FACULDADE DE TECNOLOGIA E CIÊNCIAS SOCIAIS APLICADAS – FATECS
ENGENHARIA DE COMPUTAÇÃO

CLÉSIO ALVES DE ARAÚJO
21500543

CYBER DECEPTION: uma análise de eficiência na segurança de redes

BRASÍLIA
2020

CLÉSIO ALVES DE ARAÚJO

CYBER DECEPTION: uma análise de eficiência na segurança de redes

Trabalho de Conclusão de Curso (TCC) apresentado como um dos requisitos para a conclusão do curso de Engenharia de Computação do UniCEUB– Centro Universitário de Brasília

Orientador (a)**: Prof. MsC Francisco Javier de Obaldia Diaz**

BRASÍLIA
2020

CLÉSIO ALVES DE ARAÚJO

CYBER DECEPTION: uma análise de enficiência

Trabalho de Conclusão de Curso (TCC) apresentado como um dos requisitos para a conclusão do curso de Engenharia Computação do UniCEUB – Centro Universitário de Brasília

Orientador (a)**: Francisco Javier de Obaldia Diaz**

Brasília, 2020.

BANCA EXAMINADORA

_____

Nome e titulação.
Orientador (a)

_____

Nome e titulação.
Examinador (a)

_____

Nome e titulação.
Examinador (a)

CYBER DECEPTION: uma análise de eficiência na segurança de redes
CYBER DECEPTION: an efficiency analysis in network security

Clésio Alves de Araújo[1], Francisco Javier de Obaldia Diaz[2], Ivandro da Silva Ribeiro[3], William Roberto Malvezzi[4]

RESUMO

Com a sofisticação da tecnologia, de ferramentas e com avanço de genialidade, os ciber criminosos estão cada vez mais perigosos e causando cada vez mais prejuízo as empresas e corporações pelo mundo a fora. Mas os profissionais da área de Segurança Cibernética também estão mais sofisticados, evoluindo constantemente e criando novos métodos de tornar o campo cibernético um lugar melhor para todos nós. Por esse motivo, um conceito que vem ganhando bastante fama de pouco tempo para cá é a estratégia de *Cyber Deception*. Neste trabalho, é estudado dois aspectos das técnicas, tecnologias e ferramentas de *Cyber Deception*: quais os seus impactos no comportamento dos agentes maliciosos e como podem diminuir o impacto final de um ataque cibernético. A metodologia empregada foi a de criar, uma espécie de *Capture the Flag* customizado, com quatro objetivos, e com ferramentas de *Cyber Deception* espalhadas pelo ambiente de testes. Os resultados mostraram que nenhum dos dois ataques realizados obtiveram total furtividade e também não conseguiram atingir os quatro objetivos. Também foi observado uma certa resistência dos atacantes à intimidação mas mesmo assim ficaram nítidos alguns desvios de comportamento de parte dos atacantes.

Palavras-chave: Segurança. Cibernética. Redes. Negação. Armadilha. Visibilidade.

Abstract:

With the sophistication of technology, tools and the advancement of genius, cyber criminals are increasingly dangerous and causing more damage to companies and corporations around the world. But cybersecurity professionals are also more sophisticated, constantly evolving and creating new methods of making cyberspace a better place for all of us. For this reason, a concept that has been gaining quite a fame for a short time here is the Cyber Deception strategy. In this work, two aspects of Cyber Deception techniques, technologies and tools are studied: what are their impacts on the behavior of malicious agents and how can they reduce the final impact of a cyber attack. The methodology used was to create, a kind of customized Capture the Flag, with four objectives, and with Cyber Deception tools spread throughout the testing environment. The results showed that none of the two attacks carried out achieved total stealth and also failed to achieve the four objectives. Attackers' resistance to intimidation was also observed, but some behavioral deviations on the part of the attackers were clear.

keywords: Security. Cyber. Deception. Network. Denial. Trap. Visibility.

---

[1] UniCEUB, aluno. [2] UniCEUB, orientador. [3] UniCEUB, primeiro examinador. [4] UniCEUB, segundo examinador.

1 INTRODUÇÃO

A pandemia da COVID-19 dificultou a vida dos profissionais de uma área que já é um tanto quanto difícil lidar: a área de Segurança Cibernética.

Atualmente, uma das maneiras que os profissionais de segurança cibernética tem a sua mercê e que vem ganhando cada vez mais visibilidade e investimento são as tecnologias e técnicas baseadas no paradigma de resposta – diferente do paradigma de prevenção que já está bastante consolidado no mercado.

Vale ressaltar que o objetivo do paradigma de resposta não é substituir o de prevenção e sim auxiliar servindo como uma segunda camada de segurança do ambiente para lidar com ataques em que o agente malicioso já conseguiu acesso interno de alguma maneira. Aqui, entram as técnicas e tecnologias baseadas nas estratégias já vistas em guerras a várias gerações e que, em uma tradução literal, chamamos de enganação ou, originalmente, *deception*.

Cyber Deception é o nome utilizado para designar técnicas e tecnologias baseadas no jogo da enganação que são aplicadas no campo da segurança cibernética com o objetivo de ludibriar e persuadir os agentes maliciosos para dar mais controle aos profissionais de segurança cibernética na hora de proteger as organizações de ataques cibernéticos.

Esta é uma pesquisa que tem como foco averiguar como tecnologias e técnicas de *Cyber Deception* simples e de código livre lidam com um ataque em tempo real, como isso pode mudar o impacto final do ataque e também como isso pode afetar até mesmo o próprio atacante.

O ataque será feito em um ambiente de testes com vários indicadores de sucesso relacionados ao atacante com o intuito de rastrear suas atividades e vários indicadores de gatilho para alertar sobre as atividades do atacante.

2 REVISÃO BIBLIOGRÁFICA

Com o avanço da inteligência ligada ao mundo do cibercrime, as técnicas e os grupos de ciber criminosos vem aumentando cada vez mais, sem falar nas vulnerabilidades que são geradas no desenvolvimento de sistemas e afins.

Isso, ligado aos desafios que a transformação digital trás para o ramo da segurança cibernética, faz com que os métodos de defesa tradicionais não sejam mais tão eficazes, o que abre diversas oportunidades para práticas, técnicas e procedimentos novos entrarem em ação, como por exemplo: *Cyber Deception.*

O uso de técnicas baseadas em *deception* na segurança cibernética moderna como um meio ativo de ciber defesa pós-invasão é um fenômeno emergente (BUSHBY, 2019).

Mas o que é *deception*? A definição mais bem vista é que são “Ações planejadas para enganar invasores fazendo com que tenham (ou não tenham) ações específicas que auxiliam na defesa da segurança cibernética” (YUILL, 2006).

Quais os ganhos que se tem utilizando destas técnicas? Técnicas baseadas em *deception* são válidas por quatro motivos: primeiro, estas técnicas podem desviar a atenção do adversário dos alvos e recursos principais dando aos agentes que empregaram as técnicas mais tempo e liberdade para agir; segundo, podem causar impressões no adversário fazendo com que ele aja de uma maneira que ajude os defensores em seu trabalho; terceiro, provém o elemento surpresa para os defensores; quarto, pode ajudar a manter os recursos atuais fora de perigo (HECKMAN et al., 2015).

E como trazer isso para o mundo da segurança cibernética? Pode-se trazer as técnicas de *deception* para o mundo da segurança da informação, por exemplo, através de adaptações feitas por Kristin E. Heckman et al[1] na matriz elaborada

[1] HECKMAN, Kristin E. Et al. Cyber Denial, Deception and Counter Deception. Springer International Publishing. Suíça, 2015.

originalmente para técnicas de *deception* (BENNETT; WALTZ, 2007), mas por motivos de clareza e simplicidade, convém dividir em quatro etapas, respectivamente: revelar fatos, esconder fatos, revelar ficções e esconder ficções.

De acordo com um estudo feito pela FireEye[1], em 2019, 41% dos comprometimentos investigados pelos experts da *Mandiant (FireEye)* constataram que o tempo de permanência mundial dos invasores é de 31 dias ou menos, quando os dados são analisados separadamente por região, as américas lideram o tempo de permanência em 60 dias. Obviamente o estudo também mostra que cada ano que se passa o tempo de permanência de um comprometimento tende a diminuir mas hoje ainda não temos um tempo de permanência desejável, ou seja, a resposta dos experts em segurança da informação ainda é muito lenta.

O motivo desse tempo extenso até as empresas detectarem um comprometimento sem interferência de fatores externos é a elaboração de ataques mais sofisticados, direcionados e menos repetitivos, nesses casos, a segurança baseada num paradigma de prevenção com firewalls, softwares de EDR e afins não é eficaz, sendo assim, a segurança baseada em resposta é a mais indicada.

A demanda por técnicas e processos de segurança baseada no paradigma de resposta crescerá progressivamente já que os ataques mais sofisticados, direcionados e menos repetitivos estão crescendo também (BASRKERVILLE; SPAGNOLETTI; KIM, 2014).

Há também indícios na literatura que provam que há um tipo de ataque cibernético que, atualmente, na maioria das vezes resulta em sucesso. Conceitos sobre os ataques mais sofisticados, direcionados e menos repetitivos citados anteriormente e os grupos que os praticam, designados como *APTs – Advanced Persistent Threat*, já foram definidos na literatura anteriormente, segundo os pesquisadores, esse tipo de agente malicioso apresenta um risco elevado as empresas e corporações do mundo inteiro por serem grupos maliciosos totalmente organizados, bem estruturados que aplicam ataques sofisticados e persistentes com dois objetivos principais: sabotagem ou roubo de informações, e, algumas vezes, ambos (AHMAD et al., 2019)

Por causa de sua natureza, *APTs* tem um índice de sucesso contra os mecanismos de segurança de perímetro muito elevado. O motivo disso se da pela metodologia e operação dos ataques, que exploraram vulnerabilidades *zero day* – classificadas com tal nome pois são vulnerabilidades recém descobertas e que não possuem nenhum tipo de correção por parte do desenvolvedor do sistema em questão, o que torna os mecanismos de defesa baseados em assinaturas ou inteligência artificial ineficazes (ZHISHENG, 2019).

Cyber Deception pode ser usado como ferramenta para recriar Ciberataques de Múltiplas Etapas em ambientes construídos em nuvem privada (KAHLHOFER; HÖLZL; BERGER, 2020).

É possível também utilizar *Cyber Deception* de maneira virtualizada para aplicar em ambientes como se fossem ativos de Internet das Coisas ou CPE sem ser necessário comprar outros ativos ou adquirir novas máquinas virtuais para se defender de ataques utilizando o *Framework Honware* (VETTERL; CLAYTON, 2019).

Outra utilidade também utilizada com tecnologias de *Cyber Deception* é a pesquisa aplicada a determinado serviço, protocolo ou software para poder se antecipar ao aumento de exploração ou rastrear determinada vulnerabilidade. Um exemplo disso é que os ataques a servidores de telefonia PBX, não só continuam acontecendo como estão aumentando cada vez mais, isto sugere talvez a criação de algumas *botnets* formadas somente de servidores de telefonia (MCINNES; ZALUSKA; WILLS, 2018).

Em junho deste ano, através de uma pesquisa empregando técnicas de *Cyber*

[1] FireEye. M-Trends 2020. p. 14. Disponível em: https://content.fireeye.com/m-trends/rpt-m-trends-2020

Deception emplementadas dentro do perímetro seguro, foi demonstrado que é possível distinguir entre atividades maliciosas robóticas e automáticas de atividades maliciosas realizadas com interações humanas observando como agentes maliciosos interagem com o ambiente tecnológico quando já estão dentro do perímetro seguro (CHACON et al., 2020).

O uso de *honeypots* e *honeytokens* para detectar ameaças internas também já foi estudado no passado (SPITZNER, 2003). Essas técnicas hoje fazem parte do que classificamos como *Cyber Deception*.

Sendo assim, esta pesquisa tem como foco a utilização de técnicas de *Cyber Deception* implementadas no ambiente interno, ou seja, por dentro do perímetro seguro, e como premissa, que o ponto de partida do agente malicioso não é conseguir atravessar a segurança de perímetro e sim o que fazer quando já tem acesso ao ambiente interno.

3 METODOLOGIA DO TRABALHO

Levando em consideração o que foi proposto no início, esta pesquisa de natureza aplicada tem como objetivo responder duas perguntas:

P1: Como *Cyber Deception* pode mudar o impacto final do ataque?

P2: Como *Cyber Deception* pode influênciar no comportamento do atacante?

Para isso, será proposto uma espécie de *CTF Game (Capture the Flag)* com técnicas de *Cyber Deception* em todo o ambiente.

O problema, abordado de forma qualitativa, será a falta de ferramentas, tecnologia e técnicas para diminuir, rastrear e até bloquear a ação de agentes maliciosos depois que estes já possuem acesso a rede interna de seu alvo.

Como todo jogo, este possui regras, estas listadas a seguir:

R1: O objetivo dos jogadores (atacantes) é capturar a maior quantidade de bandeiras que conseguir sem ser descoberto ou rastreado.

R2: Os jogadores (atacantes) podem ter seu acesso bloqueado a qualquer momento como consequência das suas ações.

R3: Os jogadores (atacantes) têm um tempo determinado de exatamente 2 (duas) horas para realizar a competição, ao final do período o ambiente de competição será desligado e a competição chegará ao fim.

R4: A progressão dos jogadores (atacantes) é incremental, ou seja, para passar para os próximos desafios é necessário capturar as bandeiras que contém informações necessárias para seguir a diante.

R5: Os jogadores não podem ter conhecimento sobre o escopo deste artigo, ou seja, eles não podem saber que existem técnicas e tecnologias de *Cyber Deception* espalhadas pelo ambiente. Para eles é somente um *capture the flag* customizado com as 4 regras anteriores.

3.1 ESTRUTURA DO AMBIENTE DE TESTES

O ambiente de testes foi construído, conforme ilustrado na **Figura 1*,* utilizando os serviços de computação em nuvem da *DigitalOcean*[1], utilizando $100 que contas criadas recebem para testar os serviços da provedora.

Neste ambiente, foram provisionadas 8 máquinas virtuais sendo que 7 delas estavam na mesma *VPC (Virtual Private Cloud)* de endereço 192.168.70.0/24, que nada mais é do que uma rede local em nuvem.

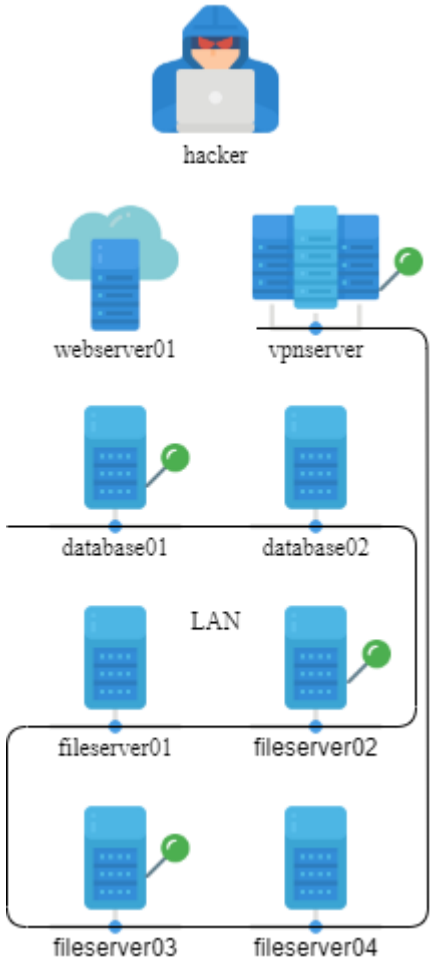
E uma única máquina não está dentro da *VPC* que é o servidor *webserver-01*.

A olustração a seguir mostra a interligação das máquinas por uma linha contínua que representa a rede local ou *VPN*.

[1] DigitalOcean: The developer Cloud. Disponível em: https://www.digitalocean.com.

* Ícones feitos por Freepik, monkik e disponibilizados em: https:// www.flaticon.com.

Figura 1: Topologia do ambiente de testes.



Fonte: Autor.

A *Tabela 1* disposta a seguir identifica cada servidor e seus respectivos endereços IP (público e privado):

Tabela 1: Endereçamento dos ativos do ambiente de testes.

Hostname	IP Privado	IP Público
vpn-server	192.168.70.2	104.131.35.237
database01	192.168.70.3	167.172.26.137
database02	192.168.70.8	142.93.184.217
fileserver01	192.168.70.4	104.248.8.40
fileserver02	192.168.70.5	167.71.182.79
fileserver03	192.168.70.6	104.131.124.167
fileserver04	192.168.70.7	167.71.173.254

Fonte: Autor.

A máquina virtual que não foi listada anteriormente, é apenas o servidor web que serve de início para o jogo. Trata-se do servidor de nome *webserver-01* que tem somente endereçamento público – 64.225.113.201, pois não há necessidade de estar na *VPC* das outras máquinas.

3.1.2 CONSTRUÇÃO DO AMBIENTE DE TESTES

Para construção o ambiente, foram seguidos os seguintes passos:

1) Provisinamento de oito máquinas virtuais, ou *Droplest*, que é como tratamos na plataforma da *DigitalOcean*, com a configuração básica de $10 por mês, as configurações são listadas na *Tabela 2*:

Tabela 2: Configurações dos Droplets utilizados.

Núcleos de processamento	1
Memória RAM	2GB
Transferência	2TB
Sistema Operacional	Ubuntu 18.04

Fonte: Autor.

2) Escolha de duas ferramentas de *Cyber Deception*, para implementação no ambiente de testes;
3) Implementação da ferramenta de *Cyber Deception OpenCanary*[1] no servidor de *hostname database01*;
4) Implementação da ferramenta de *Cyber Deception Artillery*[2] em seu modo

[1] Open Canary: Thinkst Applied Research. Disponível em: https://github.com/thinkst/opencanary
[2] Project Artillery by Binary Defense Systems. Disponível em: https://github.com/BinaryDefense/artillery

reativo no servidor de *hostname fileserver03*;
5) Implementação da ferramenta de *Cyber Deception Artillery* em seu modo de monitoramento de *filesystem*, monitorando o diretório "/home/user/secret-user/" no servidor de *hostname fileserver02*;
6) Implementação do coletor de logs *Filebeat*[3] para coletar e enviar os logs das ferramentas implementadas anterior mente para o servidor de monitoramento;
7) Implementação do banco de dados *MySQL* no servidor de *hostname database02* e criação da tabela "colaboradores" com algumas entredas de nome, cargo e CPF fictícios;
8) Implementação do serviço de FTP nos servidores de *hostname fileserver01* e *fileserver04*;
9) Implementação do serviço de VPN *OpenVPN*.
10) Implementação de um servidor web com as páginas "/index.html", "/admin.html" e "/robots.txt", com informações sobre o jogo.

Na página /admin.html do servidor web, foram plantadas 3 credenciais de acesso e o link para autenticação e download do arquivo de configuração da VPN, dentre elas, duas são verdadeiras. A finalidade disto é permitir a entrada dos atacantes já que o foco deste trabalho não é voltado para testes de invasão, mas esta estratégia também é utilizada como técnica de *Cyber Deception*, pois ao plantar credenciais falsas em lugares convenientes é possível rastrear a ação de agentes maliciosos já que elas não são utilizadas por ninguém dentro do ambiente corporativo.

3.2 TECNICAS E FERRAMENTAS DE CYBER DECEPTION

Dentro do ambiente de testes estão implementadas tecnologias/ferramentas de *Cyber Deception* que serão responsáveis por dar visibilidade da ação dos atacantes a medida que eles forem interagindo com tais tecnologias/ferramentas.

3.2.1 THINKST – OPENCANARY

O *OpenCanary*, é um projeto *open source* desenvolvido pela Thinkst que possui a finalidade de simular versões designadas "canarys" de serviços conhecidos e bastante utilizados para monitorar e alertar sempre que um desses serviços for solicitado, como por exemplo, *Secure Shell (SSH)*, *MySQL*, *File Transfer Protocol (FTP)*, etc.

No servidor *database01* estão emplementados versões canary dos serviços *SSH* (porta 22) e *MySQL* (porta 3306).

Neste artigo, a ferramenta *OpenCanary* é classificada como passiva, já que não realiza nenhuma ação além de monitorar e alertar.

3.2.2 BINARY DEFENSE – ARTILLERY

O projeto *Artillery*, desenvolvido pela *Binary Defense* é uma ferramenta que combina *honeypot*, monitoramento e alertas, além de que, no futuro, também apresentará função de inspeção e *hardening* de sistemas baseados em *UNIX* e sistemas *Windows*.

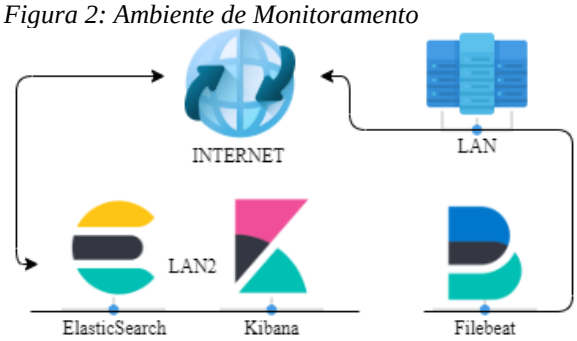
O ambiente de testes contempla duas implementações diferentes do *Artillery*, o servidor *fileserver02* conta com a função de monitoramento de *filesystem*, ou seja, é capaz de detectar modificações em um determinado diretório que foi previamente selecionado - "/home/secret-user/".

Já o servidor *fileserver03* conta com a função de "sentinela ativa", operando de maneira similar as versões *canary* de serviços mas com a habilidade de bloquear todo tráfego proveniente de um *host* que tenha interagido com alguma das versões *canary* escolhidas, neste caso, *SSH* (porta 22) e *FTP* (porta 21).

3.3 ESTRUTURA DO AMBIENTE DE MONITORAMENTO

[1] Open Canary: Thinkst Applied Research. Disponível em: https://github.com/thinkst/opencanary
[2] Project Artillery by Binary Defense Systems. Disponível em: https://github.com/BinaryDefense/artillery

Figura 2: Ambiente de Monitoramento



Fonte: Autor

Ambas as tecnologias apresentadas na sessão anterior possuem procedimentos para alertas via *e-mail* e entradas em arquivos de *log*. Para monitorar melhor e mais ativamente os alertas, este projeto utiliza um ambiente de monitoramento baseado na *Elastick Stack*, que está implementado como ilustrado na **Figura 2*.

O ambiente de monitoramento foi construído também utilizando os serviços de computação em nuvem da *Digital Ocean* e tem em sua implementação uma máquina virtual em que operam o *ElasticSearch*, para armazenamento e busca de eventos de *logs* que são enviados por agentes de envio chamados *Beats*, neste caso, existem 4 *Filebeats* instalados nas máquinas pelo ambiente de testes para dar visibilidade de toda interação que é feita com as ferramentas de *Cyber Deception* de uma maneira centralizada.

3.3.1 CONSTRUÇÃO DO AMBIENTE DE MONITORAMENTO

Para construção do ambiente de monitoramento foram seguidos os seguintes passos:

1) Provisionamento de mais um *Droplet* com as mesmas especificações dos outros;
2) Instalação do ElasticSearch 7.2;
3) Instalação do Kibana 7.2
4) Configuração das tabelas para apresentação dos logs coletados pelos *Filebeats* utilizando a ferramenta de apresentação *"Canvas"* do Kibana.

Os *Filebeats* estão separados em 4 servidores, sendo eles:

1) *database01:* para recolher as entradas de *logs* da ferramenta *OpenCanary*.
2) *fileserver02:* para recolher as entradas de *logs* da ferramenta *Artillery*.
3) *fileserver03:* para recolher as entradas de *logs* da ferramenta *Artillery*.
4) *vpn-server:* para recolher as entradas de *logs* referentes a conexão com o servidor de VPN.

Vale ressaltar que, o servidor *vpn-server* não possui nenhuma ferramenta de *Cyber Deception* implementada, é monitorado apenas por motivo de visibilidade.

A comunicação entre os *Filebeats* e o *ElasticSearch* é feita toda através da *Internet*, não sendo utilizada uma *VPC* para tal.

Todas as ocorrências são visualizadas, através de tabelas com as entradas dos *logs* coletados, renderizadas pelo *Kibana*, que consome estes dados do *ElasticSearch*.

3.4 OBJETIVOS PROPOSTOS PARA A CAPTURA DE BANDEIRAS

Existem 4 bandeiras, ou objetivos, que devem ser capturadas ou alcançados pelos atacantes, ou jogadores.

O foco da metodologia é observar, como as ferramentas de *Cyber Deception* espalhadas pelo ambiente de testes podem ajudar a rastrear e reagir as ações dos atacantes, e ao final, responder as duas perguntas propostas anteriormente. É importantes deixar claro que, é possível um atacante ser detectado somente na sua conexão com o ambiente interno, por meio da VPN, e depois disso, alcançar todos os objetivos propostos sem ser rastreado, até porque, a conexão com o ambiente interno não pode ser classificada como maliciosa pois foi feita com credenciais válidas.

* Ícones feitos por Freepik, monkik e disponibilizados em: https:// www.flaticon.com.

3.4.1 ORDEM DE AVANÇO DOS OBJETIVOS

Os objetivos estão dispostos de maneira sequêncial e dependente, como diz a **Regra 4 (R4)**. Sendo assim, o atacante deve fazer o determinado percurso, a fim de concluir o máximo de objetivos possíveis:

1) Analisar o código fonte da página "/admin.html" do servidor *webserver-01* e conseguir conectar-se a *VPC* do ambiente de testes.
1) Obter as informações da tabela "colaboradores" que está armazenada no servidor *database02.*
2) Visualizar o conteúdo do arquivo "plano-estrategico", ele revela as credencias necessárias para o último objetivo, e está armazenado no servidor *fileserver01* e está acessivel via *FTP*.
3) Por último, substituir o arquivo "clientes-prioritarios" por um arquivo de mesma estrutura mas com as modificações solicitadas descritas na página "/admin.html". Este arquivo encontra-se no servidor *fileserver04* e está acessível via *FTP*.

4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

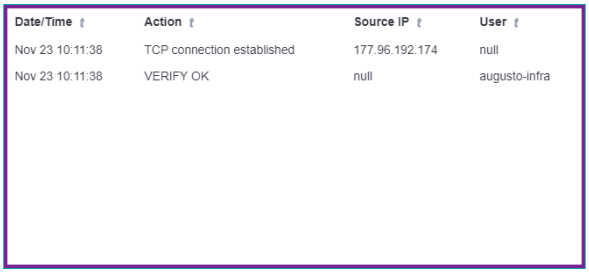
Foram realizados dois ataques ao ambiente, cada um por um profissional diferente e eles não tiveram comunicação um com o outro durante o teste, tampouco depois dos testes. Vamos analisar os ataques separadamente e descrever algumas atividades dos atacantes, visto que o ponto de destaque deste trabalho não é o ataque, nem as técnicas de invasão.

4.1 PRIMEIRO ATAQUE (10h - 33min)

O primeiro ataque, teve como fonte o endereço IP público 177.96.192.174, seu início às 10h e durou cerca de 33 minutos.

Inicialmente, o atacante passou, em média, 10 minutos analisando o *webserver-01*, logo após entender que para entrar no ambiente interno precisaria utilizar uma VPN emplementada com uso do software *OpenVPN*. Assim, como apresentado na *Figura* 3, o mesmo conseguiu conexão com o ambiente interno.

Figura 3: Logs de conexão com o servidor vpn-server.



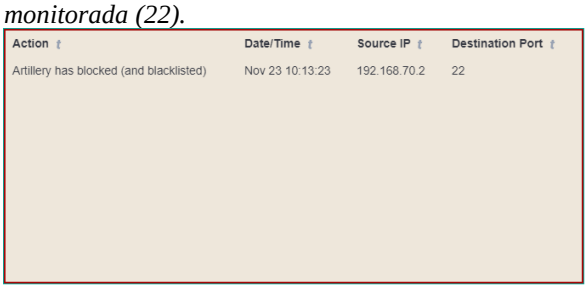
Date/Time	Action	Source IP	User
Nov 23 10:11:38	TCP connection established	177.96.192.174	null
Nov 23 10:11:38	VERIFY OK	null	augusto-infra

Fonte: Kibana, servidor de monitoramento.

Aqui, o atacante utilizou as credenciais "augusto-infra" e senha, plantadas nos comentários da página *admin.html* do *webserver-01*.

Logo mais, em questão de poucos minutos, tivemos um *log* da ferramenta *Artillery*, implementada no *fileserver03*. Por se tratar de um objetivo posterior e ter sido disparado de maneira precoce, a probabilidade deste evento ter partido de um *scaner* de portas é extremamente alta, pois o atacante não teria motivos, ainda no começo, para interagir em nenhum servidor de arquivos. A *Figura 4* o *log* gerado ao ativar o *Artillery*.

Figura 4: Log alegando interação com uma porta monitorada (22).



Action	Date/Time	Source IP	Destination Port
Artillery has blocked (and blacklisted)	Nov 23 10:13:23	192.168.70.2	22

Fonte: Kibana, servidor de monitoramento.

Uma observação importante sobre esse *log*, é que o IP apresentado como bloqueado e *blacklisted* é o endereço IP do

servidor *vpn-server*, isso acontece pois, o acesso as máquinas conectadas a *VPC* através da *VPN* acontece por meio de *Network Address Translation (NAT)*.

Network Address Translation nada mais é do que um método de traduzir endereços IP entre si. Geralmente, utilizado para traduzir endereços IP privados e não registrados em uma rede local para um ou mais endereços IP públicos e registrados para serem utilizados na internet (KUROSE, ROSS, 2017).

No próximo movimento, depois de enumerar os servidores *database01* e *database02* como servidores que estão com o *software MySQL* implementado, o atacante escolheu como primeiro alvo o servidor *database-01*, e como podemos ver nos *logs* que constam na *Figura* 5, gerados pela ferramenta *OpenCanary*, foram feitas diversas tentativas de login com as credenciais plantadas.

Figura 5: Tentativas de acesso ao banco de dados MySQL do servidor database01.

Date/Time	Username	Password	Source IP	Destination Port
2020-11-23 10:18:38.484542	joao-dba	8a03b76b7e6d4633a4d3e8d43c497f9d8c127a0e	192.168.70.2	3306
2020-11-23 10:20:59.362118	maria-araujo	a9aead91a029ebf49c6e4d43b0baa6b341c83a96	192.168.70.2	3306
2020-11-23 10:21:33.922103	augusto-infra	c73d6f121dd0fe078bbdf874cfc9464209219ad7	192.168.70.2	3306
2020-11-23 10:22:13.063561	joao-dba	8a03b76b7e6d4633a4d3e8d43c497f9d8c127a0e	192.168.70.2	3306
2020-11-23 10:22:26.615622	joao-dba	8a03b76b7e6d4633a4d3e8d43c497f9d8c127a0e	192.168.70.2	3306

Fonte: Kibana, servidor de monitoramento.

Depois disso, o atacante resolveu tentar acessar o banco de dados do servidor *database02* e obteve sucesso em capturar sua segunda bandeira, os dados da tabela "colaboradores", e agora ia para o próximo objetivo: os servidores de arquivo.

Ao começar a tentar obter informações sobre os servidores de arquivos o atacante decidiu parar com duas bandeira, e usou como argumento o fato do *host* de endereço IP 192.168.70.6, que apareceu no *scaneamento* de rede executando, na sua porta 22, um *tcpwrapper*, que é uma maneira de controlar o acesso a rede funcionando de modo a permitir ou negar conexões (HATCH et al., 2008) e por ter falhado em se autenticar no primeiro servidor que tentou acessar. Tendo em vista as regras propostas, o atacante resolveu parar neste estágio, conquistando duas bandeiras: o acesso a rede interna por meio da VPN e os dados da tabela "colaboradores" do servidor *database02*.

Segundo relato posterior do primeiro atacante, o ambiente parecia um tanto normal, mesmo após o primeiro escaneamento não haviam suspeitas de servidores falsos ou de ferramentas de *Cyber Deception*, o que começou a mudar no segundo escaneamento feito, pois um dos servidores não aparecia mais e isso era meio suspeito.

Com isso, pode-se então realizar uma análise sobre o primeiro ataque de acordo com a *Tabela 2*:

Tabela 3: Pontos principais sobre o primeiro ataque.

Completou o percurso?
Não, apenas duas bandeiras.
Foi rastreado?
Sim.
Foi bloqueado?
Sim

Fonte: Autor.

4.2 SEGUNDO ATAQUE (13h – 1h 5min)

Já o segundo ataque teve início às 13h, originou-se do IP público 189.40.77.57 e levou o dobro de tempo para ser encerrado. Em nota, o atacante relatou que conseguiu alcançar todos os objetivo.

Primeiramente, com 18 minutos de duração, o atacante conseguiu se conectar a *VPC* através da *VPN*, isso pode ser visto na *Figura 6*.

Figura 6: Logs de conexão com o servidor vpn-server

Date/Time	Action	Source IP	User
Nov 23 13:18:19	TCP connection established	189.40.77.57	null
Nov 23 13:18:20	VERIFY OK	null	augusto-infra

Fonte: Kibana, servidor de monitoramento.

Novamente, a credencial “augusto-infra” foi utilizada para realiza a autenticação e conexão com o ambiente de testes via VPN.

Segundo o atacante, durante o processo de enumeração, ele percebeu que um dos serviços de banco de dados MySQL estava em uma versão superior a outra, e assim suspeitou que pudesse ser uma armadilha.

De fato, não foi registrada nenhuma interação com o servidor *database01*, ou seja, neste cenário, o atacante conseguiu acessar o conteúdo da tabela “colaboradores” sem ser rastreado.

Seguindo, foi a vez de passar para o processo de enumeração dos outros servidores, em busca de servidores de arquivos.

Algo importante a se notar, é que, diferente do primeiro ataque, no segundo foi disparado um alerta do *Artillery*, mostrado na *Figura 7*, implementado no servidor *fileserver03* às 13:34, na porta 21, proveniente de uma tentativa de acesso *FTP*.

Figura 7: Log de bloqueio do Artillery, em fileserver03.

Action	Date/Time	Source IP	Destination Port
Artillery has blocked (and blacklisted)	Nov 23 13:33:44	192.168.70.2	21

Fonte: Kibana, servidor de monitoramento.

Pouco mais de meia hora depois, mais alertas do *Artillery*, mas dessa vez, da sua implementação no servidor *fileserver02*, onde atuou como ferramenta de monitoramento do sistema de arquivos *(filesystem)*, monitorando o diretório “/home/secret-user/”, diretório falso que simula o mesmo diretório no servidor *fileserver04*, contendo até o arquivo “clientes-prioritarios”, a diferença é que o diretório falso é monitorado, os *logs* de alteração estão listados na *Figura 8*, e que serve de distração, para ludibriar os atacantes com o útimo objetivo, que inclusive, foi o que aconteceu com o segundo atacante.

Figura 8: Alteraçõeso no diretório/home/secret-user/.

Date/Time	Action
Nov 23 14:01:25	changes were detected
Nov 23 14:01:30	changes were detected
Nov 23 14:02:45	changes were detected

Fonte: Kibana, servidor de monitoramento.

Ao verificar o arquivo “clientes-prioritarios” e ambos os servidores – *fileserver02*, *fileserver04* – fica claro que o atacante modificou o arquivo errado, pois teoricamente, o correto seria o arquivo localizado no servidor *fileserver04*, depois disso, o atacante finalizou seu ataque mas conseguiu capturar três bandeiras, ou concluir três objetivos.

Segundo relato posterior do segundo atacante, escaneamento especificamente as duas máquinas de banco de dados, foi possível notar que havia diferenças entre elas em relação aos serviços e sistemas operacionais respondidos, o que o fez escolher o servidor com os serviços mais atualizados e versão mais atual do sistema operacional. Isso ocorre por causa da implementação das ferramentas de *Cyber Deception*, e é um bom ponto a ser levando em consideração para pensar em uma padronização.

Novamente vamos preencher a *Tabela 3* com uma análisando os pontos importantes.

Tabela 4: Pontos principais sobre o segundo ataque.

Completou o percurso?
Não, capturou três bandeiras.
Foi rastreado?
Sim.
Foi bloqueado?
Sim

Fonte: Autor.

5 CONSIDERAÇÕES FINAIS

Apesar de existirem pesquisas, artigos, projetos e trabalhos sobre as técnicas, ferramentas e tecnologias que hoje foram designadas como *Cyber Deception*, desde dez anos atrás, esse assunto ainda está apenas no começo e precisa amadurecer bastante.

O fato é que, *Cyber Deception*, é necessário para melhorar as defesas cibernéticas contra agentes maliciosos, mas como criar indicadores, metas e objetivos para medir o desempenho das implementações de sistemas de segurança baseados em *Deception?*

Analisando os dados e as atitudes dos atacantes durante os dois ataques, pode-se pensar em como responder as dois perguntas propostas no início do artigo:

P1: Como *Cyber Deception* pode mudar o impacto final do ataque?

R1: No primeiro ataque, foi visto o *Artillery* entrar em ação logo nos primeiros minutos em que o atacante obteve acesso a rede interna do ambiente de testes, novamente, o que pode ter acontecido por um processo de *scanning* de rede para reconhecimento do ambiente. Ao ser ativado, o gatilho do *Artillery* em modo reativo, bloqueia toda a comunicação proveniente do endereço IP que provocou seu acionamento, se isso for feito de uma maneira global no ambiente, pode retirar o único meio de um agente malicioso navegar sobre a rede interna corporativa diminuindo, bruscamente, o impacto final de um ataque.

R1: No segundo ataque, algo de interessante foi visto também: o atacante achou que tinha capturado todas as bandeiras, mas na verdade ele foi enganado em relação a quarta e última, ou seja, se esse sentimento de certeza frustrada puder ser reproduzido com técnicas, tecnologias e ferramentas de *Cyber Deception*, é possível diminuir bastante o impacto final de um ataque, pois o agente malicioso terá certeza de que fez o seu trabalho mas na verdade causou menos dano ou até nenhum à empresa alvo.

P2: Como *Cyber Deception* pode influenciar no comportamento do atacante?

R2: Ao olhar o escopo do *Capture the Flag* como um todo, fica claro que o objetivo de aplicar as técnicas, tecnologias e ferramentas de *Cyber Deception* é justamente esse – rastrear, manipular, persuadir e enganar os agentes maliciosos para que eles ajam de uma maneira que facilita o trabalho da parte da defesa cibernética.

R2: Durante os dois ataques, não se observou muita mudança de comportamento nos atacantes, mas pode-se levar em consideração que o primeiro atacante desistiu depois de suspeitar ter sido descoberto, e o segundo atacante decidiu evitar técnicas de enumeração mais completas, pois já no início do ataque, notou que algo estava estranho entre os servidores e até conseguiu escapar de uma das técnicas de rastreamento e alerta, implementada em um dos servidores de banco de dados para executar versões *canary* dos serviços *SSH e MySQL*. Porém, o mesmo foi enganado na reta final de seu ataque, mostrando que, a diversidade estratégica das técnicas, tecnologias e ferramentas de *Cyber Deception* é um dos pontos fortes contra o crime cibernético.

Como conclusão de pensamento, também é possível criar algumas hipóteses analisando os ataques e os dados gerados, algumas hipóteses como, por exemplo:

H1: Os atacantes talvez ainda não se preocupem tanto com técnicas de defesa cibernética baseadas no paradigma de resposta, mesmo depois de esbarrar em algumas situações suspeitas, os dois atacantes avaliados estavam mais

preocupado em atingir os objetivos, o primeiro até desistiu depois de conseguir capturar a bandeira mesmo depois de já desconfiar que estava sendo rastreado.

H2: O ápice das tecnologias de defesa cibernética baseadas no paradigma de resposta virá quando começaram a desenvolver *frameworks* que possam ser repetíveis, escaláveis e centralmente gerenciados, algo na mesma linha, porém bem mais evoluído, do que foi feito neste trabalho, integrando as tecnologias de *Cyber Deception* em uma central única de armazenamento de *logs*, eventos, documentos, gerenciamento e, por último porém não menos importante, de visibilidade, como foi feito com a *Elastic Stack.*

REFERÊNCIAS

[1] AHMED, Atif. et al. **Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack.** *Computer Fraud & Security.* Elsever, 2019.

[2] BASKERVILLE, Richard. SPAGNOLETTI, Paolo. JONGWOO, Kim. **Incident-centered information security: Managing a strategic balance between prevention and response.** *Information & Management 51*. Elsevier, 2014.

[3] BENNETT, Michael. WALTZ, Edward. **Counterdeception Principles and Applications for National Security.** Artech House. Boston, 2007.

[4] BUSHBY, Andrew. **How deception can change cyersecurity defences**. *Computer Fraud & Security*. Elsever, 2019.

[5] CHACON, Joel. MCKEOWN, Sean. MACFARLANE, Richard. **Towards Identifying Human Actions, Intent, and Severity of APT Attacks Applying Deception Techniques** - An Experiment. *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*. IEEE, 2020.

[6] FireEye. **M-Trends 2020.** p. 14. Disponível em: https://content.fireeye.com/m-trends/rpt-m-trends-2020. Acesso 05 de outubro, 2020.

[7] HATCH, Brian. et al. **Hacking Exposed Linux:** Linux Security Secrets & Solutions. McGraw-Hill. Nova Iorque, 2008.

[8] HECKMAN, Kristin E. Et al. **Cyber Denial, Deception and Counter Deception.** Springer International Publishing. Suíça, 2015.

[9] HU, Zhisheng. et al. **Reinforcement Learning for Adaptive Cyber Defense Against Zero-Day Attacks.** *Adversarial and Uncertain Reasoning for Adaptive Cyber Defense.* Springer International Publishing. Suíça, 2019.

[10] KAHLHOFER, Mario. HÖLZL, Michael. BERGER, Andreas. **Towards Reconstructing Multi-Step Cyber Attacks in Modern Cloud Environments with Tripwires.** European Interdisciplinary Cybersecurity Conference. France, 2020.

[11] KUROSE, James F. ROSS, Keith W. **Computer Networking:** *A Top-Down Approach.* Pearson Education Limited. Ed 7. Londres, 2017.

[12] MCINNES, Nathaniel. ZALUSKA, Edwards. WILLS, Gary. **Analysis of threats on a VoIP Based PBX Honeypot.** University of Southampton. United Kingdom, 2018.

[13] SPITZNER, Lance. **Honeypots:** Catching the Insider Threat. *19th Annual Computer Security Applications Conference, 2003.* IEEE, 2003.

[14] VETTERL, Alexander. CLAYTON, Richard. **Honware:** *A Virtual Honeypot Framework for Capturing CPE and IoT Zero Days.* University of Cambridge. United Kingdom, 2019.

[15] YUILL, J.J. **Defensive Computer-Security Deception Operations:** Processes, Principles and Techniques. North Carolina State University, 2006.