Tempest

ACADEMY

Conference

2023

Tunelamento DNS

Ataque e Detecção usando aprendizado de máquina

Classificação da Informação: PÚBLICA
Autor da apresentação: Júlio Luz

whoami

Tempest
[ACADEMY]
Conference

Estagiário do SOC (2022) -> Estagiário de P&D (2023)

Estudante de Engenharia da Computação
CIn - UFPE

Membro do grupo de pesquisa de detecção de ciber ataques através de técnicas de inteligência artificial (CIn/Tempest)

Centro de Informática
UFPE

Tempest

whoami

Tempest
[ACADEMY]
Conference

Estagiário do SOC (2022) -> Estagiário de P&D (2023)

Estudante de Engenharia da Computação
CIn - UFPE

Membro do grupo de pesquisa de detecção de ciber ataques através de técnicas de inteligência artificial (CIn/Tempest)

→ Pesquisa: Detecção de tunelamento DNS

Centro de Informática UFPE

Tempest

01 Conceitos iniciais

02 Simulação prática do ataque

03 Detecção

04 Resultados

Tempest
[ACADEMY]
Conference

Mas afinal, o que é “tunelamento DNS”?

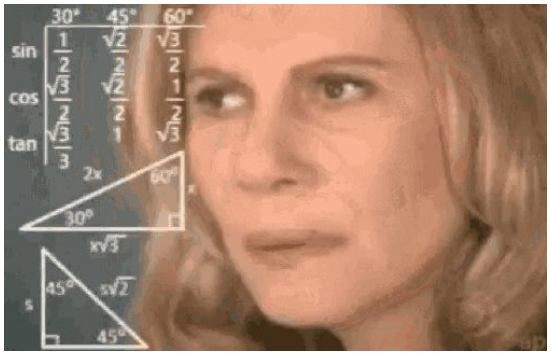
Conceitos iniciais

Mas afinal, o que é “Tunelamento DNS”?

Tempest
[ACADEMY]
Conference

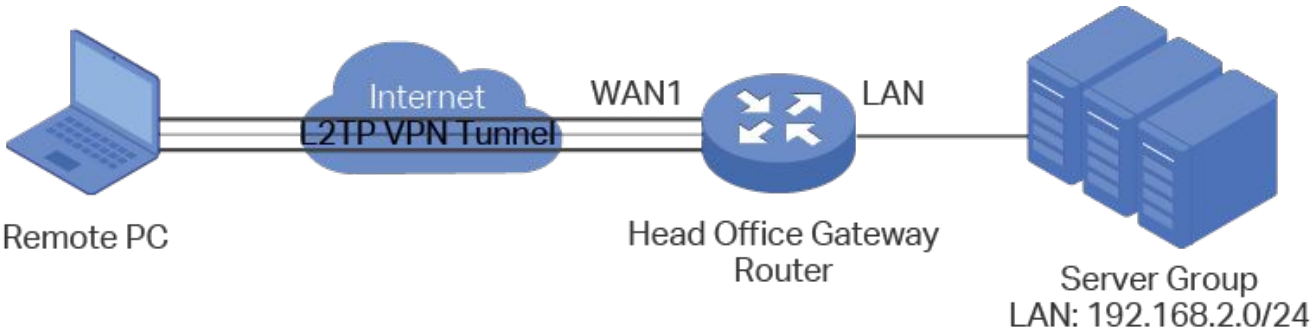
O que é tunelamento?

O que é DNS?



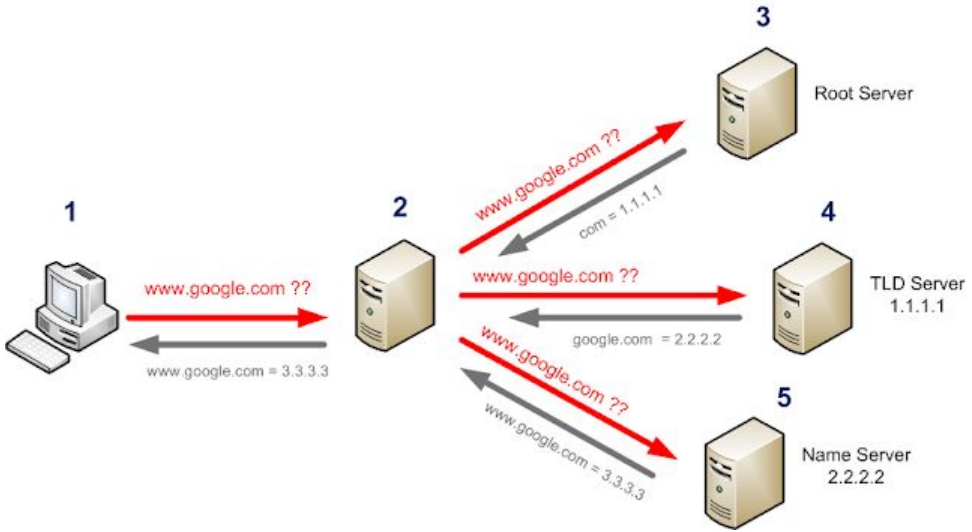
O que é tunelamento?

Método usado para transferir dados contidos por um protocolo usando outro protocolo, através do *encapsulamento* de dados.



O que é DNS?

(1) Base de dados distribuída implementada em uma hierarquia de servidores DNS, (2) Protocolo da camada de aplicação que permite que usuários consultem a base de dados distribuída.



Tempest
[ACADEMY]
Conference

Finalmente... O que é tunelamento DNS?

Finalmente... O que é tunelamento DNS?

Tempest
[ACADEMY]
Conference

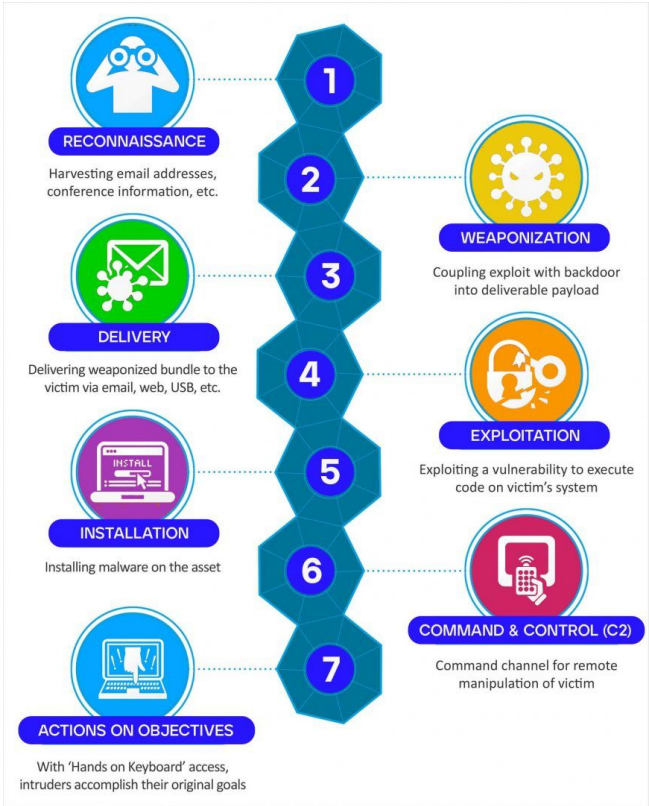
É uma técnica que consiste em usar o protocolo DNS para transferir quaisquer dados desejados, encapsulando os mesmos em campos do protocolo DNS.

Por que isso é importante?

Tempest
[ACADEMY]
Conference

Permite a criação de um canal de comunicação **discreto**, possibilitando:

- Configuração de um canal de comando e controle
- Exfiltração de dados
- Transmissão de malwares

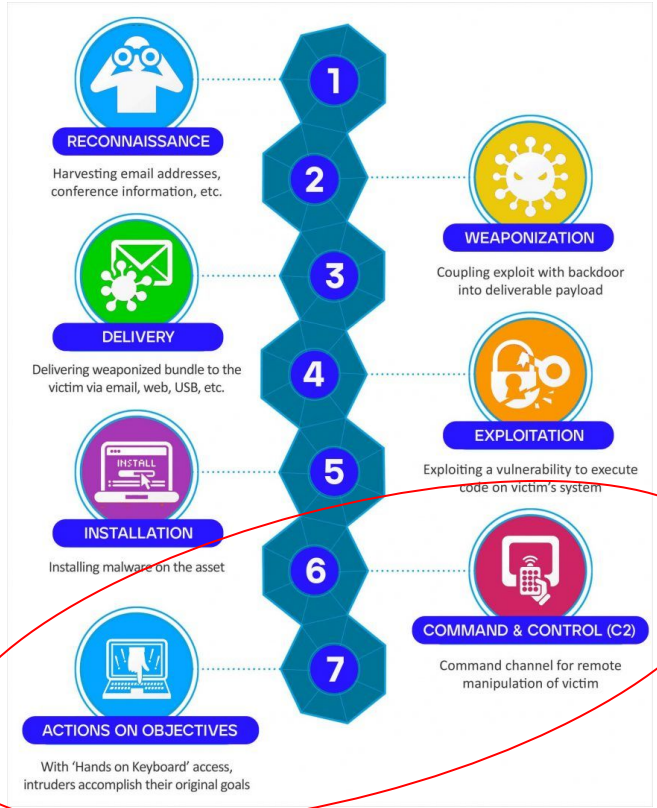


Por que isso é importante?

Tempest [ACADEMY] Conference

Permite a criação de um canal de comunicação **discreto**, possibilitando:

- Configuração de um canal de comando e controle
- Exfiltração de dados
- Transmissão de malwares



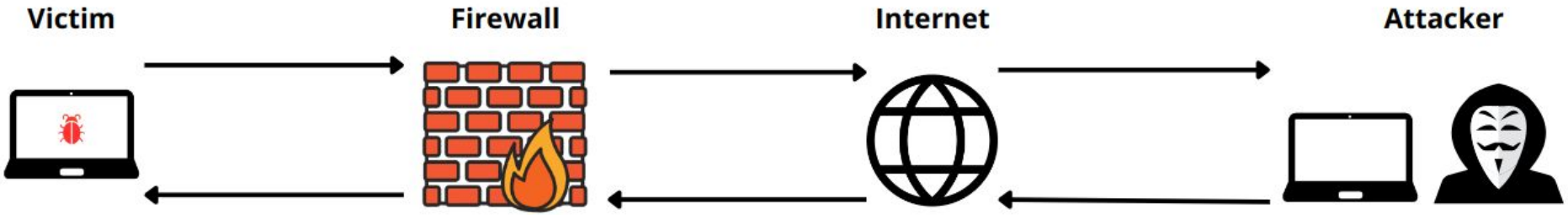
Estamos aqui

O que é tunelamento DNS?

Tempest
ACADEMY
Conference

DNS Query:
on2g63dfnzsgc5db.attacker-domain.com

Attacker receives exfiltrated data:
base32_dec(on2g63dfnzsgc5db) -> "stolendata"

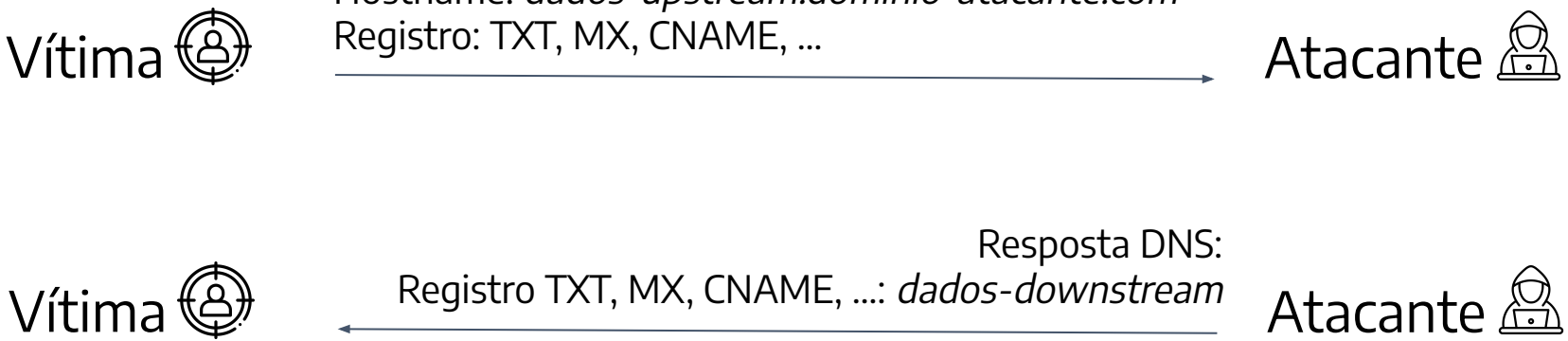


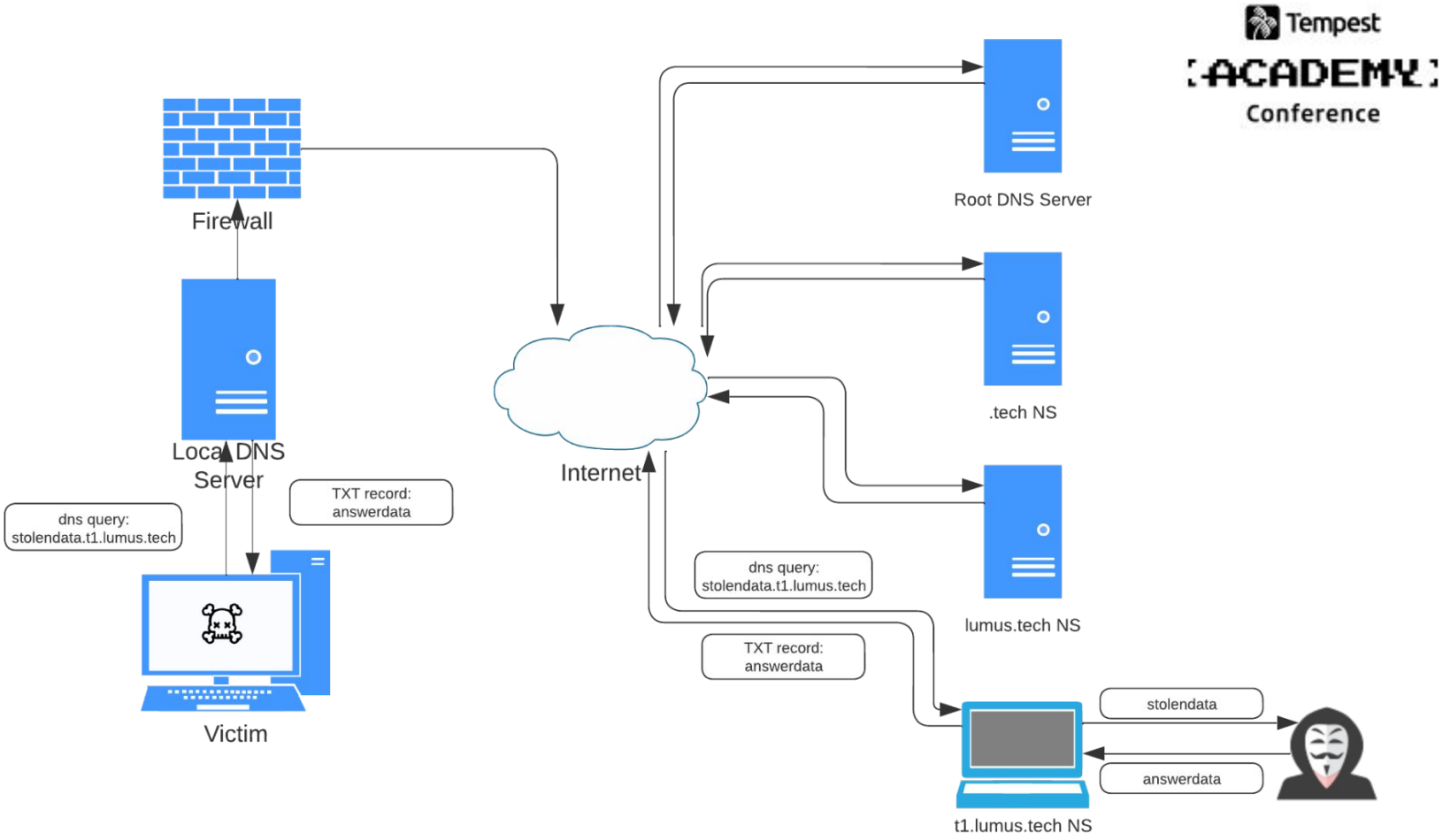
Victim receives command:
base64_dec(KNSW4ZBAMZUWYZJOOR4HI===) -> "Send file.txt"

DNS Response:
KNSW4ZBAMZUWYZJOOR4HI===

Como funciona?

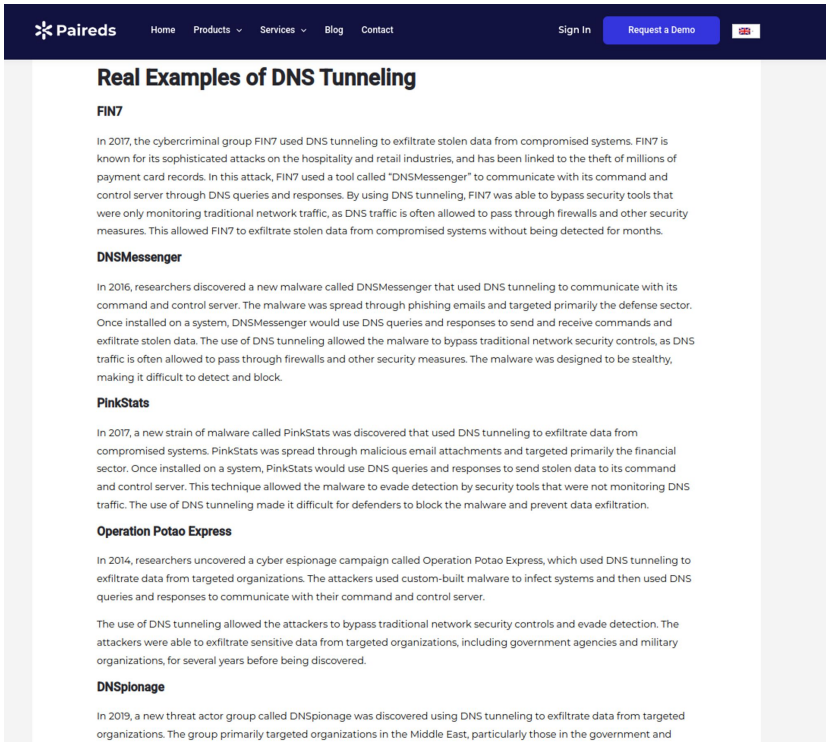
Tempest
[ACADEMY]
Conference





Exemplos de uso

Tempest
[ACADEMY]
Conference



Paireds Home Products Services Blog Contact Sign In Request a Demo

Real Examples of DNS Tunneling

FIN7

In 2017, the cybercriminal group FIN7 used DNS tunneling to exfiltrate stolen data from compromised systems. FIN7 is known for its sophisticated attacks on the hospitality and retail industries, and has been linked to the theft of millions of payment card records. In this attack, FIN7 used a tool called "DNSMessenger" to communicate with its command and control server through DNS queries and responses. By using DNS tunneling, FIN7 was able to bypass security tools that were only monitoring traditional network traffic, as DNS traffic is often allowed to pass through firewalls and other security measures. This allowed FIN7 to exfiltrate stolen data from compromised systems without being detected for months.

DNSMessenger

In 2016, researchers discovered a new malware called DNSMessenger that used DNS tunneling to communicate with its command and control server. The malware was spread through phishing emails and targeted primarily the defense sector. Once installed on a system, DNSMessenger would use DNS queries and responses to send and receive commands and exfiltrate stolen data. The use of DNS tunneling allowed the malware to bypass traditional network security controls, as DNS traffic is often allowed to pass through firewalls and other security measures. The malware was designed to be stealthy, making it difficult to detect and block.

PinkStats

In 2017, a new strain of malware called PinkStats was discovered that used DNS tunneling to exfiltrate data from compromised systems. PinkStats was spread through malicious email attachments and targeted primarily the financial sector. Once installed on a system, PinkStats would use DNS queries and responses to send stolen data to its command and control server. This technique allowed the malware to evade detection by security tools that were not monitoring DNS traffic. The use of DNS tunneling made it difficult for defenders to block the malware and prevent data exfiltration.

Operation Potao Express

In 2014, researchers uncovered a cyber espionage campaign called Operation Potao Express, which used DNS tunneling to exfiltrate data from targeted organizations. The attackers used custom-built malware to infect systems and then used DNS queries and responses to communicate with their command and control server.

The use of DNS tunneling allowed the attackers to bypass traditional network security controls and evade detection. The attackers were able to exfiltrate sensitive data from targeted organizations, including government agencies and military organizations, for several years before being discovered.

DNSpionage

In 2019, a new threat actor group called DNSpionage was discovered using DNS tunneling to exfiltrate data from targeted organizations. The group primarily targeted organizations in the Middle East, particularly those in the government and

Exemplos de uso

Tempest [ACADEMY] Conference

IronNet

Why IronNet | Platform | Industries | Company | Resources

SUNBURST: a case for DNS Tunneling

This second look leads us to the DNS Tunneling angle. The use case for DNS Tunneling is to enable communication between malware and C2 servers over the DNS protocol. Again, with SUNBURST, research around the structure and content of the DNS queries to "avsvmcloud[.]com" has shown that the lowest level subdomain label used for these queries is encoded data that corresponds to the active directory domain name of the infected network. This does not lend itself to the DGA use case, as the top domain under the registry suffix is not changing and makes blocking such traffic at the firewall trivial. This does, however, provide the threat actor with accurate information about which network — and possibly even which infected host — was making the query, a critical function when managing a vast number of infections across a broad set of environments. Furthermore, the responses to these queries are not indicative of the actual IP addresses for C2 servers. Rather, they indicate the command or action that the threat actor wants the malicious implants to take. This is exactly the way DNS Tunneling functions.

Paireds

Home | Products | Services | Blog | Contact | Sig

Real Examples of DNS Tunneling

FIN7

In 2017, the cybercriminal group FIN7 used DNS tunneling to exfiltrate stolen data from compromi
known for its sophisticated attacks on the hospitality and retail industries, and has been linked to
payment card records. In this attack, FIN7 used a tool called "DNSMessenger" to communicate wit
control server through DNS queries and responses. By using DNS tunneling, FIN7 was able to byp
were only monitoring traditional network traffic, as DNS traffic is often allowed to pass through fire
measures. This allowed FIN7 to exfiltrate stolen data from compromised systems without being de

DNSMessenger

In 2016, researchers discovered a new malware called DNSMessenger that used DNS tunneling to
command and control server. The malware was spread through phishing emails and targeted prir
Once installed on a system, DNSMessenger would use DNS queries and responses to send and rec
exfiltrate stolen data. The use of DNS tunneling allowed the malware to bypass traditional network
traffic is often allowed to pass through firewalls and other security measures. The malware was de
making it difficult to detect and block.

PinkStats

In 2017, a new strain of malware called PinkStats was discovered that used DNS tunneling to exfiltrate data from compromised systems. PinkStats was spread through malicious email attachments and targeted primarily the financial sector. Once installed on a system, PinkStats would use DNS queries and responses to send stolen data to its command and control server. This technique allowed the malware to evade detection by security tools that were not monitoring DNS traffic. The use of DNS tunneling made it difficult for defenders to block the malware and prevent data exfiltration.

Operation Potao Express

In 2014, researchers uncovered a cyber espionage campaign called Operation Potao Express, which used DNS tunneling to exfiltrate data from targeted organizations. The attackers used custom-built malware to infect systems and then used DNS queries and responses to communicate with their command and control server.

The use of DNS tunneling allowed the attackers to bypass traditional network security controls and evade detection. The attackers were able to exfiltrate sensitive data from targeted organizations, including government agencies and military organizations, for several years before being discovered.

DNSpionage

In 2019, a new threat actor group called DNSpionage was discovered using DNS tunneling to exfiltrate data from targeted organizations. The group primarily targeted organizations in the Middle East, particularly those in the government and

Exemplos de uso

Tempest [ACADEMY] Conference

IronNet

Why IronNet | Platform | Industries | Company | Resources

SUNBURST: a case for DNS Tunneling

This second look leads us to the DNS Tunneling angle. The use case for DNS Tunneling is to enable communication between malware and C2 servers over the DNS protocol. Again, with SUNBURST, to "avsvmcloud[.]com" h queries is encoded data infected network. This d under the registry suffix trivial. This does, howev which network — and po critical function when m environments. Furtherm actual IP addresses for the threat actor wants t Tunneling functions.

Paireds

Home | Products | Services | Blog | Contact | Sig

Real Examples of DNS Tunneling

FIN7

In 2017, the cybercriminal group FIN7 used DNS tunneling to exfiltrate stolen data from compromi known for its sophisticated attacks on the hospitality and retail industries, and has been linked to payment card records. In this attack, FIN7 used a tool called "DNSMessenger" to communicate wit control server through DNS queries and responses. By using DNS tunneling, FIN7 was able to byp were only monitoring traditional network traffic, as DNS traffic is often allowed to pass through fire measures. This allowed FIN7 to exfiltrate stolen data from compromised systems without being de

DNSMessenger

In 2016, researchers discovered a new malware called DNSMessenger that used DNS tunneling to command and control server. The malware was spread through phishing emails and targeted prir Once installed on a system, DNSMessenger would use DNS queries and responses to send and rec exfiltrate stolen data. The use of DNS tunneling allowed the malware to bypass traditional network traffic is often allowed to pass through firewalls and other security measures. The malware was de making it difficult to detect and block.

PinkStats

In 2017, a new strain of malware called PinkStats was discovered that used DNS tunneling to exfiltrate data from compromised systems. PinkStats was spread through malicious email attachments and targeted primarily the financial sector. Once installed on a system, PinkStats would use DNS queries and responses to send stolen data to its command and control server. This technique allowed the malware to evade detection by security tools that were not monitoring DNS traffic. The use of DNS tunneling made it difficult for defenders to block the malware and prevent data exfiltration.

Operation Potao Express

In 2014, researchers uncovered a cyber espionage campaign called Operation Potao Express, which used DNS tunneling to exfiltrate data from targeted organizations. The attackers used custom-built malware to infect systems and then used DNS queries and responses to communicate with their command and control server.

The use of DNS tunneling allowed the attackers to bypass traditional network security controls and evade detection. The attackers were able to exfiltrate sensitive data from targeted organizations, including government agencies and military organizations, for several years before being discovered.

DNSpionage

In 2019, a new threat actor group called DNSpionage was discovered using DNS tunneling to exfiltrate data from targeted organizations. The group primarily targeted organizations in the Middle East, particularly those in the government and

UNIT 42

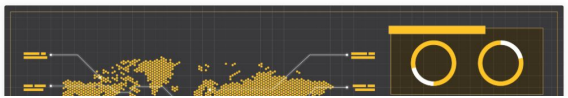
About Unit 42 | Services | Unit 42 Threat Research | Partners | Resources

xHunt Campaign: Newly Discovered Backdoors Using Deleted Email Drafts and DNS Tunneling for Command and Control

49,722 people reacted | 35 | 12 min. read

SHARE

By Robert Falcone
November 9, 2020 at 12:00 AM
Category: Unit 42
Tags: backdoor, C2, CASHY200, dns tunneling, Snugy, TriFive, xHunt



Exemplos de uso

Tempest [ACADEMY] Conference

IronNet

Why IronNet | Platform | Industries | Company | Resources

SUNBURST: a case for DNS Tunneling

This second look leads us to the DNS Tunneling angle. The use case for DNS Tunneling is to enable communication between malware and C2 servers over the DNS protocol. Again, with SUNBURST, to "avsvmcloud[.]com" h queries is encoded data infected network. This d under the registry suffix trivial. This does, howev which network — and pc critical function when m environments. Furtherm actual IP addresses for the threat actor wants th Tunneling functions.

Paireds

Home | Products | Services | Blog | Contact | Sig

Real Examples of DNS Tunneling

FIN7

In 2017, the cybercriminal group FIN7 used DNS tunneling to exfiltrate stolen data from compromi known for its sophisticated attacks on the hospitality and retail industries, and has been linked to t payment card records. In this attack, FIN7 used a tool called "DNSMessenger" to communicate wit control server through DNS queries and responses. By using DNS tunneling, FIN7 was able to bypa were only monitoring traditional network traffic, as DNS traffic is often allowed to pass through fire measures. This allowed FIN7 to exfiltrate stolen data from compromised systems without being de

DNSMessenger

In 2016, researchers discovered a new malware called DNSMessenger that used DNS tunneling to command and control server. The malware was spread through phishing emails and targeted prir Once installed on a system, DNSMessenger would use DNS queries and responses to send and rec exfiltrate stolen data. The use of DNS tunneling allowed the malware to bypass traditional network traffic is often allowed to pass through firewalls and other security measures. The malware was de making it difficult to detect and block.

PinkStats

In 2017, a new strain of malware called PinkStats was discovered that used DNS tunneling to exfiltrate data from compromised systems. PinkStats was spread through malicious email attachments and targeted primarily the financial sector. Once installed on a system, PinkStats would use DNS queries and responses to send stolen data to its command and control server. This technique allowed the malware to evade detection by security tools that were not monitoring DNS traffic. The use of DNS tunneling made it difficult for defenders to block the malware and prevent data exfiltration.

Operation Potao Express

In 2014, researchers uncovered a cyber espionage campaign called Operation Potao Express, which used DNS tunneling to exfiltrate data from targeted organizations. The attackers used custom-built malware to infect systems and then used DNS queries and responses to communicate with their command and control server.

The use of DNS tunneling allowed the attackers to bypass traditional network security controls and evade detection. The attackers were able to exfiltrate sensitive data from targeted organizations, including government agencies and military organizations, for several years before being discovered.

DNSpionage

In 2019, a new threat actor group called DNSpionage was discovered using DNS tunneling to exfiltrate data from targeted organizations. The group primarily targeted organizations in the Middle East, particularly those in the government and

UNIT 42 | About Unit 42 | Services | Unit 42 Threat Research | Partners | Resources

xHunt Campaign: Newly Discovered Backdoors Using Deleted Email Drafts and DNS Tunneling for Command and Control

49,722 people reacted | 35 | 12 min. read | SHARE

By Robert Falcone
November 9, 2020 at 12:00 AM
Category: Unit 42
Tags: backdoor, C2, CASHY200, dns tunneling, Snugy, TriFive, xHunt

UNIT 42 | About Unit 42 | Services | Unit 42 Threat Research | Partners | Resources

DNS Tunneling in the Wild: Overview of OilRig's DNS Tunneling

69,973 people reacted | 9 | 37 min. read | SHARE

By Robert Falcone
April 16, 2019 at 9:00 AM
Category: Unit 42
Tags: ALMA Communicator, BONDUPDATER, dns tunneling, Helminth, ISMAgent, OilRig, QUADAGENT

This post is also available in: 日本語 (Japanese)

On March 15, Unit 42 published a blog providing an overview of DNS tunneling and how malware can use DNS queries and answers to act as a command and control channel. To supplement this blog, we have decided to describe a collection of tools that rely on DNS tunneling used by an adversary known as OilRig.

Simulação do Ataque

Simulação do ataque

Tempest
[ACADEMY]
Conference

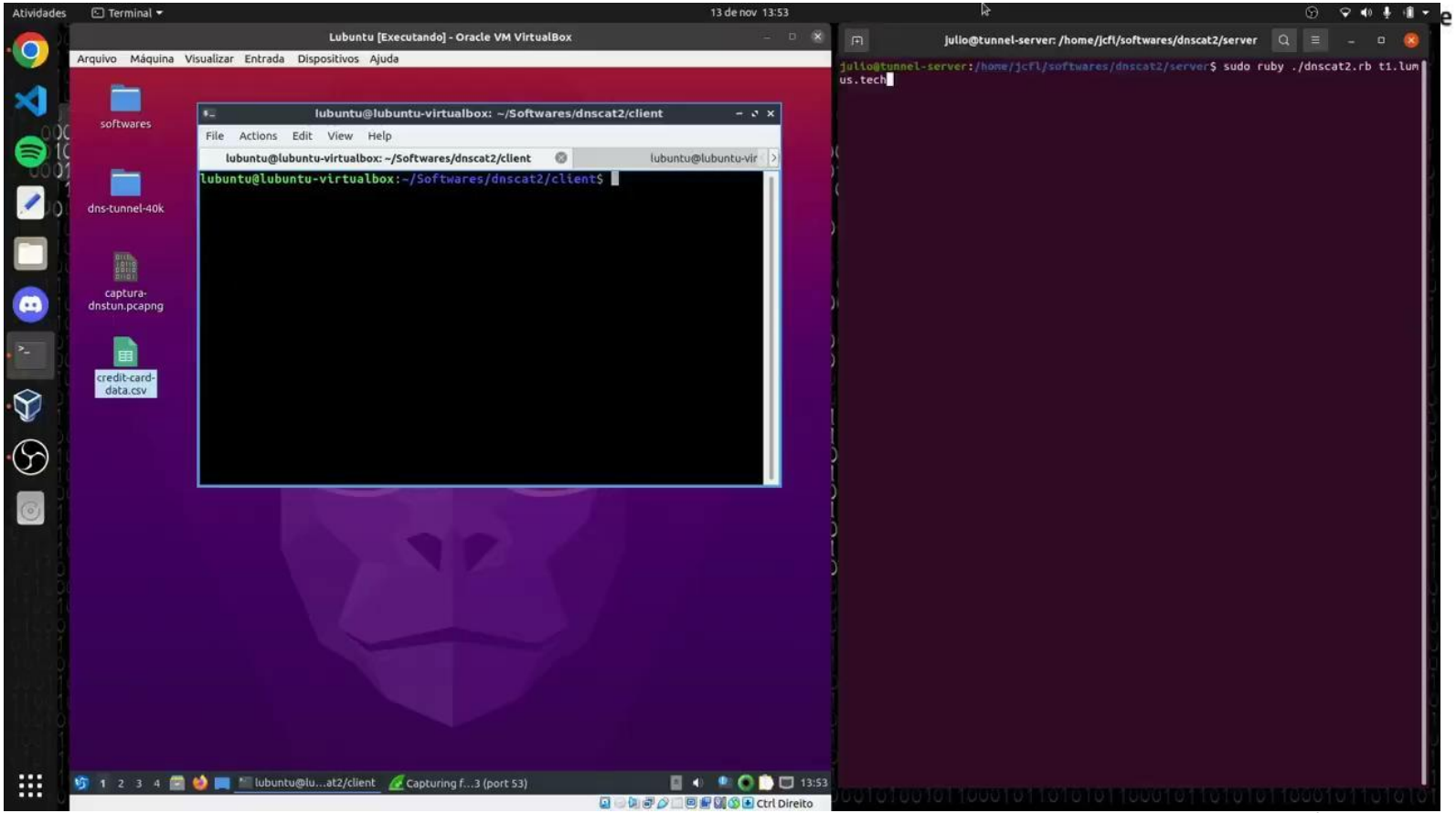
1. Configuração do servidor DNS do atacante, delegando a ele a zona t1.lumus.tech

2. Iniciação do servidor DNS do atacante com o dnscat2

3. Estabelecimento do túnel DNS pela vítima, com o dnscat2

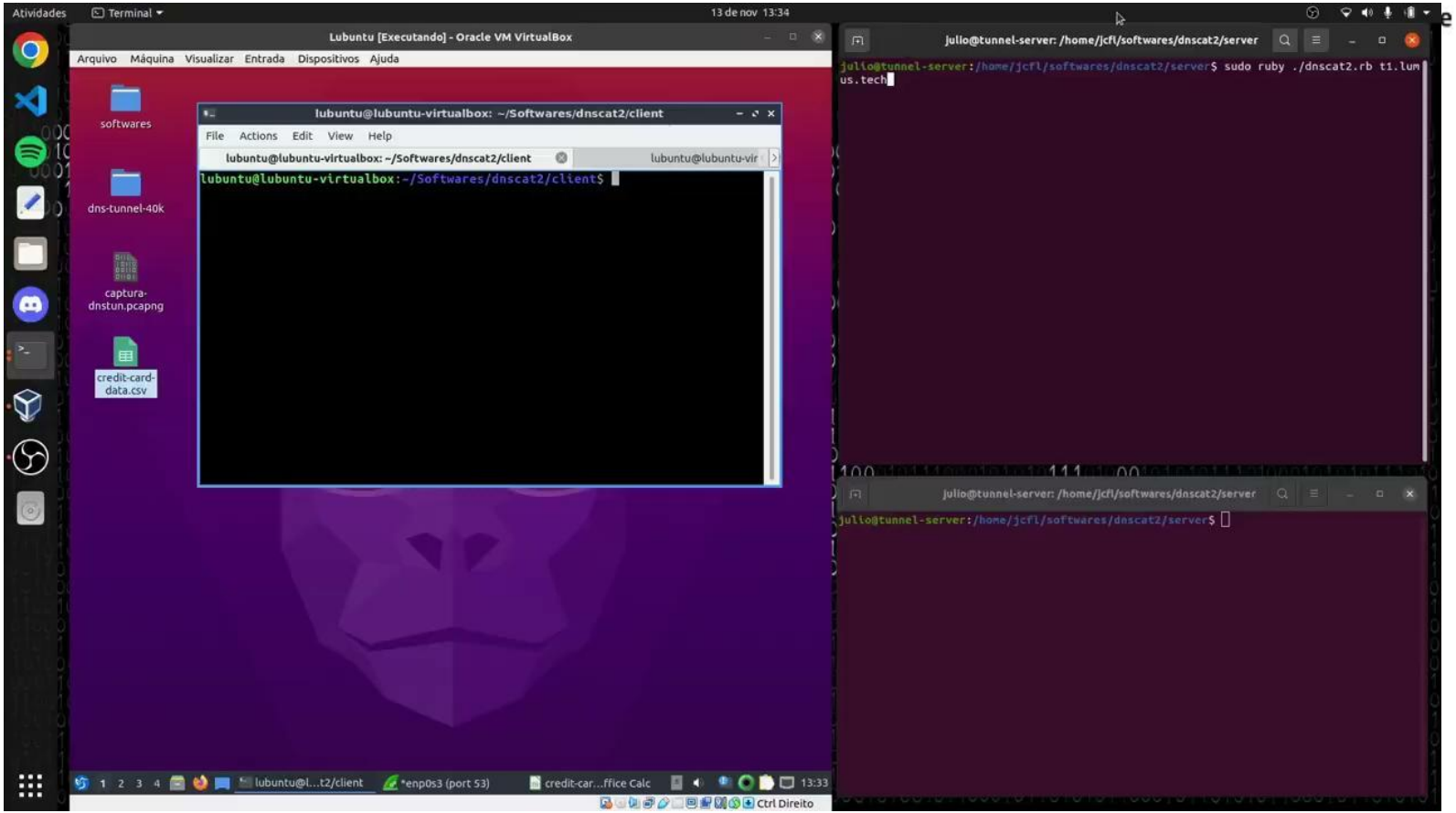
4. Execução de comandos do atacante através do túnel DNS firmado

Executando comandos

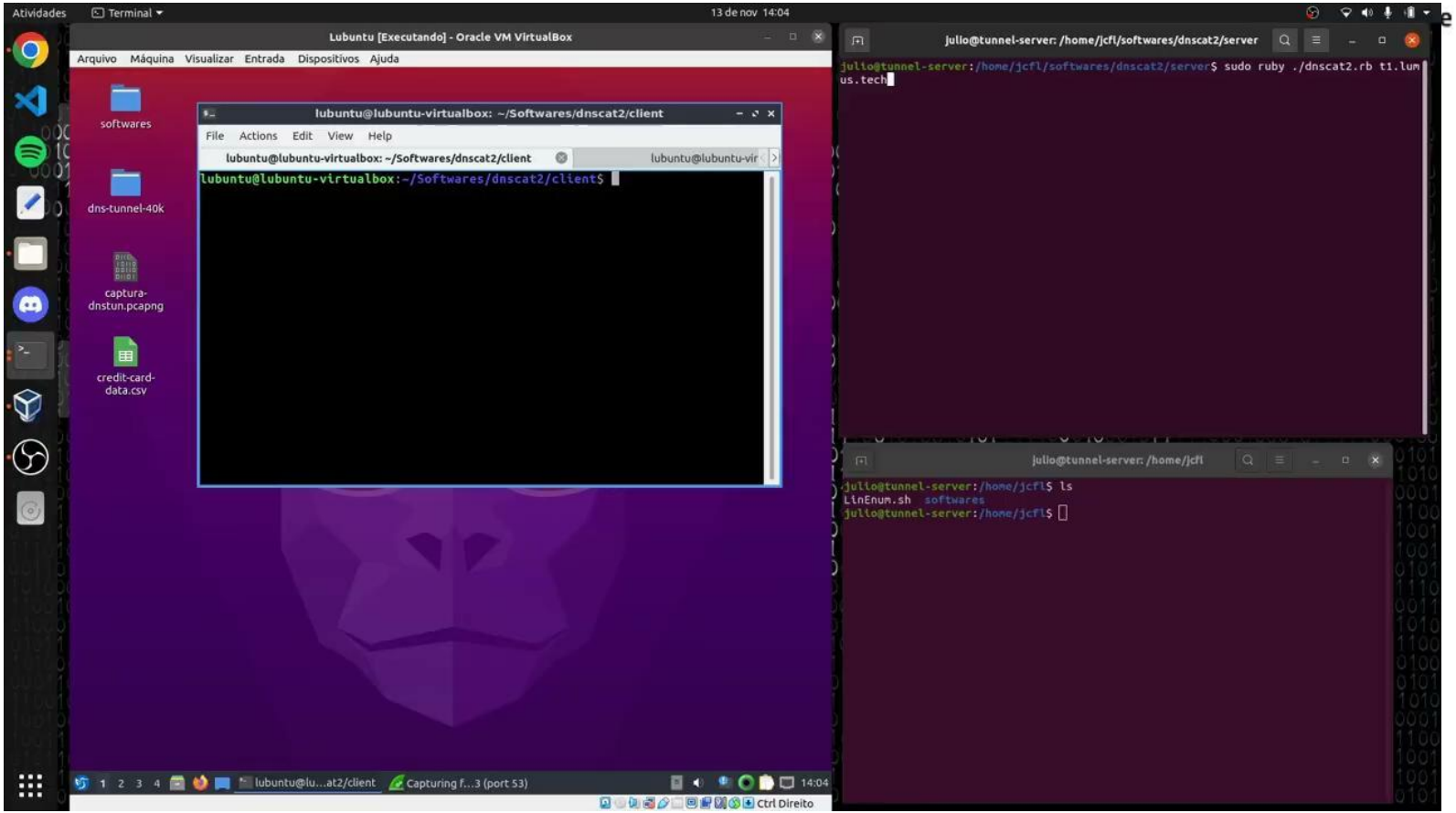


Exfiltrando arquivos

Tempest ACADEMY



Upload de script malicioso



Tempest

ACADEMY

Conference

Detecção

Detecção (dnscat2)

Tempest
ACADEMY
Conference

Para o **dnscat2**, poderíamos criar uma regra de detecção definindo thresholds de tamanho e formato da query (análise baseada em payload) ou thresholds considerando a volumetria do tráfego DNS (análise baseada em tráfego)

Análise baseada em payload

Tempest
[ACADEMY]
Conference

Queries maliciosas

FQDN
11f103285af295b8b025130000ad0a04269703fc6ee9898939f580db3750.c28f4765d2...
d50801285a64c3c49c08221a35aee97c04.t1.lumus.tech
d18801c4dce24f03071d98005b1e738b34dc596fb211cdc31b57cadd41b2.8e147045de...
a54001c4dc0bc387c3bdfb005ce5d9e685cda8a172eefa90eb49777cbc52.dfc469d834...
e08401c4dc31104f3af8e5005d86308fb7020165e7092bb4fb1ee9936cf7.db635ca7ad...
3fcb01c4dcd7c9ecf532c7004739944716a818f82bc843b98f5617b7e045.2321d11e7c...
d05b03c4dc00000000d7199875ad52ecb948797748514109a02400517c9d.f9ca94073d...
758f01285a730b14ef21761c48ec8373d1.t1.lumus.tech
11f103285af295b8b025130000ad0a04269703fc6ee9898939f580db3750.c28f4765d2...
11f103285af295b8b025130000ad0a04269703fc6ee9898939f580db3750.c28f4765d2...
a32b03c4dc191108726d1c0000bfa266888a9901a591ae9fd850832e09d9.860dba14ff...
0c5a01c4dc18e783942e820046097d034a9317cd0fd1b661614ad5644372.39f38ebe9f...

Análise baseada em payload

Tempest
[ACADEMY]
Conference

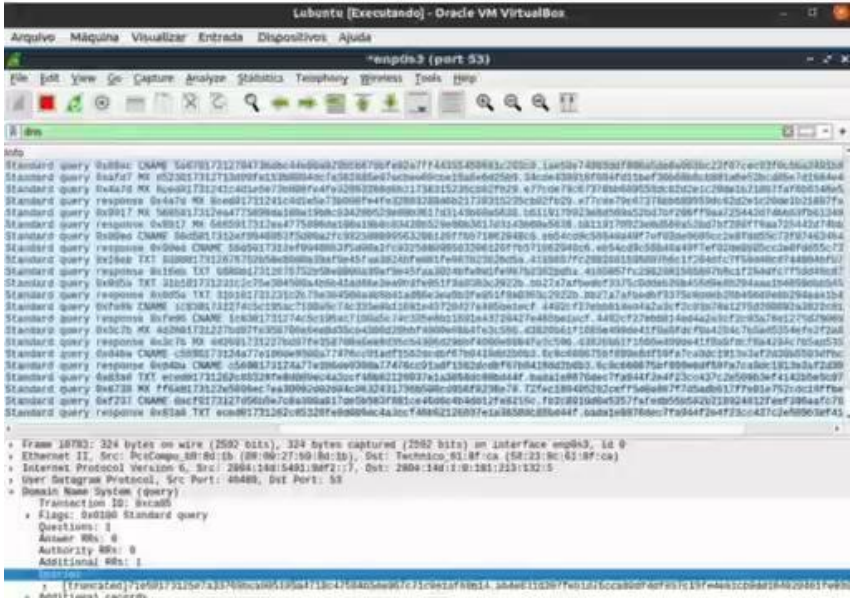
Queries maliciosas

FQDN
11f103285af295b8b025130000ad0a04269703fc6ee9898939f580db3750.c28f4765d2...
d50801285a64c3c49c08221a35aee97c04.t1.lumus.tech
d18801c4dce24f03071d98005b1e738b34dc596fb211cdc31b57cadd41b2.8e147045de...
a54001c4dc0bc387c3bdfb005ce5d9e685cda8a172eefa90eb49777cbc52.dfc469d834...
e08401c4dc31104f3af8e5005d86308fb7020165e7092bb4fb1ee9936cf7.db635ca7ad...
3fcb01c4dcd7c9ecf532c7004739944716a818f82bc843b98f5617b7e045.2321d11e7c...
d05b03c4dc00000000d7199875ad52ecb948797748514109a02400517c9d.f9ca94073d...
758f01285a730b14ef21761c48ec8373d1.t1.lumus.tech
11f103285af295b8b025130000ad0a04269703fc6ee9898939f580db3750.c28f4765d2...
11f103285af295b8b025130000ad0a04269703fc6ee9898939f580db3750.c28f4765d2...
a32b03c4dc191108726d1c0000bfa266888a9901a591ae9fd850832e09d9.860dba14ff...
0c5a01c4dc18e783942e820046097d034a9317cd0fd1b661614ad5644372.39f38ebe9f...

Queries legítimas

FQDN
212.156.118.6.00-atakoy-xrs-t2-1.00-ebgp-atakoy1-k.statik.turktelekom.c...
mac-d8-b1-22-95-27-00.ipv4-080-081-202-040.pas-25638.500mega.de-cix.muc...
telefonica-peru-hu0-0-0-19-100-grtlurem2.net.telefonicaglobalsolutions.com
asn-cxa-all-cci-22773-rdc-as22773.100gigabitethernet10-1.core1.lax1.he.net
81.212.223.45.21-diyarbakir-t2-2.72-batman-t3-1.statik.turktelekom.com.tr
81.212.244.249.78-safranbolu-t3-1.78-kayabasi-h4-1.statik.turktelekom.c...
mitchell-seaforth-cable-tv-ltd.10gigabitethernet1-1-50.switch2.tor2.he.net
telefonica-peru-grtmratw1-0-0-1-2-101.net.telefonicaglobalsolutions.com
pt-parsaoran-global-datatrans.10gigabitethernet1-1-36.switch2.sin1.he.net
mccluskey-chevrolet-geo-guaranteedcarcredit-66-42-142-76.static.fuse.net
45-248-194-67.STATIC.Skyline\x5fInfonet\x5fPrivate\x5fLimited.skylinein...
81.212.24.133.01-karsiyaka-t3-1.01-pttevleri-t4-1.statik.turktelekom.co...

Análise baseada em tráfego



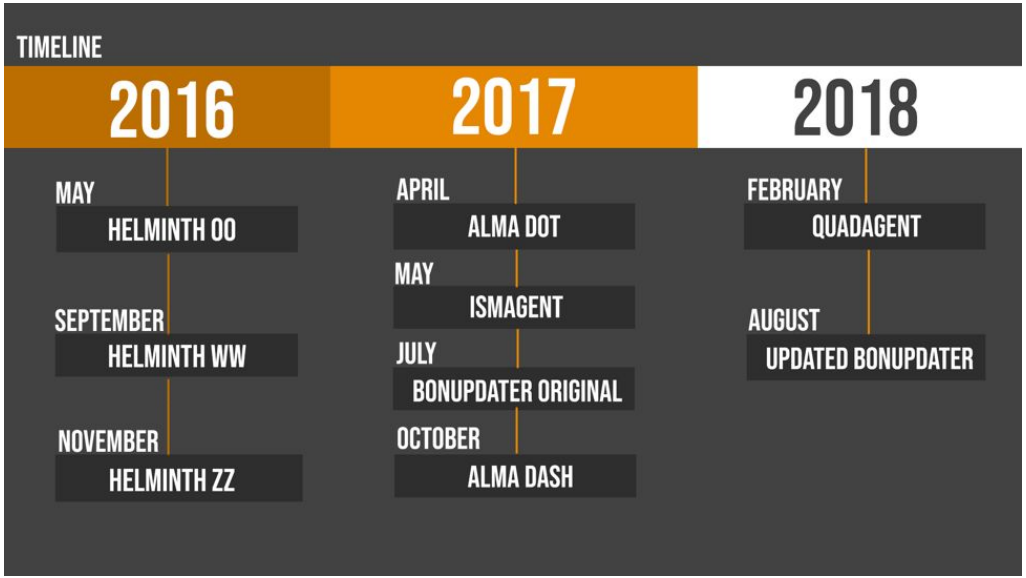
Mas e para detectar túneis DNS de outras ferramentas?

Tempest
ACADEMY
Conference

Mas e para detectar túneis DNS de outras ferramentas?

Tempest [ACADEMY] Conference

Caso relevante: Evolução das ferramentas usadas para tunelamento DNS do grupo OilRig



Tool	DNS Type	DNS Query method	Example C2 domain
Helminth	A	[System.Net.DNS]::GetHostByName	go0gie[.]com
ISMAgent	AAAA	DnsQuery_A	ntpupdateserver[.]com
ALMACommunicator	A	DnsQuery_W	prosalar[.]com
BONDUPDATER	A, TXT	[System.Net.Dns]::GetHostAddresses, System.Net.Sockets.UdpClient	poison-frog[.]club, withyourface[.]com
QUADAGENT	AAAA	nslookup.exe, Resolve-DnsName	acrobatverify[.]com

Table 2. DNS type and query method used by OilRig's tools using DNS tunneling for C2

Tempest
[ACADEMY]
Conference

Porém, como poderíamos criar uma detecção do caráter anômalo gerado pelo ataque de tunelamento DNS?

Mecanismos de defesa

Tempest
[ACADEMY]
Conference

Potenciais mecanismos de defesa e limitações

- Firewalls
- IDS baseados em assinaturas
- IDS supervisionados
- IDS não supervisionados

Mecanismos de defesa

Tempest
[ACADEMY]
Conference

Potenciais mecanismos de defesa e limitações

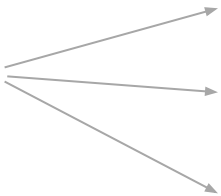
- Firewalls
- IDS baseados em assinaturas
- IDS supervisionados
- **IDS não supervisionados**

Mecanismos de defesa

Tempest
[ACADEMY]
Conference

Potenciais mecanismos de defesa e limitações

- Firewalls
- IDS baseados em assinaturas
- IDS supervisionados
- **IDS não supervisionados**



- Maior capacidade de generalização
- Possibilidade de detectar *zero-days*
- Menos custos de obtenção de dados rotulados

Vantagens de usar *machine learning*

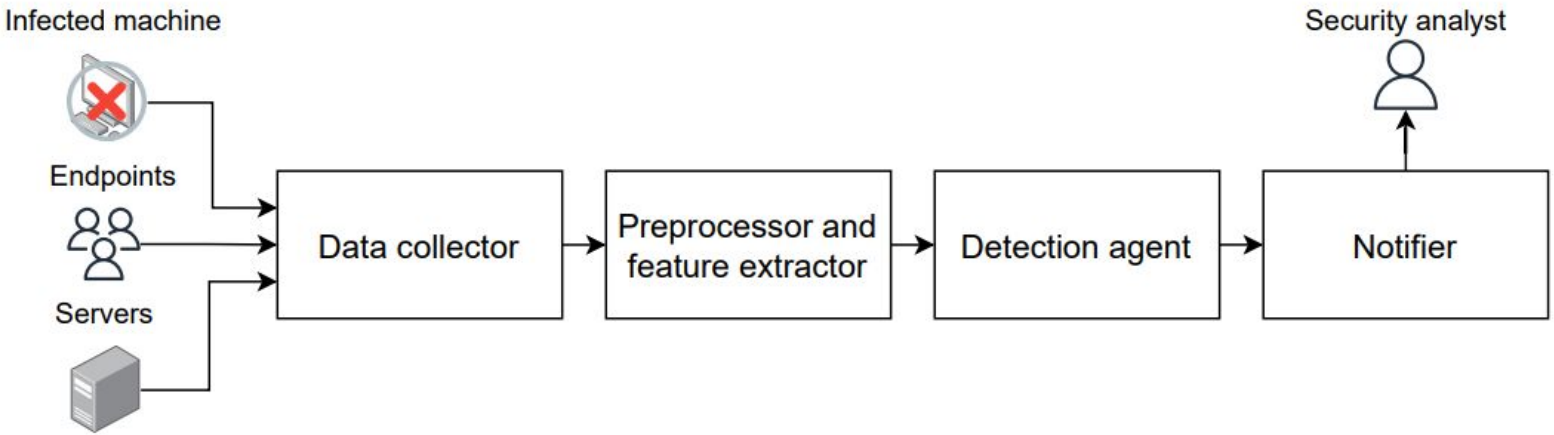
- Maior generalização da detecção
- Detecção baseada em desvio do comportamento normal
- Possibilidade de detectar *zero-days*
- Adaptabilidade a diferentes ambientes através do retreinamento

Trabalhos relacionados

	Contribuições	Limitações	Avaliado em tráfego DNS de redes empresariais reais?
[Lambion et al. 2020]	Modelo CNN + Random Forest	Requer dados maliciosos rotulados	✓
[Nguyen et al. 2020]	Modelo DBSCAN	Avaliação foi realizada em um dataset sem a documentação de como foram realizadas atividades de tunelamento DNS	x
[Campbell and Zincir-Heywood 2020]	Modelo SOM	Nenhum mecanismo é sugerido para mitigar os falsos positivos	x

Sistema de Detecção Proposto

Visão geral



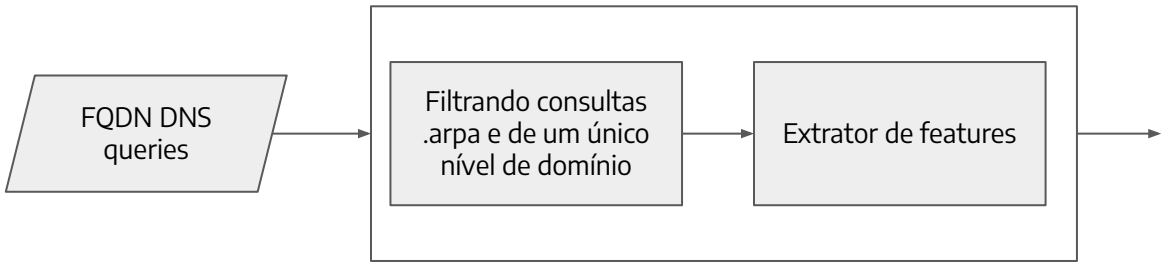
Sistema de Detecção Proposto

Pré processamento e extração de features

Dados de entrada

Pré processamento e extração de features

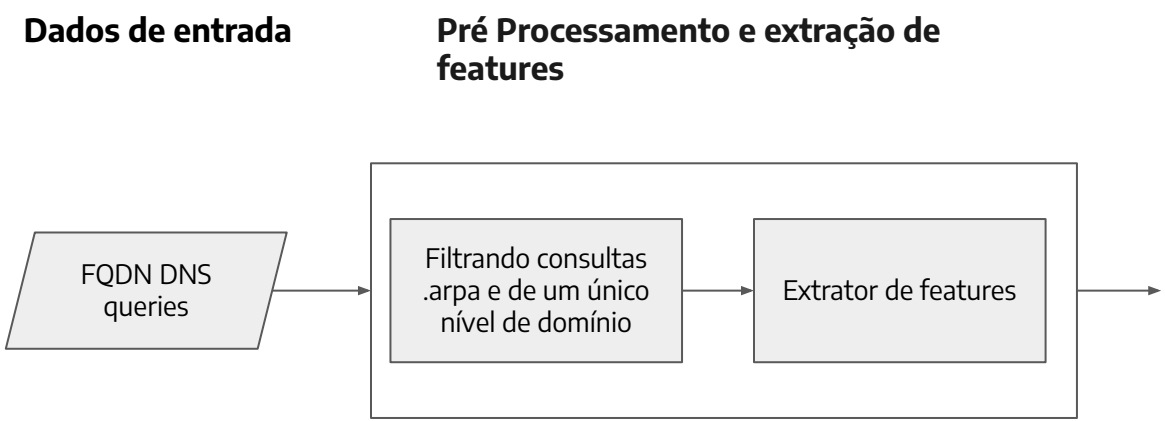
Features extraídas



Entropia
Número de subdomínios
Tamanho
Tamanho máximo de subdomínio
Tamanho da maior seq. de dígitos
Tamanho da maior seq. de letras
Taxas de dígitos, caracteres especiais e vogais
Contagem de dígitos, caracteres especiais e vogais
Reputation value
Reputation value por *n*-gram

Sistema de Detecção Proposto

Pré processamento e extração de features



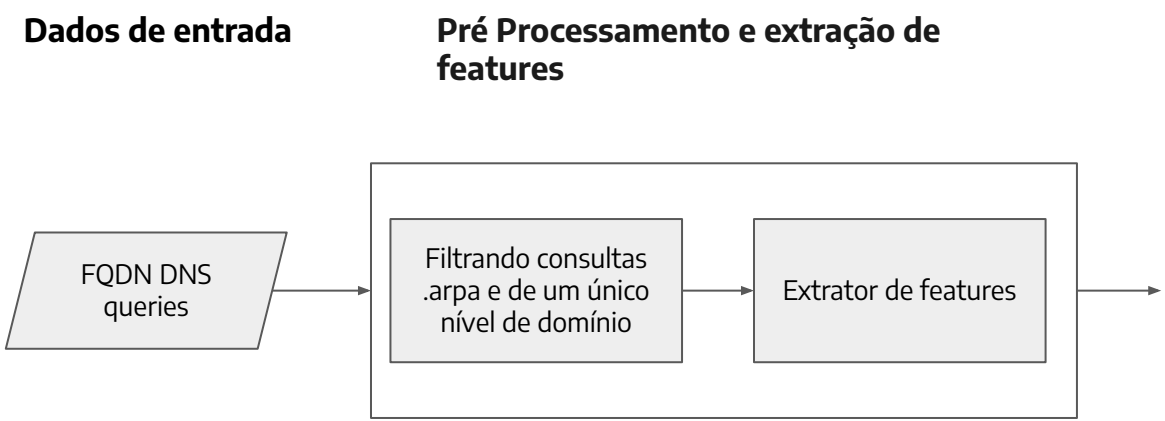
Features extraídas

(Análise baseada em payload)

Entropia
Número de subdomínios
Tamanho
Tamanho máximo de subdomínio
Tamanho da maior seq. de dígitos
Tamanho da maior seq. de letras
Taxas de dígitos, caracteres especiais e vogais
Contagem de dígitos, caracteres especiais e vogais
Reputation value
Reputation value por *n*-gram

Sistema de Detecção Proposto

Pré processamento e extração de features



Features extraídas

(Análise baseada em payload)

Entropia
Número de subdomínios
Tamanho
Tamanho máximo de subdomínio
Tamanho da maior seq. de dígitos
Tamanho da maior seq. de letras
Taxas de dígitos, caracteres especiais e vogais
Contagem de dígitos, caracteres especiais e vogais
Reputation value
Reputation value por *n*-gram

Sistema de Detecção Proposto

Pré processamento e extração de features

Tempest ACADEMY Conference

Reputation Value

	FQDNs mais consultados
1	google.com
2	tempest.com.br
3	github.com
4	facebook.com.br
...	...
100.000	theguardian.com

N-grama	Contagem
goo	230
oo	600
temp	250
guard	90
...	...
0af	0
qkzw	0

$$W_{N-Gram_{(i)}} = log_2(N \times C_{N-Gram_{(i)}})$$

$$\text{Reputation-value} = \sum_{i=1}^{t} W_{N-Gram_{(i)}}$$

Sistema de Detecção Proposto

Pré processamento e extração de features

Tempest ACADEMY Conference

Reputation Value

	FQDNs mais consultados
1	google.com
2	tempest.com.br
3	github.com
4	facebook.com.br
...	...
100.000	theguardian.com

N-grama	Contagem
goo	230
oo	600
temp	250
guard	90
...	...
0af	0
qkzw	0

$$W_{N-Gram_{(i)}} = log_2(N \times C_{N-Gram_{(i)}})$$

$$\text{Reputation-value} = \sum_{i=1}^{t} W_{N-Gram_{(i)}}$$

login.yahoo.com

7afb0af2183d3698ce.t1.lumus.tech

Sistema de Detecção Proposto

Pré processamento e extração de features

Tempest ACADEMY Conference

Reputation Value

	FQDNs mais consultados
1	google.com
2	tempest.com.br
3	github.com
4	facebook.com.br
...	...
100.000	theguardian.com

N-grama	Contagem
goo	230
oo	600
temp	250
guard	90
...	...
0af	0
qkzw	0

$$W_{N-Gram_{(i)}} = log_2(N \times C_{N-Gram_{(i)}})$$

$$\text{Reputation-value} = \sum_{i=1}^{t} W_{N-Gram_{(i)}}$$

login.yahoo.com — Maior *reputation value*

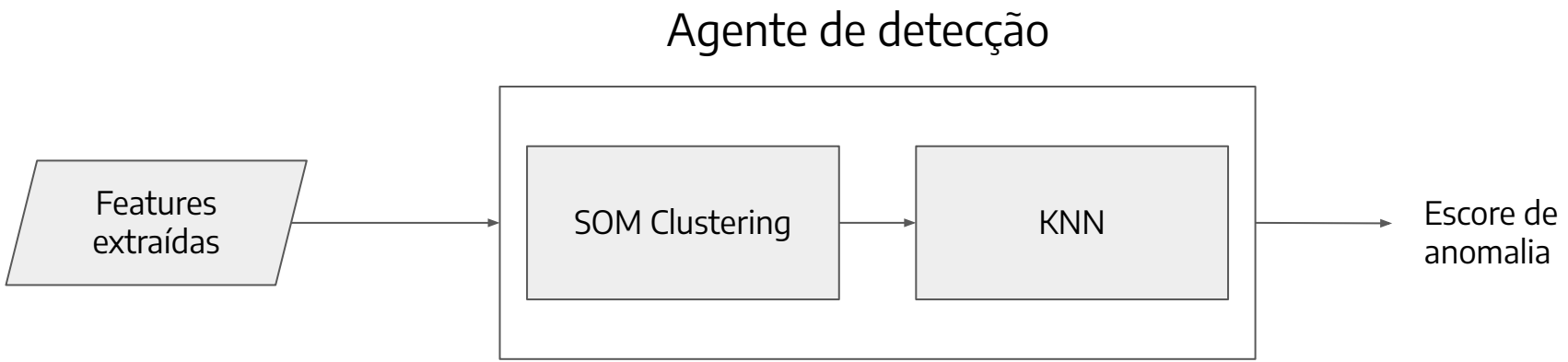
7afb0af2183d3698ce.t1.lumus.tech — Menor *reputation value*

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Detecção de anomalias usando Self-Organizing Maps (SOM) e K-Nearest Neighbors (KNN)

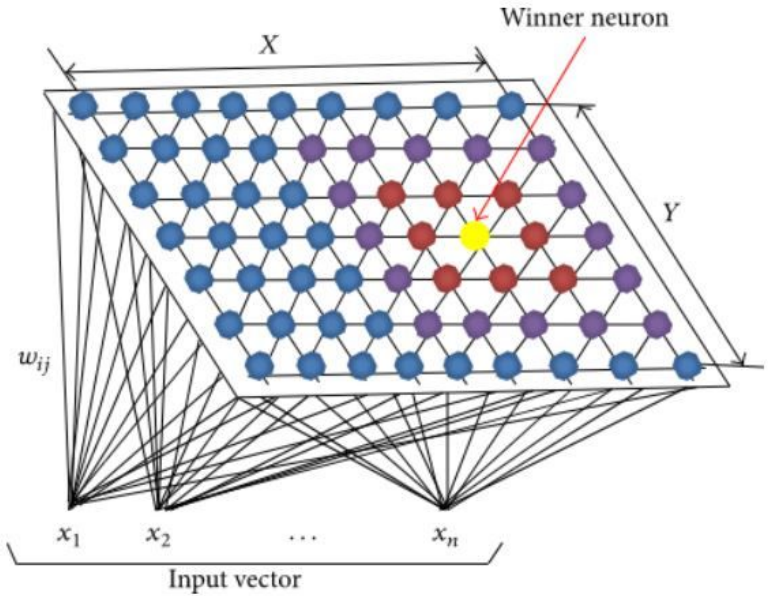


Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM)

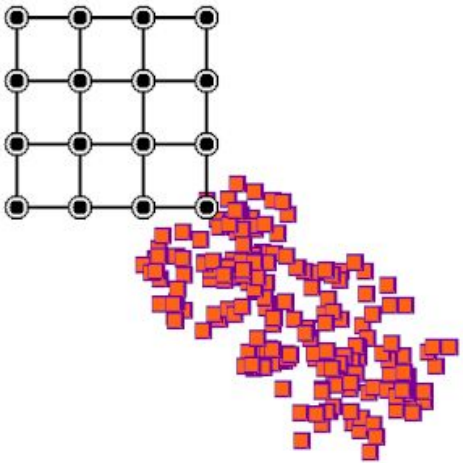


Sistema de Detecção Proposto

Agente de detecção

Tempest
ACADEMY
Conference

Self-Organizing Maps (SOM)

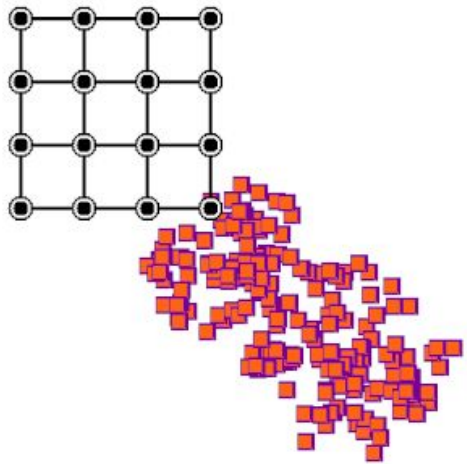


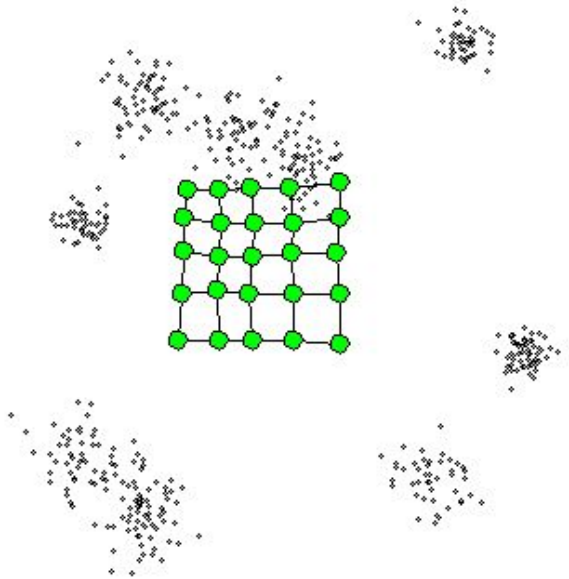
Sistema de Detecção Proposto

Agente de detecção

Tempest
ACADEMY
Conference

Self-Organizing Maps (SOM)



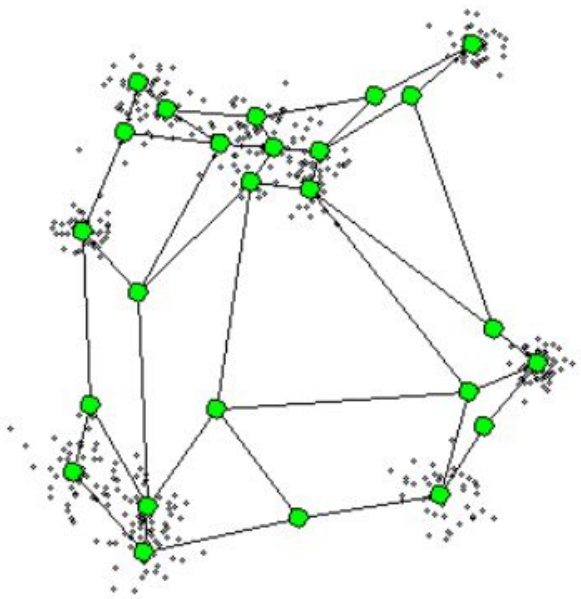


Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



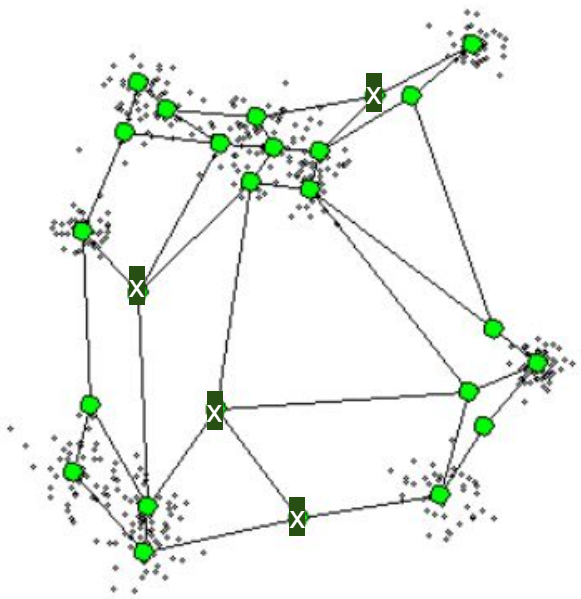
E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



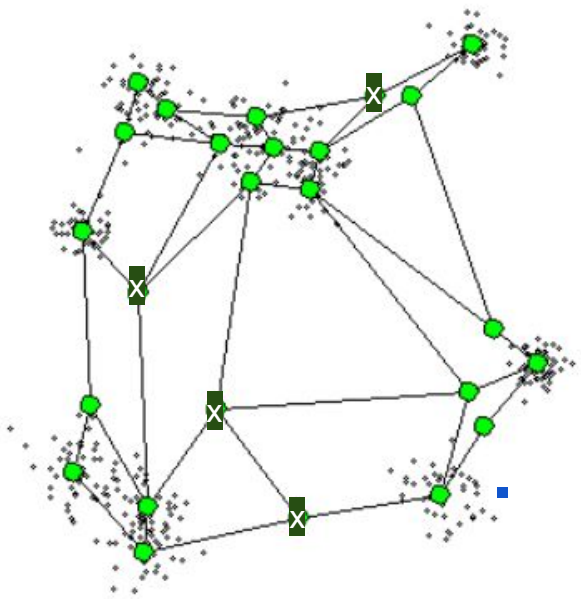
E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



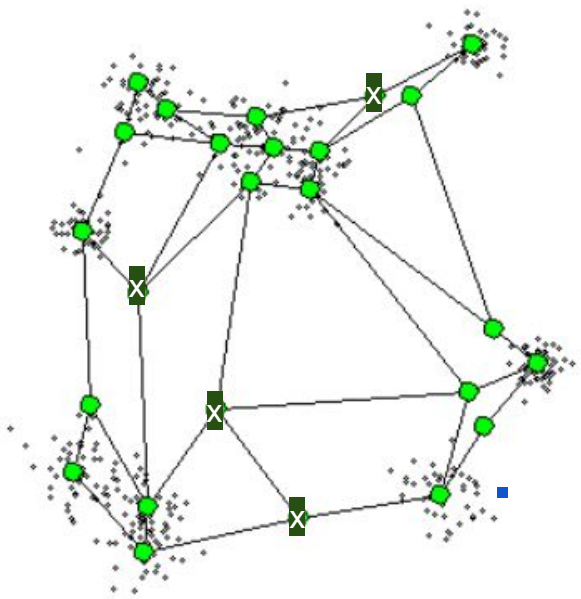
E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

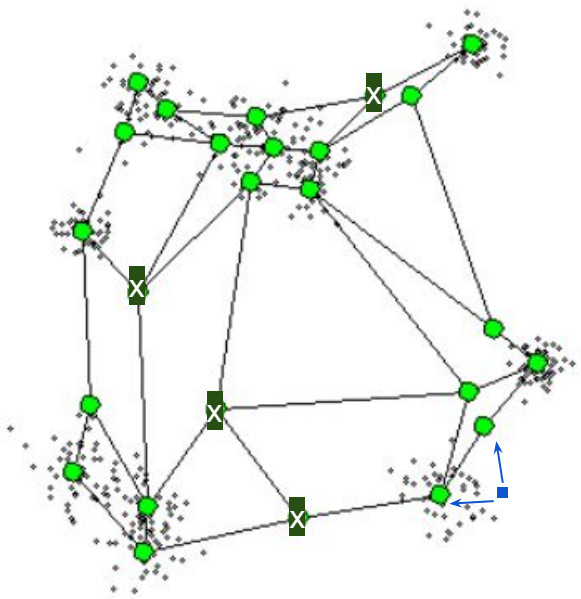
1. Calcula-se a distância média para os k neurônios mais próximos. (Escore de anomalia)

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

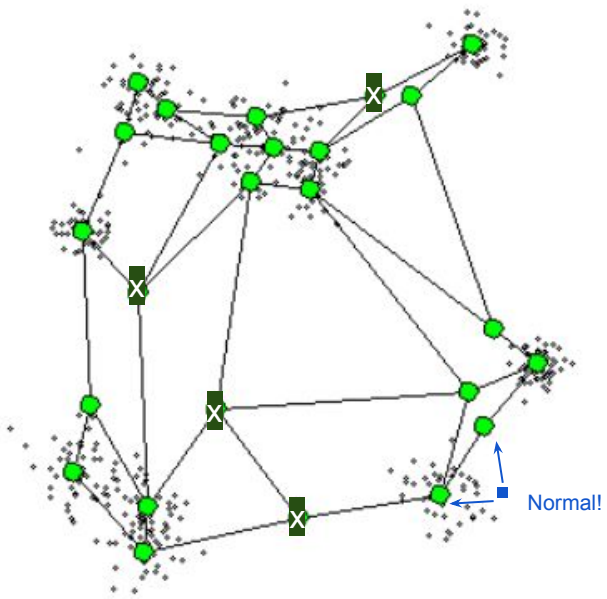
1. Calcula-se a distância média para os k neurônios mais próximos. (Escore de anomalia)
2. Caso o escore de anomalia seja maior do que um determinado limiar, acuse-o como anomalia

Sistema de Detecção Proposto

Agente de detecção

Tempest ACADEMY Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

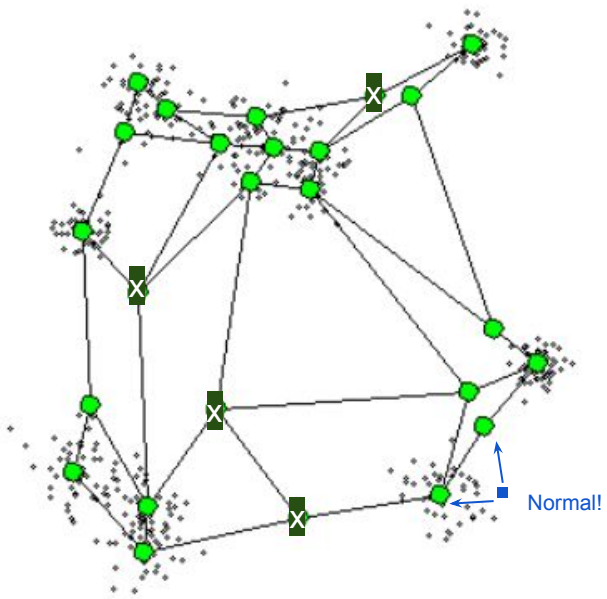
1. Calcula-se a distância média para os k neurônios mais próximos. (Escore de anomalia)
2. Caso o escore de anomalia seja maior do que um determinado limiar, acuse-o como anomalia

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

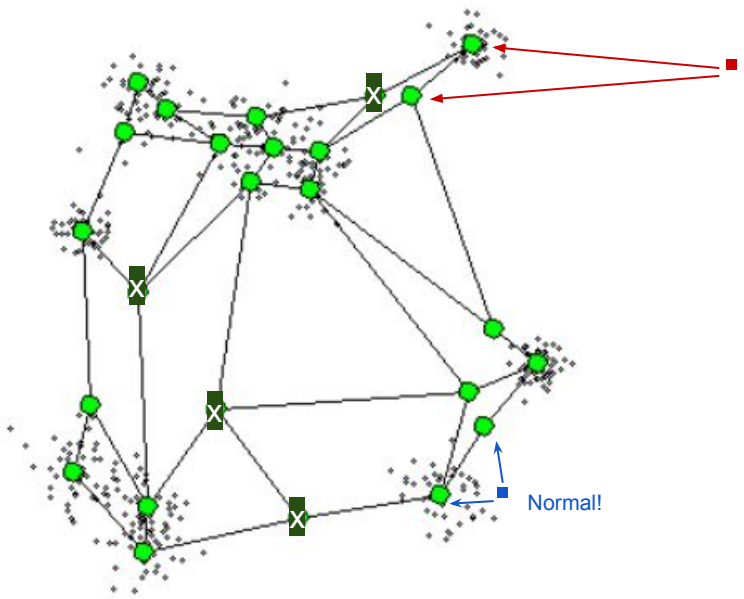
1. Calcula-se a distância média para os k neurônios mais próximos. (Escore de anomalia)
2. Caso o escore de anomalia seja maior do que um determinado limiar, acuse-o como anomalia

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

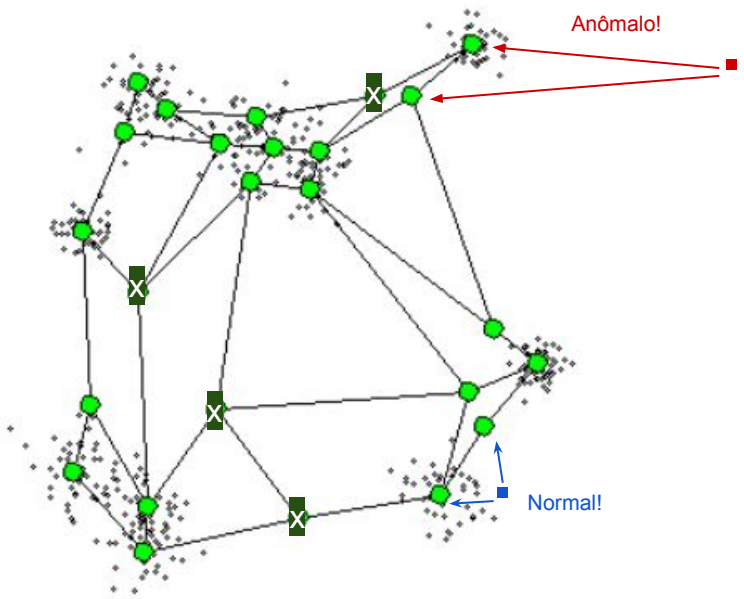
1. Calcula-se a distância média para os k neurônios mais próximos. (Escore de anomalia)
2. Caso o escore de anomalia seja maior do que um determinado limiar, acuse-o como anomalia

Sistema de Detecção Proposto

Agente de detecção

Tempest
[ACADEMY]
Conference

Self-Organizing Maps (SOM) + K-Nearest Neighbors



E depois de ter treinado a rede SOM, como a detecção de anomalias ocorre?

1. Calcula-se a distância média para os k neurônios mais próximos. (Escore de anomalia)
2. Caso o escore de anomalia seja maior do que um determinado limiar, acuse-o como anomalia

Tempest

ACADEMY

Conference

Resultados

Experimentos

Dataset público

Dados benginos: CAIDA UCSD IPv4 Routed /24 DNS Names Dataset
Dados maliciosos: Consultas DNS das ferramentas dns2tcp, dnscapy, iodine, e tuns

Dataset privado (Tráfego DNS real da Tempest Security Intelligence)

Dados benignos: Consultas DNS coletadas de março e abril de 2023
Dados maliciosos: Consultas DNS das ferramentas iodine e DNSExfiltrator

Experimentos

Tempest
[ACADEMY]
Conference

Resumo dos datasets

Number of domains	Public dataset			Real DNS traffic dataset		
	Training set	Validation set	Testing set	Training set	Validation set	Testing set
Number of benign domains	200,000	8,000	8,000	2,345,219	128,820,871	210,136,888
Number of malicious domains	0	8,000	8,000	0	3,385	247,501

Métricas

Tempest
ACADEMY
Conference

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$F1\ Score = 2 \times \frac{recall \times precision}{recall + precision}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

Resultados

Dados públicos

Tempest
[ACADEMY]
Conference

Resultados comparando nossa proposta contra [Campbell and Zincir-Heywood 2020]

Model	Accuracy	F1-Score	Precision	Recall	TPR	FPR	TNR	FNR
Proposed architecture	**94.38%**	**0.9460**	**0.9107**	**0.9841**	**98.41%**	**9.65%**	**90.35%**	**1.59%**
[Campbell and Zincir-Heywood 2020]	92.24%	0.9265	0.8800	0.9783	97.83%	13.34%	86.66%	2.16%

Resultados

Dados privados

Tempest
[ACADEMY]
Conference

Resultados da nossa proposta para os dados privados

Tool	Exfiltrated file size	Accuracy	F1 Score	Precision	Recall	TPR	FPR	TNR	FNR
Iodine	1.6 MB	99.99%	0.8524	0.9982	0.7438	74.38%	0.00006%	99.99994%	25.62%
	200 KB	99.99%	0.8295	0.9899	0.7138	71.38%	0.00006%	99.99994%	28.62%
	32 KB	99.99%	0.7623	0.9745	0.6261	62.61%	0.00006%	99.99994%	37.39%
DNSExfiltrator	1.6 MB	99.99%	0.9993	0.9989	0.9998	99.98%	0.00006%	99.99994%	0.02%
	200 KB	99.99%	0.9959	0.9930	0.9989	99.89%	0.00006%	99.99994%	0.11%
	32 KB	99.99%	0.9904	0.9850	0.9959	99.59%	0.00006%	99.99994%	0.41%

Por outro lado, [Campbell and Zincir-Heywood 2020] mostrou uma taxa de falsos positivos de 1,13%, resultando em mais de 2M falsos positivos

Conclusão

- IDS não supervisionado com funcionamento em tempo real para detecção de tunelamento DNS
- Resultados interessantes em uma rede empresarial real
- F1-Score de 0.946 para dados públicos
- Baixa taxa de falsos positivos para tráfego DNS real

Obrigado!

Linkedin

Tempest
ACADEMY
Conference