sua capacidade de canalizar o acesso e melhorar a reflexão, beneficiando as tecnologias MIMO e OFDM presentes no Wi-Fi. Este cenário está representando o que poderia ser a entrada em infraestruturas críticas.

B. Equipamentos Utilizados

Para a coleta dos dados, foram utilizados dois dispositivos Raspberry Pi 4 modelo B como receptores (Rx), equipados com processadores quad-core Cortex-A72 de 64 bits, 8 GB de RAM LPDDR4, conectividade wireless padrão 802.11b/g/n/ac, Bluetooth 5.0, capacidade PoE, e consumo de energia de 5V/3A, alimentados por *Powerbank* com conexão USB-C. Esses dispositivos foram escolhidos por seu baixo consumo de energia, tamanho compacto (25x52x10mm) e adequação para aplicações de automação em IoT e comunicação M2M. A utilização de dispositivos de baixo consumo energético está em conformidade com as diretrizes de Proteção Cibernética, estabelecidas pela Doutrina Militar de Defesa Cibernética.

Os Raspberry Pi foram controlados remotamente por um notebook DELL Inspiron 15 Gaming 7567, com sistema operacional Windows 10 Home, processador Intel octa-core i7-7700HQ 2.80GHz, 16GB de memória RAM e acesso por SSH. Para simular o sinal Wi-Fi de uma rede fictícia, foi utilizado um roteador TP-Link Archer C60 configurado como transmissor (TX), operando em uma rede de 5GHz a 80MHz, no canal 36. O protocolo de captura seguiu o exibido na Tabela 2, com os objetos de interesse posicionados a 75 cm de distância entre os dois dispositivos.

C. Protocolo de Captura

Foram monitorados dez usuários com características físicas variadas, sendo cinco homens e cinco mulheres. Para garantir precisão e replicabilidade, os usuários variaram entre 1,62 m e 1,85 m de altura, 60 a 93 quilos, e tinham entre 18 e 43 anos. Um dispositivo celular foi utilizado para gerar tráfego,

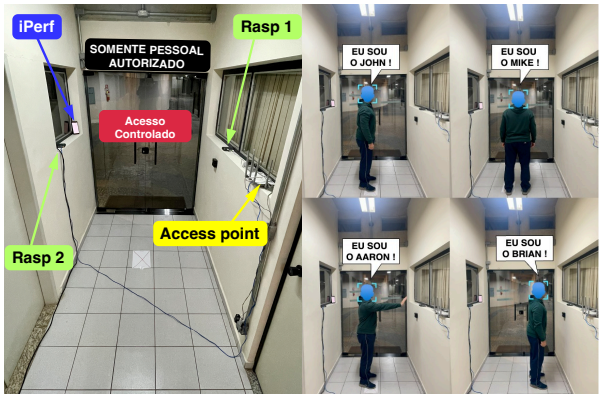

Fig. 3: Captura dos dados CSI no ambiente monitorado.

TABELA 2: PROTOCOLO DE CAPTURA.

Usuário	Altura	Peso	Gênero	Atividade	Total de registros por Dispositivo	Tempo de atividade (s)	Total de capturas
1	1.78 m	79 kg	M	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
2	1.66 m	77 kg	M	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
3	1.85 m	93 kg	M	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
4	1.73 m	90 kg	M	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
5	1.83 m	90 kg	M	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
6	1.65 m	64 kg	F	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
7	1.63 m	65 kg	F	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
8	1.68 m	63 kg	F	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
9	1.65 m	62 kg	F	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50
10	1.67 m	66 kg	F	A, B, C, D e E	Rasp 1: 25, Rasp 2: 25	20	50

empregando a aplicação *iperf2* com uma taxa de aproximadamente 1000 pacotes UDP por segundo, executado em um sistema *Android* com os parâmetros: `-c 192.168.1.1 -u -b 500M -t 60 -i 1 -l 1400`.

Durante as capturas, cinco atividades distintas foram realizadas. Nas atividades *(A), (B), (C) e (D)*, os usuários mantiveram uma distância de 75 cm dos dispositivos, enquanto na atividade *(E)*, estavam a 30 cm do Raspberry 1 e a 1,20 m do Raspberry 2. Na atividade (A), os usuários posicionavam-se de frente para o Raspberry 1 e de costas para o Raspberry 2. Na atividade (B), ficavam com a lateral direita voltada para o Raspberry 1 e a esquerda para o Raspberry 2. A atividade (C) exigia que os usuários ficassem de frente para o Raspberry 1, levantando o braço direito com a mão aberta em direção ao dispositivo. Na atividade (D), ficavam de frente para o Raspberry 1 e levantavam o braço direito com a mão aberta em direção ao peito esquerdo. Por fim, na atividade (E), os usuários posicionavam-se de frente para o Raspberry 1 e pronunciavam a palavra *(ALOHOMORA)*. Ao todo, o conjunto de dados analisado é composto por 500 capturas de 20 segundos, somando aproximadamente 10.000.000 de instâncias que registraram as reflexões, atenuações e difrações do sinal eletromagnético no ambiente analisado.

Cada atividade teve um *delay* inicial de 5 segundos antes do início da gravação, e a ferramenta foi calibrada para que cada sessão de captura tivesse a duração de 20 segundos, garantindo consistência e precisão nos dados coletados. Utilizou-se o modo *Line-of-Sight* (LOS), evitando a presença de obstáculos entre o ponto de transmissão e o ponto de recepção. Dessa forma, os cenários apresentaram menor atenuação e dispersão do sinal, proporcionando uma comunicação mais limpa e robusta.

V. RESULTADOS

Os dados CSI utilizados nos experimentos foram coletados utilizando a ferramenta Nexmon, apresentada inicialmente no estudo [21]. Cada captura continha informações de 256 subportadoras, totalizando aproximadamente 20.000 pacotes por arquivo. Esses dados foram extraídos do *payload* de pacotes UDP presentes nos arquivos PCAP gerados pela ferramenta e posteriormente convertidos em arquivos CSV para facilitar a análise. Os atributos considerados relevantes para a análise focada na autenticação dos usuários foram os dados CSI complexos.

O atributo *CSI Data* forneceu números complexos derivados de cálculos matriciais aplicados aos dados brutos do CSI. Após a transformação matemática, restaram 234 subportadoras úteis para análise, considerando que algumas subportadoras eram nulas ou pilotos.

Para melhorar a qualidade dos dados CSI, foram aplicadas técnicas de pré-processamento como o Filtro de Hampel, Filtro Passa-Baixa, Transformada Discreta de Hilbert, Média Móvel, Filtro de Savitzky-Golay, Filtro de Kalman e Transformada Wavelet Discreta (DWT). Entre essas técnicas, os filtros de *Hampel* e *Savitzky-Golay* mostraram-se mais eficazes na suavização do sinal, preservando suas características.

A. Análise das Subportadoras

A Fig. 4 ilustra as amplitudes médias das subportadoras afetadas por cada usuário durante a atividade ”A”. Observa-se que cada usuário influencia as subportadoras de maneira distinta, permitindo a diferenciação entre os indivíduos e também entre os gêneros, masculino em azul e feminino em vermelho.

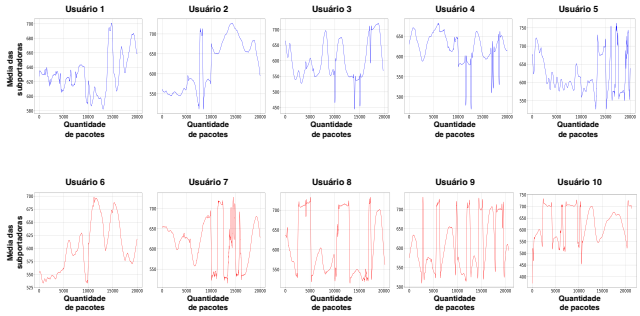

Fig. 4: Influência dos usuários nas subportadoras.

B. Desempenho do Modelo

Para avaliar o desempenho do modelo de autenticação, foram utilizados os algoritmos *KNN*, *RF*, *SVM*, *J48* e *NB*. A Tabela 3 compara o desempenho do modelo com e sem a aplicação do filtro de Hampel. Verifica-se que o KNN teve a melhor acurácio e que o uso de filtros, em geral, melhora a precisão e reduz o tempo de construção do modelo.

C. Função de Distribuição Acumulada

Os altos valores de desempenho dos modelos são justificados com o uso da Função de Distribuição Cumulativa (CDF), conforme ilustrado na Fig. 5. A CDF mostra que as características individuais dos usuários do gênero masculino resultam em padrões de influência no sinal distintos dos apresentados pelos usuários do gênero feminino, permitindo uma identificação precisa.

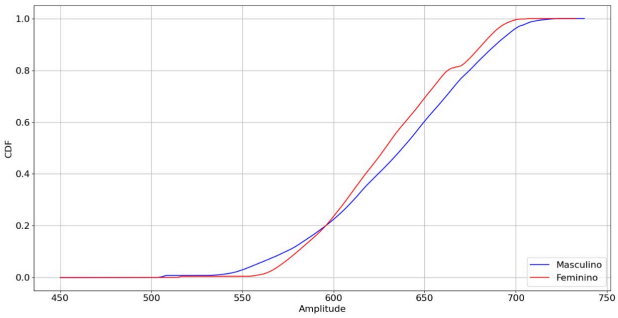

Fig. 5: CDF dos usuários masculinos e femininos.

TABELA 3: RESULTADO DOS CLASSIFICADORES

Métricas	Com Filtro					Sem Filtro				
	KNN	**RF**	**SVM**	**J48**	**NB**	**KNN**	**RF**	**SVM**	**J48**	**NB**
Acurácia Média (%)	99.99	99.99	99.80	99.94	94.05	99.99	99.99	96.93	99.90	94.06
F1-Score (%)	100.00	100.00	100.00	99.90	94.10	100.00	100.00	97.00	99.00	94.00
Tempo de Construção (s)	0.10	123.48	6.91	244.03	10.79	0.10	97.91	189.25	329.09	11.07
Kappa	0.999	1.000	0.996	0.998	0.881	0.999	0.999	0.938	0.998	0.881
MCC	1.00	1.00	0.996	0.999	0.881	1.00	1.000	0.944	0.998	0.884

VI. CONCLUSÃO E TRABALHOS FUTUROS

Este estudo introduz uma nova abordagem para autenticação de usuários em sistemas de controle de acesso físico, utilizando dados de Wi-Fi. Empregando dispositivos Raspberry Pi num ambiente controlado e aplicando técnicas de aprendizado supervisionado, alcançamos uma acurácia de 99,99% na distinção entre homens e mulheres, utilizando o classificador *K-Nearest Neighbors (KNN)*. A pesquisa demonstrou que características biofísicas e comportamentais capturadas pelo CSI são eficazes para autenticação de usuários, oferecendo uma alternativa viável aos métodos tradicionais de controle de acesso.

No âmbito da Defesa, este trabalho está alinhado com as diretrizes de proteção e exploração cibernética, contribuindo para a segurança das infraestruturas críticas e a eficácia das operações de inteligência. A capacidade de distinguir usuários com base em características biofísicas é importante para a exploração cibernética, permitindo mapear e identificar com precisão os indivíduos em espaços cibernéticos de interesse, podendo ser útil tanto em operações defensivas quanto ofensivas. Apesar dos resultados promissores, a aplicação prática da tecnologia proposta enfrenta desafios, como a necessidade de ajustes em ambientes com alta interferência eletromagnética.

Como trabalhos futuros, visualiza-se aumentar a base de dados, integrar mais dispositivos Raspberry Pi para capturar uma gama maior de características biofísicas, avaliar o modelo em diferentes cenários e explorar a viabilidade de utilizar aprendizado não supervisionado. Essas implementações podem contribuir ainda mais com a produção de conhecimento de inteligência para apoiar a tomada de decisão em níveis estratégicos, operacionais e táticos.

AGRADECIMENTOS

Este trabalho tem apoio financeiro do Programa de Pós-graduação em Aplicações Operacionais—PPGAO/ITA, da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) processo #2020/09850-0 e #2022/00741-0, do CNPq e da CAPES.

REFERÊNCIAS

[1] M. da Defesa, *Doutrina Militar de Defesa Cibernética*, Estado-Maior Conjunto das Forças Armadas, 2023, 2ª Edição. [Online]. Available: https://mdlegis.defesa.gov.br/pesquisar_normas/

[2] D. D. C. Brito, R. P. Ferreira, and A. J. Chaves, "Perspectivas globais militares sobre tecnologias quânticas," *Aplicações Operacionais em Áreas de Defesa*, vol. 24, no. 1, pp. 54–60, Sep. 2023. [Online]. Available: https://spectrum.ita.br/index.php/spectrum/article/view/389

[3] M. da Defesa, *Estratégia de Inteligência de Defesa*, Estado-Maior Conjunto das Forças Armadas, 2023, 1ª Edição. [Online]. Available: https://mdlegis.defesa.gov.br/pesquisar_normas/

[4] ——, *Política de Inteligência de Defesa*, Estado-Maior Conjunto das Forças Armadas, 2023, 1ª Edição. [Online]. Available: https://mdlegis.defesa.gov.br/pesquisar_normas/

[5] S. W. Shah and S. S. Kanhere, "Wi-auth: Wifi based second factor user authentication," 2017, pp. 393–402. [Online]. Available: https://doi.org/10.1145/3144457.3144468

[6] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," ser. UbiComp '16, 2016, pp. 363–373. [Online]. Available: https://doi.org/10.1145/2971648.2971670

[7] L. Guo, L. Wang, J. Liu, W. Zhou, B. L. T. Liu, G. Li, and C. Li, "A novel benchmark on human activity recognition using wifi signals," in *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 2017, pp. 1–6.

[8] D. Zhang, H. Wang, and D. Wu, "Toward centimeter-scale human activity sensing with wi-fi signals," *Computer Society*, vol. 50, no. 1, pp. 48–57, 2017.

[9] K. Niu, F. Zhang, Z. Chang, and D. Zhang, "A fresnel diffraction model based human respiration detection system using cots wi-fi devices," ser. UbiComp '18, 2018. [Online]. Available: https://doi.org/10.1145/3267305.3267561

[10] F. Zhang, K. Niu, J. Xiong, B. Jin, T. Gu, Y. Jiang, and D. Zhang, "Towards a diffraction-based sensing approach on human activity recognition," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 1, mar 2019.

[11] H. Kong, L. Lu, J. Yu, Y. Chen, X. Xu, F. Tang, and Y.-C. Chen, "Multiauth: Enable multi-user authentication with single commodity wifi device," ser. MobiHoc '21, 2021, pp. 31–40.

[12] A. Afshar, V. T. Vakili, and S. Daei, "Active user detection and channel estimation for spatial-based random access in crowded massive mimo systems via blind super-resolution," *IEEE Signal Processing Letters*, vol. 29, pp. 1072–1076, 2022.

[13] F. Meneghello, D. Garlisi, N. D. Fabbro, I. Tinnirello, and M. Rossi, "Sharp: Environment and person independent activity recognition with commodity ieee 802.11 access points," *IEEE Transactions on Mobile Computing*, vol. 22, no. 10, pp. 6160–6175, 2023.

[14] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A survey on behavior recognition using wifi channel state information," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 98–104, 2017.

[15] C. Lin, J. Hu, Y. Sun, F. Ma, L. Wang, and G. Wu, "Wiau: An accurate device-free authentication system with resnet," in *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2018, pp. 1–9.

[16] H. Zou, Y. Zhou, J. Yang, H. Jiang, L. Xie, and C. J. Spanos, "Deepsense: Device-free human activity recognition via autoencoder long-term recurrent convolutional network," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–6.

[17] Y. Gu, H. Yan, M. Dong, M. Wang, X. Zhang, Z. Liu, and F. Ren, "Wione: One-shot learning for environment-robust device-free user authentication via commodity wi-fi in man–machine system," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 3, pp. 630–642, 2021.

[18] Y. Gu, Y. Wang, M. Wang, Z. Pan, Z. Hu, Z. Liu, F. Shi, and M. Dong, "Secure user authentication leveraging keystroke dynamics via wi-fi sensing," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2784–2795, 2022.

[19] Y. Ma, G. Zhou, and S. Wang, "Wifi sensing with channel state information: A survey," *ACM Comput. Surv.*, vol. 52, no. 3, p. 46, jun 2019. [Online]. Available: https://doi.org/10.1145/3310194

[20] D. Wu, D. Zhang, C. Xu, H. Wang, and X. Li, "Device-free wifi human sensing: From pattern-based to model-based approaches," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 91–97, 2017.

[21] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your csi: A channel state information extraction platform for modern wi-fi chipsets," ser. WiNTECH '19, 2019, pp. 21–28.