UFU-Do53-EXF e UFU-DoH-EXF: conjuntos de dados para detecção de exfiltração via DNS

Cristiano L. M. Borges[1] **, Rodrigo S. Miani**[1]

[1]Faculdade de Computação (FACOM) – Universidade Federal de Uberlândia (UFU)

{cristiano.borges,miani}@ufu.br

Abstract. *Data exfiltration through the Domain Name System (DNS) remains a relevant cybersecurity challenge, as DNS is an essential network service that can be misused as a covert channel for unauthorized data transfer. This paper introduces two complementary datasets, UFU-Do53-EXF and UFU-DoH-EXF, designed to support the study of data exfiltration in conventional DNS over port 53 and DNS over HTTPS. The proposed methodology combines controlled generation of benign and malicious traffic, PCAP packet captures, a reproducible PCAP-to-Flow processing pipeline, and protocol-specific feature extraction. After normalization, UFU-Do53-EXF comprises 5,351,064 samples, whereas UFU-DoH-EXF comprises 1,048,575 TCP/443 flows. Together, the datasets support the analysis of exfiltration behavior in both clear-text and encrypted DNS traffic. Supervised baseline evaluations using widely adopted machine learning classifiers indicate that the proposed representations are consistent and capture discriminative patterns for distinguishing benign traffic from exfiltration activity.*

Resumo. *A exfiltração de dados via Domain Name System (DNS) permanece como um problema relevante em segurança cibernética, pois o DNS é um serviço essencial ao funcionamento das redes e pode ser abusado como canal para transferência não autorizada de dados. Este artigo apresenta dois conjuntos de dados complementares, UFU-Do53-EXF e UFU-DoH-EXF, concebidos para apoiar o estudo da exfiltração em DNS convencional sobre a porta 53 e DNS sobre HTTPS. A proposta combina geração controlada de tráfego benigno e malicioso, capturas de pacotes em formato PCAP, um pipeline reprodutível de processamento PCAP-to-Flow e extração de atributos específicos por protocolo. Após normalização, o conjunto UFU-Do53-EXF contém 5.351.064 amostras, enquanto o UFU-DoH-EXF contém 1.048.575 fluxos TCP/443. Em conjunto, os dados apoiam a análise do comportamento de exfiltração tanto em tráfego DNS em texto claro quanto em tráfego criptografado. Avaliações supervisionadas de referência, com classificadores amplamente utilizados na literatura, indicam que as representações propostas são consistentes e capturam padrões discriminativos para distinguir tráfego benigno de atividades de exfiltração.*

1. Introdução

A exfiltração de dados corresponde à transferência não autorizada de informações de um ambiente comprometido para uma infraestrutura controlada pelo atacante. Em bases de conhecimento amplamente adotadas na área, essa etapa é reconhecida como um mecanismo recorrente de efetivação da exfiltração em campanhas de intrusão, inclusive por

meio de canais de comando e controle ou do emprego de protocolos alternativos de comunicação [MITRE ATT&CK 2026a, MITRE ATT&CK 2026b, MITRE ATT&CK 2026c]. Entre os protocolos suscetíveis a esse abuso, o DNS ocupa posição de destaque por ser indispensável ao funcionamento dos serviços de rede e, em muitos ambientes, menos restritivo do que outros canais de saída [Steadman and Scott-Hayward 2018].

No protocolo DNS tradicional, a exfiltração pode ocorrer mediante codificação de dados em subdomínios, tunelamento ou uso do protocolo como canal furtivo para transmissão de fragmentos de informação [Steadman and Scott-Hayward 2018, Ahmed et al. 2019, Abualghanam et al. 2023]. No *DNS over HTTPS* (DoH), a dificuldade de detecção aumenta porque as mensagens DNS passam a ser encapsuladas em HTTPS/TLS, usualmente sobre HTTP/2, reduzindo a inspeção direta do conteúdo e deslocando a análise para estatísticas de fluxo, direção dos pacotes, tamanhos observáveis e comportamento temporal [Hoffman and McManus 2018, Rescorla 2018, Zhan and others 2022].

A literatura recente tem mostrado que a detecção de exfiltração via DNS por aprendizado de máquina é viável [Elaoumari 2025], mas permanece condicionada pela qualidade dos conjuntos de dados empregados, pela rastreabilidade do processo de rotulagem e pela aderência entre a unidade amostral e o fenômeno observado [Sabir et al. 2021, Garcia and Valeros 2023]. Nesse contexto, este trabalho apresenta dois conjuntos de dados: *UFU-Do53-EXF*, voltado ao DNS convencional, e *UFU-DoH-EXF*, voltado ao DNS sobre HTTPS. Ambos foram gerados sob o mesmo desenho experimental, com geração controlada de tráfego benigno e malicioso, preservação de artefatos em PCAP e fluxo de rede, definição explícita dos atributos e documentação metodológica.

As principais contribuições deste trabalho são:

- a proposição de dois conjuntos de dados orientados especificamente à detecção de exfiltração via DNS, contemplando os cenários Do53 e DoH;
- a definição de um arcabouço experimental comum, com geração controlada de tráfego benigno e malicioso, preservação de PCAPs, conversão reprodutível para fluxo de rede e rotulagem explícita;
- a caracterização de atributos específicos para DNS em texto claro e para DNS sobre HTTPS, respeitando as diferenças de observabilidade entre os protocolos;
- a apresentação de avaliações supervisionadas de referência para os conjuntos *UFU-Do53-EXF* e *UFU-DoH-EXF*, utilizando classificadores amplamente empregados em problemas de detecção.

2. Exfiltração de dados por DNS

A exfiltração via DNS explora a presença constante desse protocolo nas comunicações de rede e sua importância operacional para a resolução de nomes. Por participar de grande parte das interações entre clientes, aplicações e serviços, o DNS oferece ao atacante um meio para transportar dados sem recorrer, necessariamente, a canais mais visados pelos mecanismos tradicionais de inspeção. Em consequência, o protocolo passou a ocupar posição relevante na superfície de ataque, tanto em operações de comando e controle quanto em cenários de tunelamento e exfiltração de informações [MITRE ATT&CK 2026a, MITRE ATT&CK 2026b, MITRE ATT&CK 2026c].

Em sua forma mais comum, a exfiltração por DNS consiste em fragmentar o conteúdo de interesse, aplicar uma codificação apropriada e inserir esses fragmentos em consultas dirigidas a um domínio sob controle do adversário. A infraestrutura autoritativa associada a esse domínio pode então reconstituir os dados recebidos. Esse mecanismo se beneficia da flexibilidade do DNS, especialmente no uso de rótulos de subdomínio, e pode ser adaptado para reduzir suspeitas mediante ajustes de frequência, tamanho e distribuição das consultas [Steadman and Scott-Hayward 2018, Güneş Gürsoy et al. 2024].

No contexto de Do53, a observação do conteúdo da mensagem DNS permite explorar atributos associados à estrutura do nome consultado e ao comportamento de resolução. Comprimento da consulta, profundidade do nome de domínio, entropia dos rótulos, proporção de dígitos, tipos de registros e estatísticas agregadas por domínio constituem sinais úteis para diferenciar tráfego benigno de padrões artificiais de tunelamento e exfiltração [Ahmed et al. 2020, Tang et al. 2020, Palau et al. 2019, Nadler et al. 2019]. Em tráfego benigno, é mais comum observar nomes regulares, recorrência de domínios associados a aplicações conhecidas e combinações coerentes de tipos de consulta. Em tráfego malicioso, por outro lado, aumentam a ocorrência de cadeias codificadas, a irregularidade lexical e a presença de padrões produzidos para transporte de dados [Abualghanam et al. 2023, Ahmed et al. 2019].

Sob a perspectiva defensiva, abordagens baseadas exclusivamente em *blocklists* apresentam limitações importantes. Listas estáticas de bloqueio podem ser úteis como mecanismo complementar, mas tendem a responder de forma insuficiente a domínios novos, infraestrutura rotativa e consultas maliciosas que procuram se aproximar do comportamento legítimo. Por esse motivo, exige inspeção aprofundada de consultas, análise contextual e detecção de padrões anômalos relacionados a comprimento, frequência, volume, entropia e estrutura lexical dos nomes consultados [Ahmed et al. 2019, Ahmed et al. 2020, Steadman and Scott-Hayward 2018]. Essa perspectiva é consistente com a compreensão de que ataques bem-sucedidos evoluem mais rapidamente do que mecanismos de bloqueio puramente reativos conseguem acompanhar.

Em DoH, a natureza do problema muda de forma substancial. Como o conteúdo DNS passa a trafegar encapsulado em HTTPS/TLS, a inspeção direta das consultas deixa de estar disponível em texto claro para os mecanismos tradicionais de monitoramento. Com isso, a detecção precisa se apoiar em efeitos laterais do comportamento de comunicação, como duração do fluxo, contagens de pacotes, estatísticas de tamanho em cada direção, intervalos entre pacotes e informações temporais associadas ao estabelecimento e à transmissão da conexão TLS [Hoffman and McManus 2018, Rescorla 2018, Zhan and others 2022, Moure-Garrido et al. 2023, Salat et al. 2023]. Em termos práticos, o avanço do DNS criptografado amplia a dificuldade de monitoramento e impõe maior dependência de estatísticas de fluxo, padrões temporais e controles complementares para fins de detecção e aplicação de políticas de segurança [Zhan and others 2022, Moure-Garrido et al. 2023].

Essa distinção entre Do53 e DoH é metodologicamente relevante porque afeta diretamente a formulação do problema de detecção. No primeiro caso, a modelagem pode explorar atributos derivados da própria consulta DNS. No segundo, a análise precisa ser deslocada para sinais indiretos associados ao transporte criptografado. Em ambos os cenários, porém, a exfiltração via DNS permanece como uma forma tecni-

camente viável de abuso do protocolo, o que justifica a construção de conjuntos de dados específicos e metodologicamente consistentes para apoiar estudos de detecção [Sabir et al. 2021, Steadman and Scott-Hayward 2018, Zhan and others 2022].

3. Trabalhos Relacionados

3.1. Métodos de detecção de exfiltração via DNS

A detecção de exfiltração via DNS depende da visibilidade sobre o protocolo e dos atributos analisados. No DNS convencional sobre a porta 53 (Do53), em que o conteúdo das consultas permanece observável, são comuns abordagens baseadas em atributos léxicos dos domínios, metadados do protocolo e estatísticas comportamentais agregadas. Trabalhos anteriores investigam desde a detecção em tempo real em redes corporativas até o uso de aprendizado supervisionado e profundo para identificar comunicações encobertas e ameaças associadas ao abuso do DNS [Ahmed et al. 2019, Tang et al. 2020, Palau et al. 2019]. Outros estudos destacam desafios como exfiltração de baixa vazão, tunelamento furtivo e limitações de métricas volumétricas, reforçando a importância de atributos estruturais, entropia e informações contextuais dos domínios consultados [Nadler et al. 2019, Steadman and Scott-Hayward 2018, Abualghanam et al. 2023].

De modo geral, os métodos propostos para Do53 convergem para a ideia de que o abuso do DNS tende a produzir sinais discriminativos observáveis no próprio conteúdo da consulta ou em padrões de resolução associados. Comprimento do nome, profundidade hierárquica, distribuição de caracteres, presença de cadeias codificadas, proporção de dígitos e tipos de registros estão entre os sinais mais explorados para diferenciar tráfego benigno de comportamentos artificiais de tunelamento e exfiltração [Ahmed et al. 2020, Tang et al. 2020, Nadler et al. 2019, Palau et al. 2019]. Essa linha mostra que a detecção supervisionada pode alcançar bons resultados quando a representação dos dados preserva adequadamente as propriedades relevantes da comunicação DNS.

No contexto de DoH, o problema assume características distintas, pois as mensagens DNS passam a ser encapsuladas em HTTPS/TLS, o que reduz a possibilidade de inspeção semântica direta das consultas. Nesse cenário, a literatura desloca a modelagem para atributos indiretos, como duração de fluxo, contagem de pacotes, estatísticas de tamanho por direção, intervalos temporais e informações associadas ao estabelecimento e à transmissão da conexão TLS. Zhan et al. demonstram que a detecção de exfiltração em DoH permanece viável mesmo sem acesso ao conteúdo DNS em texto claro, desde que sejam utilizados atributos estatísticos adequados [Zhan and others 2022]. Em sentido semelhante, Moure-Garrido et al. e Salat et al. mostram que o tráfego DoH malicioso tende a introduzir alterações mensuráveis em padrões temporais e direcionais, ainda que a criptografia imponha limitações adicionais à interpretabilidade dos resultados [Moure-Garrido et al. 2023, Salat et al. 2023].

Em termos mais amplos, revisões e estudos comparativos indicam que o desempenho dos métodos de detecção não depende apenas do algoritmo empregado, mas também da qualidade da rotulagem, da aderência entre a unidade amostral e o fenômeno observado e da consistência metodológica da coleta [Sabir et al. 2021, Garcia and Valeros 2023, Luz et al. 2023, Ozery et al. 2024]. Assim, embora haja evidências consistentes de que a detecção de exfiltração via DNS seja factível tanto em Do53 quanto em DoH, a utilidade prática desses métodos permanece fortemente condicionada à disponibilidade de

conjuntos de dados bem documentados, representativos e compatíveis com o protocolo analisado [Sabir et al. 2021, Garcia and Valeros 2023].

3.2. Conjuntos de dados para detecção de exfiltração via DNS

Além da evolução dos métodos de detecção, a literatura também evidencia a importância dos conjuntos de dados utilizados na avaliação experimental. No caso de exfiltração via DNS, essa questão é particularmente relevante porque diferentes bases públicas adotam objetivos distintos, unidades amostrais não equivalentes e níveis variados de documentação metodológica. Como consequência, comparações entre estudos nem sempre refletem apenas diferenças entre modelos, mas também diferenças estruturais entre os próprios dados empregados [Garcia and Valeros 2023, Sabir et al. 2021].

Entre os conjuntos frequentemente citados na literatura, o *DNS Exfiltration Dataset* reúne grande volume de amostras voltadas ao cenário de exfiltração em DNS tradicional, mas não disponibiliza, de forma estruturada, todos os artefatos necessários para rastrear integralmente o processo de geração e transformação dos dados. O *CIC-Bell-DNS-EXF-2021* representa uma referência importante para Do53 ao combinar artefatos de captura e representação em fluxo de rede, embora não tenha sido concebido para estabelecer um paralelo metodológico direto com o cenário de DNS criptografado. Já o *CIRA-CIC-DoHBrw-2020* se destaca como base relevante para estudos em DoH, disponibilizando capturas, fluxos e scripts, mas com foco distinto do problema específico de exfiltração sob o mesmo desenho experimental adotado para DNS convencional.

A Tabela 1 sintetiza essa comparação entre conjuntos de dados relacionados e os conjuntos propostos neste trabalho. O objetivo não é apenas contrastar volume amostral, mas destacar critérios que afetam diretamente a utilidade científica da base, como cobertura protocolar, presença de artefatos em PCAP, representação em fluxo de rede e disponibilidade de scripts que favoreçam reprodutibilidade.

Tabela 1. Comparação entre conjuntos de dados relacionados e os conjuntos propostos.

Conjunto de dados	Do53	DoH	PCAP	Fluxo	Scripts	Comparabilidade	Amostras
DNS Exfiltration Dataset	✓	✗	✗	✓	✗	✗	1.583.163
CIC-Bell-DNS-EXF-2021	✓	✗	✓	✓	✗	✗	1.019.318
CIRA-CIC-DoHBrw-2020	✗	✓	✓	✓	✓	✗	269.643
UFU-Do53-EXF	✓	✗	✓	✓	✓	✓	5.351.064
UFU-DoH-EXF	✗	✓	✓	✓	✓	✓	1.048.575

Observa-se que as bases existentes oferecem contribuições importantes, mas também apresentam limitações recorrentes quando consideradas sob a perspectiva da comparação metodológica entre Do53 e DoH. Em muitos casos, os conjuntos cobrem apenas um dos cenários, utilizam unidades amostrais diferentes, não preservam simultaneamente capturas e representações estruturadas, ou não disponibilizam documentação suficiente para reconstituir o processo experimental com precisão. Essas limitações afetam não apenas a reprodutibilidade, mas também a comparabilidade entre estudos que buscam compreender o impacto da mudança de observabilidade entre DNS tradicional e DNS

criptografado. Nesse contexto, a motivação para a criação dos conjuntos *UFU-Do53-EXF* e *UFU-DoH-EXF* decorre da necessidade de reduzir essa lacuna metodológica. Ao estabelecer um desenho experimental comum, com geração controlada de tráfego benigno e malicioso, preservação de artefatos em PCAP e fluxo de rede, rotulagem coerente e atributos aderentes às características de cada protocolo, busca-se fornecer uma base mais consistente para o estudo comparativo da exfiltração via DNS em cenários Do53 e DoH [Garcia and Valeros 2023, Sabir et al. 2021, Srivastava et al. 2022].

4. Descrição dos conjuntos de dados propostos

Esta seção descreve o processo de construção dos conjuntos *UFU-Do53-EXF* e *UFU-DoH-EXF*. A metodologia foi concebida para manter coerência entre os dois protocolos analisados, preservando comparabilidade entre cenários benignos e maliciosos, rastreabilidade da geração do tráfego e consistência na transformação das capturas em representações adequadas para modelagem. Em ambos os casos, a construção dos conjuntos partiu de um mesmo desenho experimental, com definição prévia dos cenários de tráfego, topologia laboratorial reprodutível, captura em formato PCAP e posterior conversão para representações estruturadas no nível de fluxo. Com objetivo de viabilizar a reprodutibilidade dos experimentos, todos os artefatos, scripts e procedimentos metodológicos empregados neste estudo foram disponibilizados publicamente no repositório GitHub do projeto: `https://github.com/CborgesUFU/DNS-360-Dataset`.

4.1. Visão geral dos conjuntos UFU-Do53-EXF e UFU-DoH-EXF

Os conjuntos propostos foram organizados de forma complementar. O *UFU-Do53-EXF* foi concebido para representar o problema da exfiltração em DNS convencional, preservando atributos diretamente observáveis nas consultas e respostas trocadas na porta 53. O *UFU-DoH-EXF*, por sua vez, foi construído para representar o mesmo problema em um contexto de menor observabilidade semântica, no qual as mensagens DNS são encapsuladas em HTTPS/TLS e a análise depende predominantemente de estatísticas de fluxo e informações temporais do tráfego TLS.

Cada conjunto incorpora dois cenários principais. O primeiro corresponde ao tráfego benigno, gerado de forma controlada para representar resoluções legítimas compatíveis com o protocolo analisado. O segundo corresponde ao tráfego malicioso com exfiltração, produzido por ferramentas e rotinas específicas para transporte encoberto de dados. Essa organização em cenários benigno e malicioso foi mantida tanto em Do53 quanto em DoH, permitindo comparar protocolos distintos sob um mesmo arcabouço metodológico. No caso do *UFU-Do53-EXF*, o cenário benigno é composto por consultas DNS tradicionais direcionadas a resolvedores legítimos, com diversidade de domínios e tipos de registros. O cenário malicioso, por sua vez, envolve a exfiltração de dados por meio de consultas DNS manipuladas, nas quais fragmentos de informação são encapsulados em nomes de domínio e encaminhados a uma infraestrutura controlada pelo atacante. No caso do *UFU-DoH-EXF*, o cenário benigno é formado por consultas DoH realizadas por meio de requisições HTTPS/TLS para resolvedores públicos, enquanto o cenário malicioso consiste em um cliente customizado capaz de exfiltrar dados por meio de requisições compatíveis com o paradigma de *DNS over HTTPS*.

4.2. Parâmetros experimentais

A construção dos cenários benignos partiu de uma seleção controlada de domínios, com o objetivo de reduzir vieses de escolha e favorecer reprodutibilidade. Para isso, as consultas benignas foram baseadas em domínios extraídos da *Tranco Top 1 Million List*, concebida especificamente para uso em pesquisa e enriquecida contra manipulação na composição do ranking [Le Pochat et al. 2019]. Não existe garantia de que os domínios incluídos sejam benignos, pois sua seleção baseia-se em popularidade e não em critérios de segurança ou reputação. Todavia, essa escolha contribui para evitar listas arbitrárias ou pouco documentadas e melhora a qualidade do *baseline* benigno, tanto em Do53 quanto em DoH.

Além da seleção dos domínios, a geração dos cenários de exfiltração exigiu a definição de um conjunto heterogêneo de arquivos capaz de representar artefatos comuns em estações de trabalho. Para esse fim, foi utilizado um conjunto de arquivos de áudio, vídeo, imagem, texto, executáveis binários e arquivos compactados. Esses arquivos foram organizados em dois perfis, denominados *light* e *heavy*, de modo a induzir variações controladas no volume de dados exfiltrados. Para fins de organização experimental, consideraram-se como *light* os arquivos com tamanho de até aproximadamente 1 MB, e como *heavy* aqueles com tamanho superior a esse valor. Essa estratégia permitiu produzir tráfego malicioso com diferentes intensidades, sem comprometer a consistência entre os experimentos conduzidos em Do53 e DoH.

A adoção dos mesmos insumos lógicos para os dois protocolos foi importante para reduzir interferências externas à comparação. Em vez de alterar simultaneamente tipo de arquivo, protocolo e topologia, optou-se por manter estáveis os artefatos utilizados na exfiltração e variar apenas os mecanismos de transporte e observação correspondentes a cada protocolo. Com isso, tornou-se mais viável comparar os dois contextos sem introduzir diferenças metodológicas desnecessárias.

4.3. Topologia do laboratório

A geração das capturas foi realizada em uma topologia laboratorial implementada em ambiente de nuvem, na Amazon Web Services (AWS). O desenho do laboratório foi orientado por três objetivos principais: isolar os papéis de atacante e vítima em segmentos distintos, garantir reprodutibilidade de endereçamento e regras de conectividade, e permitir coleta em pontos complementares da comunicação. A infraestrutura foi organizada como uma *Virtual Private Cloud* (VPC) com segmentação lógica em duas redes privadas. Uma rede corresponde ao ambiente do atacante e a outra ao ambiente da vítima. Essa separação foi adotada para preservar clareza na atribuição dos papéis experimentais e facilitar a repetição dos cenários em condições controladas.

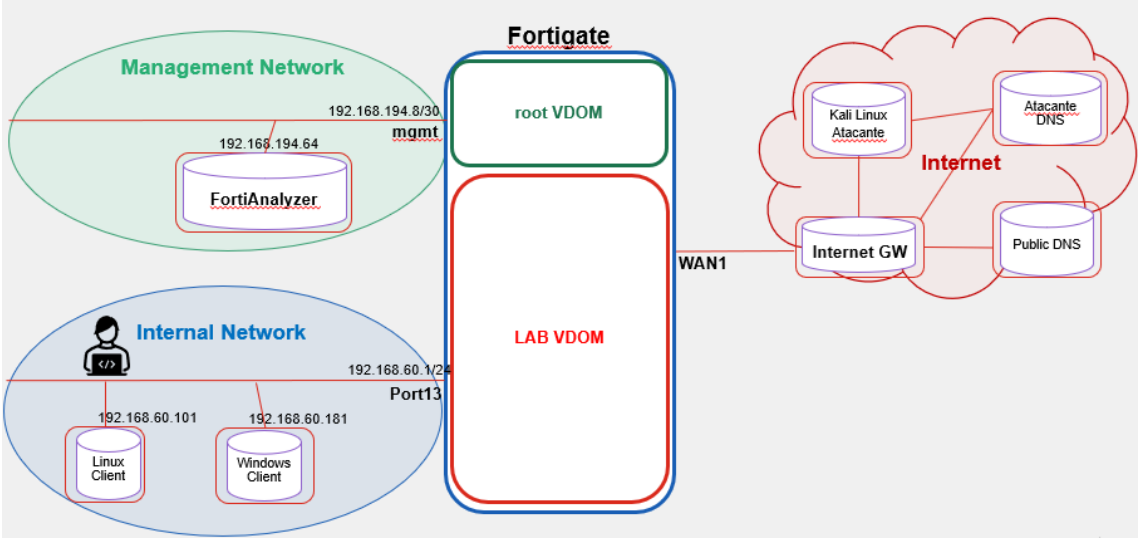

Figura 1. Topologia do laboratório utilizada na geração dos cenários Do53 e DoH.

Como ilustrado na Figura 1, entre os componentes do laboratório, o FortiGate com *Evaluation License* foi utilizado como elemento de roteamento e coleta de tráfego, permitindo apoio à observação das comunicações e, quando necessário, captura por mecanismos complementares. O FortiAnalyzer foi empregado para consolidação de logs e apoio à correlação dos eventos. O ambiente também contou com um servidor Windows para funções auxiliares, uma máquina do atacante baseada em Kali Linux e uma máquina da vítima baseada em Ubuntu. A máquina do atacante executou as ferramentas de exfiltração e de comando e controle, incluindo o `dnscat2` para o cenário Do53 e o cliente ou rotinas do cenário DoH. Essa topologia, embora deliberadamente reduzida em relação a ambientes corporativos reais, foi suficiente para reproduzir fluxos benignos e maliciosos em Do53 e DoH, registrar as comunicações em formato PCAP e sustentar a etapa posterior de transformação para representações no nível de fluxo.

4.4. Cenários de tráfego e geração das capturas

A geração das capturas foi organizada em quatro cenários experimentais. Dois deles correspondem ao protocolo DNS tradicional e dois ao *DNS over HTTPS*, sempre contemplando um cenário benigno e um cenário de exfiltração. Essa organização permite avaliar o comportamento do tráfego em dois contextos de observabilidade distintos, mas preservando uma lógica experimental comum.

No cenário benigno Do53, a máquina da vítima executa consultas DNS convencionais via 53/UDP para resolvedores legítimos. O objetivo desse cenário é estabelecer uma linha de base controlada para o comportamento de resolução de nomes. Para evitar um benigno excessivamente restrito, tanto o volume quanto a temporalidade das consultas foram controlados por scripts específicos, e a diversidade dos destinos consultados foi orientada pela lista Tranco [Le Pochat et al. 2019]. Além disso, diferentes tipos de registros podem ser utilizados ao longo da geração, ampliando a variabilidade dos campos observáveis no protocolo. No cenário malicioso Do53, a máquina da vítima estabelece um canal de comando e controle sobre DNS por meio da ferramenta `dnscat2`. Nessa configuração, dados são fragmentados e encapsulados em nomes de domínio, que passam a ser transmi-

tidos em consultas DNS dirigidas a uma infraestrutura controlada pelo atacante. A escolha dessa ferramenta foi adequada ao objetivo metodológico do trabalho, pois permite repetibilidade, controle do experimento e geração de tráfego coerente com a literatura sobre tunelamento e exfiltração via DNS [Bowes 2022, Steadman and Scott-Hayward 2018].

No cenário benigno DoH, a vítima realiza consultas a resolvedores públicos de *DNS over HTTPS* por meio de requisições HTTPS/TLS na porta 443. Nesse caso, a observação do tráfego já não se apoia no conteúdo DNS em texto claro, mas em efeitos laterais do transporte criptografado, como duração dos fluxos, contagem de pacotes, tamanhos em cada direção, estatísticas de fluxo e informações temporais associadas ao estabelecimento e à transmissão da conexão TLS. Para evitar um *baseline* enviesado, buscou-se diversidade de domínios e resolvedores, preservando o caráter controlado da coleta. No cenário malicioso DoH, um cliente customizado realiza a exfiltração dos arquivos por meio de requisições HTTPS compatíveis com o paradigma de *DNS over HTTPS*. O conteúdo é fragmentado e distribuído em múltiplas requisições, de forma que o efeito da exfiltração passe a se manifestar principalmente nas estatísticas bidirecionais de fluxo e nas informações temporais associadas ao tráfego TLS, e não no conteúdo das mensagens DNS propriamente dito. Esse cenário é particularmente relevante porque representa uma condição de menor observabilidade semântica, mais próxima dos desafios impostos por DNS criptografado em ambientes operacionais.

4.5. Pipeline de conversão de PCAP para Flow

A conversão das capturas de rede para os conjuntos de dados finais foi conduzida por meio de um pipeline reprodutível de processamento *PCAP-to-Flow*, implementado em Python. O objetivo dessa etapa foi transformar arquivos PCAP/PCAPNG em fluxos para viabilidade dos experimentos de aprendizado de máquina. Embora os conjuntos *UFU-Do53-EXF* e *UFU-DoH-EXF* tenham sido construídos a partir de rotinas específicas para cada protocolo, ambos seguiram a mesma lógica metodológica. A implementação utilizou a biblioteca `dpkt` para leitura e interpretação dos arquivos PCAP/PCAPNGs. As bibliotecas `NumPy` e `pandas` foram empregadas, respectivamente, para o cálculo das estatísticas numéricas e para a organização tabular das amostras extraídas. Essa escolha permitiu manter o pipeline independente de ferramentas proprietárias e favoreceu a reprodutibilidade do processo de transformação das capturas em dados estruturados.

No cenário Do53, o pipeline processa pacotes DNS convencionais associados à porta 53, contemplando tráfego sobre UDP e TCP quando aplicável. Para cada mensagem DNS válida, são extraídos atributos diretamente relacionados ao nome consultado, aos metadados da consulta e às respostas observadas. No cenário DoH, o pipeline processa comunicações TCP associadas à porta 443. Como o conteúdo DNS está encapsulado em HTTPS/TLS, a extração não depende da inspeção direta das consultas DNS. Em vez disso, os pacotes são organizados em fluxos bidirecionais, considerando os sentidos cliente→servidor e servidor→cliente. Para cada unidade amostral, são calculadas estatísticas de tamanho dos pacotes, intervalos temporais, quantidade de pacotes por direção, duração da comunicação e volume total de bytes.

A rotulagem das amostras foi realizada a partir do cenário experimental de origem de cada captura. Arquivos associados aos cenários benignos receberam o rótulo *Benign*, enquanto arquivos gerados nos cenários de exfiltração receberam o rótulo *Exfiltration*.

Essa estratégia foi adotada porque o tráfego foi produzido em ambiente controlado, com separação explícita entre os cenários benignos e maliciosos. Dessa forma, o rótulo de cada amostra deriva do procedimento experimental que originou o arquivo PCAP/PCAPNG, preservando a rastreabilidade entre captura, cenário e instância tabular.

Portanto, o pipeline *PCAP-to-Flow* constitui uma etapa central da metodologia proposta. Ele garante que os conjuntos *UFU-Do53-EXF* e *UFU-DoH-EXF* sejam derivados de capturas brutas de rede por meio de um processo explícito, rastreável e reprodutível. Ao mesmo tempo, preserva a especificidade de cada protocolo: no Do53, explora atributos diretamente observáveis nas mensagens DNS e no DoH, utiliza estatísticas de fluxo compatíveis com o cenário de tráfego criptografado.

4.6. Atributos Do53

Os 29 atributos do *UFU-Do53-EXF* foram definidos com base no conjunto de atributos proposto por Mahdavifar et al. [Mahdavifar et al. 2021], que apresenta uma abordagem híbrida e leve para detecção de exfiltração de dados via DNS por meio da combinação de atributos *stateless* e *stateful*. No presente trabalho, esse conjunto foi adaptado ao desenho experimental adotado e à estrutura dos artefatos gerados, abrangendo propriedades léxicas do nome consultado, metadados DNS e estatísticas agregadas por domínio. Essa escolha é coerente com a literatura que associa o abuso de DNS à irregularidade lexical, ao aumento de entropia, à presença de codificação artificial em rótulos de subdomínio e a padrões de resolução não usuais [Ahmed et al. 2020, Tang et al. 2020, Palau et al. 2019, Nadler et al. 2019, Mahdavifar et al. 2021]. A Tabela 2 apresenta o conjunto de atributos adotado para o *UFU-Do53-EXF*.

Atributo	**Descrição**
`query_length`	Comprimento do nome de domínio completo (*FQDN*).
`subdomain_length`	Comprimento do subdomínio, correspondente à primeira *label* do *FQDN*.
`num_labels`	Número de *labels* presentes no *FQDN*.
`max_label_length`	Comprimento máximo entre as *labels* do *FQDN*.
`avg_label_length`	Comprimento médio das *labels* do *FQDN*.
`entropy`	Entropia do *FQDN*, utilizada como medida de aleatoriedade dos caracteres.
`ratio_digits`	Proporção de dígitos no nome consultado, desconsiderando os pontos.
`ratio_uppercase`	Proporção de caracteres maiúsculos no nome consultado, desconsiderando os pontos.
`ratio_lowercase`	Proporção de caracteres minúsculos no nome consultado, desconsiderando os pontos.
`ratio_special_chars`	Proporção de caracteres não alfanuméricos, excetuando ponto e hífen.
`rr_type`	Tipo do registro DNS consultado ou retornado.
`query_class`	Classe da consulta DNS.
`ttl`	Valor de *Time To Live* observado nos registros retornados, quando aplicável.
`contains_base64_pattern`	Indicador da presença de padrão compatível com codificação Base32/Base64 no domínio.
`contains_hex_pattern`	Indicador da presença de sequências hexadecimais longas no domínio.
`contains_dictionary_word`	Indicador da ocorrência de palavras comuns.
`ratio_vowels`	Proporção de vogais no nome consultado.
`ratio_consonants`	Proporção de consoantes no nome consultado.
`num_distinct_characters`	Número de caracteres distintos presentes no nome consultado.
`mean_query_length_per_domain`	Comprimento médio das consultas agrupadas por domínio-base.
`std_query_length_per_domain`	Desvio padrão do comprimento das consultas agrupadas por domínio-base.
`mean_entropy_per_domain`	Entropia média das consultas agrupadas por domínio-base.
`num_unique_ips`	Número de endereços IP distintos observados por domínio-base.
`num_unique_asns`	Número de sistemas autônomos distintos observados por domínio-base.
`num_unique_ttls`	Número de valores de TTL distintos observados por domínio-base.
`mean_ttl`	Valor médio de TTL por domínio-base.
`var_ttl`	Variância do TTL por domínio-base.
`num_queries_per_domain`	Número total de consultas observadas por domínio-base.
`ratio_failed_queries`	Proporção de consultas com falha por domínio-base.

Tabela 2. Conjunto de 29 atributos definido para o *UFU-Do53-EXF*.

4.7. Atributos DoH

Os 28 atributos do *UFU-DoH-EXF* foram definidos a partir de uma adaptação da abordagem de Zhan et al. [Zhan and others 2022], voltada à detecção de exfiltração via DNS

sobre HTTPS. Como, em DoH, o conteúdo semântico das consultas e respostas DNS não está diretamente disponível para inspeção, os autores exploram características observáveis da comunicação, como tamanhos de registros TLS, intervalos temporais e direção dos pacotes. Neste trabalho, essa lógica foi adaptada ao *UFU-DoH-EXF*, priorizando atributos de fluxo relacionados à direção dos pacotes, tamanhos observáveis, duração, volume e estatísticas temporais.

Com a padronização do TLS Encrypted Client Hello pelo RFC 9849 em 2026, parte das informações tradicionalmente extraídas do *ClientHello* em claro tende a perder disponibilidade para mecanismos de inspeção passiva. Em especial, o SNI real, o ALPN real e o *fingerprint* JA3 associado ao *ClientHello* efetivo podem ficar indisponíveis, pois passam a integrar o *ClientHelloInner* criptografado. Nesses cenários, sensores de rede observarão apenas o *ClientHelloOuter*, que pode conter valores públicos, genéricos ou inócuos, reduzindo a confiabilidade dessas *features* para classificação. Assim, abordagens de detecção de exfiltração via DoH devem priorizar atributos independentes do conteúdo do *ClientHello*, como estatísticas de fluxo, direção, tamanho, duração, volume e comportamento temporal [Rescorla et al. 2026].

Essa escolha metodológica é coerente com a premissa de que, em cenários DoH, a detecção deve explorar sinais indiretos do comportamento da comunicação, em vez de depender da inspeção direta dos nomes de domínio consultados. Assim, os atributos do *UFU-DoH-EXF* foram estruturados para capturar padrões associados à distribuição direcional dos pacotes, às variações de tamanho, à regularidade temporal, à duração dos fluxos e ao volume observado, elementos que podem diferenciar tráfego benigno de fluxos utilizados para exfiltração de dados. A Tabela 3 apresenta o conjunto de atributos adotado para o *UFU-DoH-EXF* e a Tabela 4 apresenta o sumário quantitativo dos dois conjuntos de dados construídos.

Atributo	**Descrição**
`src2dst_min_length`	Tamanho mínimo dos pacotes no sentido cliente→servidor.
`src2dst_max_length`	Tamanho máximo dos pacotes no sentido cliente→servidor.
`src2dst_mean_length`	Tamanho médio dos pacotes no sentido cliente→servidor.
`src2dst_stddev_length`	Desvio padrão do tamanho dos pacotes no sentido cliente→servidor.
`dst2src_min_length`	Tamanho mínimo dos pacotes no sentido servidor→cliente.
`dst2src_max_length`	Tamanho máximo dos pacotes no sentido servidor→cliente.
`dst2src_mean_length`	Tamanho médio dos pacotes no sentido servidor→cliente.
`dst2src_stddev_length`	Desvio padrão do tamanho dos pacotes no sentido servidor→cliente.
`bidirectional_max_length`	Maior tamanho de pacote considerando ambas as direções do fluxo.
`bidirectional_mean_length`	Tamanho médio de pacote considerando ambas as direções do fluxo.
`bidirectional_stddev_length`	Desvio padrão do tamanho dos pacotes no fluxo bidirecional.
`src2dst_min_time`	Menor intervalo temporal entre pacotes no sentido cliente→servidor.
`src2dst_max_time`	Maior intervalo temporal entre pacotes no sentido cliente→servidor.
`src2dst_mean_time`	Intervalo temporal médio entre pacotes no sentido cliente→servidor.
`src2dst_stddev_time`	Desvio padrão do intervalo temporal entre pacotes no sentido cliente→servidor.
`dst2src_min_time`	Menor intervalo temporal entre pacotes no sentido servidor→cliente.
`dst2src_max_time`	Maior intervalo temporal entre pacotes no sentido servidor→cliente.
`dst2src_mean_time`	Intervalo temporal médio entre pacotes no sentido servidor→cliente.
`dst2src_stddev_time`	Desvio padrão do intervalo temporal entre pacotes no sentido servidor→cliente.
`bidirectional_min_time`	Menor intervalo temporal observado no fluxo bidirecional.
`bidirectional_max_time`	Maior intervalo temporal observado no fluxo bidirecional.
`bidirectional_mean_time`	Intervalo temporal médio considerando o fluxo bidirecional.
`bidirectional_stddev_time`	Desvio padrão dos intervalos temporais no fluxo bidirecional.
`num_src2dst`	Número de pacotes no sentido cliente→servidor.
`num_dst2src`	Número de pacotes no sentido servidor→cliente.
`flow_duration`	Duração total do fluxo TLS.
`byte_count`	Total de bytes observados no fluxo bidirecional.
`handshake_time`	Tempo aproximado do *handshake* TLS.

Tabela 3. Conjunto de 28 atributos definido para o *UFU-DoH-EXF*.

Em relação ao conjunto de atributos originalmente explorado por Zhan et

al. [Zhan and others 2022], este trabalho preserva a ênfase em estatísticas de tamanho, direção e intervalos temporais, mas incorpora atributos adicionais relacionados à contagem de pacotes, duração, volume e tempo aproximado de estabelecimento da comunicação. Especificamente, foram acrescentadas as features `num_src2dst`, `num_dst2src`, `flow_duration`, `byte_count` e `handshake_time`. Essa ampliação é justificada pelo desenho experimental adotado neste trabalho, no qual as amostras DoH são extraídas a partir de janelas temporais dentro das comunicações TCP/443. Essa escolha reduz o risco de que conexões longas agreguem comportamentos heterogêneos em uma única instância e permite capturar variações locais associadas à intensidade, regularidade e assimetria do tráfego de exfiltração.

Tabela 4. Distribuição das amostras nos conjuntos de dados propostos.

Conjunto de dados	Atributos	Amostras	Benignos	Exfiltrados	% Exfiltrados
UFU-Do53-EXF	29	5.351.064	4.812.119	538.945	10,0706%
UFU-DoH-EXF	28	1.048.575	1.016.438	32.137	3,0648%

5. Estudos de caso

Os estudos de caso têm como objetivo fornecer uma avaliação supervisionada de referência para os conjuntos propostos. Em ambos os cenários, foi adotado particionamento estratificado 80/20, preservando a proporção entre as classes. As métricas foram escolhidas considerando o desbalanceamento dos dados: além da acurácia, são reportadas acurácia balanceada, métricas macro e os números absolutos de falsos positivos e falsos negativos. Essa escolha evita que o desempenho seja interpretado apenas a partir da classe majoritária. Além disso, é importante observar que as duas avaliações apresentadas aqui constituem experimentos intra-base, baseados em particionamentos estratificados aleatórios. Assim, os resultados devem ser interpretados como uma avaliação inicial da consistência e da capacidade discriminativa dos atributos extraídos, e não como evidência definitiva de generalização para outros ambientes.

5.1. *UFU-Do53-EXF*

O primeiro estudo de caso avalia o *UFU-Do53-EXF* em uma tarefa supervisionada de classificação binária entre tráfego *Benign* e *Exfiltration*. Foram utilizados os modelos *Random Forest*, *XGBoost* e *Support Vector Machine* (SVM), amplamente empregados em detecção baseada em aprendizado de máquina e em estudos de exfiltração via DNS [Mahdavifar et al. 2021].

Os experimentos utilizaram particionamento estratificado de 80% para treinamento e 20% para teste, preservando a proporção entre as classes. O conjunto de teste contém 1.070.214 instâncias, sendo 962.436 da classe *Benign* e 107.778 da classe *Exfiltration*. A variável-alvo `label` foi codificada por meio de *Label Encoding*, e o conjunto de teste foi alinhado ao esquema do conjunto de treinamento. Para o SVM, aplicou-se adicionalmente padronização dos atributos numéricos com *StandardScaler*, em razão da sensibilidade do modelo à escala dos dados. O *Random Forest* foi configurado com 100 árvores, *random state* igual a 50 e `class_weight="balanced"`. O *XGBoost* utilizou 300 estimadores, profundidade máxima igual a 6, taxa de aprendizado de 0,1, `subsample`

de 0,8, `colsample_bytree` de 0,8 e *random state* igual a 50. O SVM foi configurado com *kernel* radial, $C = 1,0$, *gamma = scale* e `class_weight="balanced"`. As métricas consideradas foram *Accuracy*, *Balanced Accuracy*, *Precision macro*, *Recall macro*, *F1-score macro*, *F1-score* ponderado e matriz de confusão. A Tabela 5 sumariza os resultados encontrados.

Tabela 5. Comparação entre os experimentos supervisionados conduzidos sobre o *UFU-Do53-EXF*.

Modelo	Acc.	Bal. Acc.	Prec.	Rec.	F1	F1 pond.	FP	FN
Random Forest	0,999996	0,999981	0,999998	0,999981	0,999990	0,999996	0	4
XGBoost	0,999998	0,999991	0,999999	0,999991	0,999995	0,999998	0	2
SVM	0,999915	0,999817	0,999714	0,999817	0,999765	0,999915	58	33

Os resultados indicam desempenho elevado na avaliação intra-base do *UFU-Do53-EXF*. O *XGBoost* obteve o melhor resultado, com apenas 2 falsos negativos e nenhum falso positivo no conjunto de teste, enquanto o *Random Forest* apresentou 4 falsos negativos e também nenhum falso positivo. O SVM manteve desempenho competitivo, mas gerou maior número de erros residuais. Em termos de detecção, os poucos falsos negativos observados sugerem que os atributos léxicos, estatísticos e comportamentais extraídos das consultas DNS preservam sinais relevantes da exfiltração. Contudo, como o experimento utiliza particionamento estratificado aleatório, os resultados devem ser interpretados como evidência de consistência intra-base, e não como garantia de generalização.

5.2. *UFU-DoH-EXF*

O segundo estudo de caso avalia o *UFU-DoH-EXF* em uma tarefa supervisionada de classificação binária entre fluxos *Benign* e *Exfiltration*. Diferentemente do Do53, o cenário DoH impõe menor observabilidade semântica, pois as consultas DNS são encapsuladas em HTTPS/TLS. Assim, a avaliação utiliza atributos estatísticos de fluxo, direção dos pacotes, tamanho, duração, volume e comportamento temporal, em linha com abordagens de detecção em tráfego DoH criptografado [Zhan and others 2022]. Foram avaliados os modelos *Random Forest*, *XGBoost* e *Support Vector Machine*. Os experimentos utilizaram particionamento estratificado de 80% para treinamento e 20% para teste. Antes da divisão, foram removidas duplicatas exatas, resultando em 1.022.257 fluxos TCP/443, dos quais 992.111 pertencem à classe *Benign* e 30.146 à classe *Exfiltration*. O conjunto de teste contém 204.452 instâncias, sendo 198.423 benignas e 6.029 de exfiltração.

A variável-alvo `label` foi codificada por *Label Encoding*. Para *Random Forest* e *XGBoost*, os atributos numéricos foram tratados com imputação pela mediana quando necessário. Para o Linear SVM, aplicou-se também padronização com *StandardScaler*. O *Random Forest* foi configurado com 100 árvores, *random state* igual a 42 e `class_weight="balanced"`. O *XGBoost* utilizou 300 estimadores, profundidade máxima igual a 6, taxa de aprendizado de 0,1, `subsample` de 0,8, `colsample_bytree` de 0,8, objetivo `binary:logistic`, métrica `logloss` e `scale_pos_weight` de aproximadamente 32,91. O Linear SVM foi configurado com $C = 1{,}0$, `max_iter = 10000` e `class_weight="balanced"`. A Tabela 6 sumariza os resultados encontrados.

Os resultados também indicam desempenho elevado para o cenário DoH, mesmo sem acesso ao conteúdo semântico das consultas DNS. *Random Forest* e *XGBoost* apre-

Tabela 6. Comparação entre os experimentos supervisionados conduzidos sobre o *UFU-DoH-EXF*.

Modelo	Acc.	Bal. Acc.	Prec.	Rec.	F1	F1 pond.	FP	FN
Random Forest	0,999985	0,999832	0,999912	0,999832	0,999872	0,999985	1	2
XGBoost	0,999985	0,999832	0,999912	0,999832	0,999872	0,999985	1	2
SVM	0,999863	0,999688	0,997927	0,999688	0,998806	0,999863	25	3

sentaram a mesma matriz de confusão, mas não erraram exatamente as mesmas instâncias, enquanto o Linear SVM produziu mais falsos positivos. Esse resultado sugere que atributos de fluxo, tamanho, direção, duração e temporalidade capturam efeitos observáveis da exfiltração mesmo em tráfego criptografado. Ainda assim, por se tratar de uma avaliação intra-base, os resultados devem ser compreendidos como uma linha de base supervisionada para o conjunto proposto. Avaliações futuras devem considerar particionamentos por sessão, por período de captura, por domínio ou por ferramenta de exfiltração, a fim de medir generalização sob cenários mais próximos de implantação real.

6. Conclusão

Este artigo apresentou dois conjuntos de dados complementares para detecção de exfiltração via DNS: *UFU-Do53-EXF* e *UFU-DoH-EXF*. A principal contribuição do trabalho é metodológica e experimental. Em vez de tratar Do53 e DoH como problemas isolados, amparados por bases heterogêneas e pouco comparáveis, a proposta estabelece um arcabouço comum de geração, captura e transformação dos dados, com tráfego benigno e malicioso controlado, preservação de artefatos em PCAP, conversão reprodutível para fluxo de rede, definição explícita de atributos e foco direto no problema de exfiltração.

Os resultados quantitativos evidenciam que ambos os conjuntos possuem escala relevante para estudos de detecção baseados em aprendizado de máquina. O *UFU-Do53-EXF* reúne 5.351.064 amostras após normalização e apresenta perfil adequado para avaliação supervisionada sob desbalanceamento de classes. O *UFU-DoH-EXF*, por sua vez, contém 1.048.575 fluxos TCP/443, sendo 1.016.438 amostras benignas e 32.137 amostras de exfiltração. Em particular, os resultados experimentais com o *UFU-DoH-EXF* mostram que, mesmo sem acesso ao conteúdo semântico das consultas DNS, atributos estatísticos de fluxo, direção, tamanho, duração, volume e comportamento temporal podem capturar efeitos observáveis da exfiltração em tráfego criptografado.

Como trabalhos futuros, destacam-se quatro frentes. A primeira visa fortalecer a avaliação experimental por meio de validação cruzada estratificada com 5 e 10 *folds*, além de particionamentos por sessão, período de captura e ferramenta de exfiltração, permitindo distinguir a consistência dos resultados da capacidade de generalização. A segunda propõe incorporar e avaliar, por meio de análise de ablação, os metadados TLS derivados do *ClientHello* (SNI, ALPN e JA3), para medir sua influência na separabilidade das classes e verificar se a detecção permanece baseada em características comportamentais mesmo em cenários com *Encrypted Client Hello* [Rescorla et al. 2026]. A terceira contempla avaliar a sensibilidade da detecção em cenários DoH sob diferentes condições de rede e implementação, incluindo latência, *jitter*, resolvedores, pilhas TLS/HTTP, além de testes de robustez com ruído e reamostragem. Por fim, pretende-se realizar validação entre bases públicas e privadas e desenvolver uma arquitetura integrada para detecção de exfiltração via DNS baseada em aprendizado de máquina.

Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) projetos 409743/2025-9 e 408432/2024-1.

Referências

Abualghanam, O., Alazzam, H., Elshqeirat, B., Qatawneh, M., and Almaiah, M. A. (2023). Real-Time Detection System for Data Exfiltration over DNS Tunneling Using Machine Learning. *Electronics*, 12(6).

Ahmed, J., Gharakheili, H. H., Raza, Q., Russell, C., and Sivaraman, V. (2019). Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks. In *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, pages 649–653.

Ahmed, J., Habibi Gharakheili, H., Raza, Q., Russell, C., and Sivaraman, V. (2020). Monitoring Enterprise DNS Queries for Detecting Data Exfiltration From Internal Hosts. *IEEE Transactions on Network and Service Management*, 17(1):265–279.

Bowes, R. (2022). dnscat2.

Elaoumari, A. (2025). *Evasion-Resilient Detection of DNS-over-HTTPS Data Exfiltration: A Practical Evaluation and Toolkit*. MSc Dissertation, University of Kent, School of Computing.

Garcia, S. and Valeros, V. (2023). Towards a better labeling process for network security datasets. _eprint: 2305.01337.

Güneş Gürsoy, A., Varol, A., and Nasab, A. (2024). DNS Tunnel Problem In Cybersecurity. In *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*. IEEE.

Hoffman, P. and McManus, P. (2018). DNS Queries over HTTPS (DoH). Issue: 8484 Type: RFC Published: Internet Requests for Comments.

Le Pochat, V., Van Goethem, T., Tajalizadehkhoob, S., Korczyński, M., and Joosen, W. (2019). Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*.

Luz, J., Araujo-Filho, P., Arcoverde, H., and Campelo, D. (2023). Unsupervised SOM-Based Intrusion Detection System for DNS Tunneling Attacks. In *Anais do XXIII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais*, pages 516–521, Porto Alegre, RS, Brasil. SBC.

Mahdavifar, S., Salem, A., Victor, P., Razavi, A., Garzon, M., Hellberg, N., and Habibi Lashkari, A. (2021). Lightweight Hybrid Detection of Data Exfiltration using DNS based on Machine Learning. In *ICCNS 2021: The 11th International Conference on Communication and Network Security*, pages 80–86.

MITRE ATT&CK (2026a). T1041 – Exfiltration Over C2 Channel.

MITRE ATT&CK (2026b). T1048 – Exfiltration Over Alternative Protocol.

MITRE ATT&CK (2026c). T1048.003 – Exfiltration Over Unencrypted Non-C2 Protocol.

Moure-Garrido, M., Campo, C., and Garcia-Rubio, C. (2023). Real time detection of malicious DoH traffic using statistical analysis. *Computer Networks*, 234:109910.

Nadler, A., Aminov, A., and Shabtai, A. (2019). Detection of malicious and low throughput data exfiltration over the DNS protocol. *Computers & Security*, 80:36–53.

Ozery, Y., Nadler, A., and Shabtai, A. (2024). Information-Based Heavy Hitters for Real-Time DNS Data Exfiltration Detection. In *Network and Distributed System Security Symposium (NDSS) 2024*.

Palau, F., Catania, C., Guerra, J., García, S. J., and Rigaki, M. (2019). Detecting DNS Threats: A Deep Learning Model to Rule Them All. In *XX Simposio Argentino de Inteligencia Artificial (ASAI 2019) - JAIIO 48*, pages 90–101, Salta, Argentina. Universidad Nacional de La Plata. Backup Publisher: Sociedad Argentina de Informática e Investigación Operativa.

Rescorla, E. (2018). The Transport Layer Security (TLS) Protocol Version 1.3. Published: RFC 8446.

Rescorla, E., Oku, K., Sullivan, N., and Wood, C. A. (2026). TLS Encrypted Client Hello. RFC 9849, RFC Editor.

Sabir, B., Ullah, F., Babar, M. A., and Gaire, R. (2021). Machine Learning for Detecting Data Exfiltration: A Review. *ACM Comput. Surv.*, 54(3). Place: New York, NY, USA.

Salat, L., Davis, M., and Khan, N. (2023). DNS Tunnelling, Exfiltration and Detection over Cloud Environments. *Sensors*, 23(2760).

Srivastava, G., Jhaveri, R. H., Bhattacharya, S., Pandya, S., Rajeswari, Maddikunta, P. K. R., Yenduri, G., Hall, J. G., Alazab, M., and Gadekallu, T. R. (2022). XAI for Cybersecurity: State of the Art, Challenges, Open Issues and Future Directions. _eprint: 2206.03585.

Steadman, J. and Scott-Hayward, S. (2018). DNSxD: Detecting Data Exfiltration Over DNS. In *2018 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, pages 1–6.

Tang, R., Huang, C., Zhou, Y., Wu, H., Lu, X., Sun, Y., Li, Q., Li, J., Huang, W., Sun, S., and others (2020). A practical machine learning-based framework to detect dns covert communication in enterprises. In *International Conference on Security and Privacy in Communication Systems*, pages 1–21. Springer.

Zhan, M. and others (2022). Detecting DNS over HTTPS based data exfiltration. *Computer Networks*.