UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
CENTRO MULTIDISCIPLINAR DE ANGICOS
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

ANDERSON CIRILO VALENTIM

TEXUGO SCAN: Coletor de Informações Livre para Testes de Intrusão

ANGICOS
2019

ANDERSON CIRILO VALENTIM

TEXUGO SCAN: Coletor de Informações Livre para Testes de Intrusão

Monografia apresentada a Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Sistemas de Informação.

Orientador: Prof. Dr. José Gildo de Araújo Jr.

ANGICOS

2019

C156t Cirilo Valentim, Anderson .
TEXUGO SCAN: Coletor de Informações Livre para Testes de Intrusão / Anderson Cirilo Valentim. - 2019.
36 f. : il.

Orientador: José Gildo de Araujo Júnior.
Monografia (graduação) - Universidade Federal Rural do Semi-árido, Curso de Sistemas de Informação, 2019.

1. Texugo Scan. 2. Framework ISSAF. 3. Coleta de Informações. 4. Testes de Intrusão. I. Araujo Júnior, José Gildo de, orient. II. Título.

Anderson Cirilo Valentim

TEXUGO SCAN: Coletor de Informações Livre para Testes de Intrusão

Monografia apresentada a Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Sistemas de Informação.

Defendida em: 14/ 08 / 2019.

BANCA EXAMINADORA

JOSE GILDO DE ARAUJO JUNIOR:06492397410
Assinado de forma digital por JOSE GILDO DE ARAUJO JUNIOR:06492397410
Dados: 2021.06.09 07:31:28 -03'00'

José Gildo de Araújo Júnior, Prof. Dr. (UFERSA)
Presidente

Assinado digitalmente por FRANCISCO DE ASSIS PEREIRA VASCONCELOS DE ARRUDA:04700207418
Localização: CMA - UFERSA
Data: 2021.06.08 23:59:42-03'00'

Francisco de Assis Pereira Vasconcelos de Arruda, Prof. Dr. (UFERSA)
Membro Examinador

JOEMIA LEILANE GOMES DE MEDEIROS:04586400420
Assinado de forma digital por JOEMIA LEILANE GOMES DE MEDEIROS:04586400420
Dados: 2021.06.09 10:10:57 -03'00'

Joêmia Leilane Gomes de Medeiros, Profa. Drª. (UFERSA)
Membro Examinador

RESUMO

Este trabalho teve por objetivo o desenvolvimento de uma ferramenta livre que unifica as etapas de coleta de informações e geração de relatórios presentes no *framework* ISSAF, amplamente utilizado por profissionais de segurança da informação na condução de testes de intrusão. Apesar de algumas ferramentas disponíveis no mercado auxiliarem na etapa de coleta de informação para testes intrusivos, muitas das suas funcionalidades são pagas, ou incompletas e seus códigos fonte não estão disponíveis para consulta e modificação, dificultando, assim, o trabalho dos profissionais com escassos recursos financeiros que, então, recorrem às suas próprias soluções. Neste sentido, a ferramenta denominada *Texugo Scan*, proposta neste trabalho, apresenta-se como uma alternativa que contempla as funcionalidades pagas e oportuniza à comunidade interessada todos os serviços necessários para a realização das etapas iniciais dos testes intrusivos de forma grátis e livre.

Palavras-chave: Texugo Scan; Framework ISSAF; Coleta de Informações; Testes de Intrusão.

ABSTRACT

This work aims to develop a free tool that unifies in a single project the steps of information gathering and reporting in the *ISSAF* framework, widely used by information security professionals in the conduct of penetration tests. Although some tools available in the market assist in the collecting information for intrusive tests, many of its functionalities are paid and its source codes are not available for consultation or modification, thus hindering the work of the professionals with scarce financial resources who must to build their own solutions. In this sense, the tool called *Texugo Scan*, proposed by this work, presents as an alternative that contemplates the paid functionalities and gives the interested community all the necessary services to perform the initial steps of the intrusive tests without costs.

Keywords: Texugo Scan; ISSAF Framework; Information Gathering, Intrusion Tests.

Sumário

1 INTRODUÇÃO

Segundo dados da *We Are Social (2018)***,** atualmente cerca de 4 bilhões de pessoas estão conectadas à Internet interagindo com uma infinidade de sistemas de informação. Esses sistemas, por sua vez, coletam, armazenam e posteriormente processam dados desses usuários com a finalidade de conhecê-los melhor, o que lhes permitem melhorar a qualidade dos serviços prestados, aumentar o poder de influência sobre eles e consequentemente a lucratividade das empresas. O potencial de utilização de dados pessoais de usuários é imenso: para recomendações em serviços (Google), músicas (Spotify), vídeos (Youtube, Netflix), compras (Amazon, Ebay), rede social para procura de emprego (LinkedIn) e diversas outras aplicabilidades. Garantir a segurança dessas informações é um desafio constante, uma vez que qualquer violação pode comprometer a imagem, a integridade, a confiabilidade e, por fim, o valor de mercado da empresa.

Uma pesquisa realizada pela Ernst & Young (2017), com 1.200 executivos da área de cibersegurança constatou que apenas 4% das empresas se sentem preparadas para enfrentar ataques cibernéticos. Ainda de acordo com essa pesquisa, 70% dos entrevistados afirmam que as companhias aumentariam seus investimentos na área caso algum incidente ocorresse e causasse danos significativos aos negócios. Segundo previsão da Gartner (2018) os investimentos em segurança da informação no ano de 2018 chegaram a US$ 96 bilhões sendo considerando um mercado em franca expansão.

Segundo McGraw (2006) e posteriormente Kissel et al. (2008), um software seguro, é aquele que satisfaz os requisitos implícitos e explícitos de segurança em condições normais de operação e em situações decorrentes de atividade maliciosa de usuário. Para Dantas (2011), um sistema de informação seguro necessita garantir três características fundamentais:

- **Integridade**: é a garantia da exatidão e completeza da informação e dos métodos de processamento;

- **Disponibilidade:** é a garantia de que os usuários autorizados obtenham acesso à informação e aos ativos correspondentes sempre que necessário;
- **Confidencialidade:** é a garantia de que a informação é acessível somente por pessoas autorizadas a terem acesso.

Estas características devem ser preservadas, pois são tidas como princípios da segurança da informação. No intuito de garantir que as três características não sejam violadas, muitas empresas contratam profissionais na área de segurança da informação, conhecidos como *pentesters*, para testá-las em seus sistemas de modo a mitigar possíveis falhas que podem resultar em vulnerabilidades, e, por fim, comprometer tanto as informações do usuário quanto a infraestrutura em que seus sistemas estão inseridos.

Para a realização desses testes, comumente chamado *Pentest (acrônimo para testes de intrusão)*, é necessário entender o escopo da organização, rede ou sistema a ser testado. Tais testes contém etapas que de forma genérica podem ser definidos nas seguintes atividades: preparação, coleta de informações, modelagens de ameaças, análise e exploração de vulnerabilidades, e relatório do teste. Contudo, devido a diversidade de organizações e sistemas de informações faz-se necessário diferentes metodologias, *frameworks*, e modelos de testes de segurança específicos a cada cenário.

Na pesquisa realizada por Bertoglio e Zorzo (2016) foi possível identificar as seguintes metodologias, frameworks e modelos de testes de segurança, são eles:

- OSSTMM (*Open Source Security Testing Methodology Manual*);
- ISSAF (*Information Systems Security Assessment Framework*);
- PTES (*Penetration Testing Execution Standard*);
- *GNST (Guideline on Network Security Testing)*;
- *OWASP Testing Guide*.

Já Pandrini (2010) identificou e comparou as seguintes metodologias, frameworks e modelos de testes de segurança, são eles: ISSAF (*Information Systems Security Assessment*

Framework), OSSTMM (*Open Source Security Testing Methodology Manual*), *BHM (Black Hat Methodology)*, e *GNST (Guideline on Network Security Testing)* como mostra a Figura 1 a seguir.

Figura 1. Comparação de metodologias para testes de intrusão

	ISSAF	OSSTM	BHM	GNST
Modelagem	+	=	-	-
Planejamento	+	-	-	-
Flexibilidade	-	-	-	+
Adaptação	=	+	+	=
Relatório	=	=	=	+
Orientação	-	=	-	+
Granularidade	+	=	-	-

Descrição: + Bom, = Médio, - Limitado ou nenhum

Fonte: Adaptado de Pandrini (2010)

O primeiro critério foi a modelagem, onde somente o *framework* ISSAF ofereceu características para elaboração de um cenário de teste de intrusão específico. No critério planejamento, o ISSAF também apresentou vantagem devido a sua possibilidade de associar a tarefa a ser executada com a técnica a ser utilizada.

No critério flexibilidade, o ISSAF apresentou desvantagem por não apresentar subjetividade em como deve ser realizada cada etapa do teste. Isso leva o *pentester* a sempre seguir o mesmo roteiro de execução do teste de intrusão, o que pode inviabilizar a descoberta de problemas ao seguir diferentes caminhos. No critério de adaptação, o ISSAF, mostrou-se moderado. A utilização de técnicas bem definidas pode ser executada em ferramentas diferentes das recomendadas, o que, em situações específicas pode ser mais eficiente do que o recomendado. Em relação aos relatórios o ISSAF apresentou falta de clareza em como deve ser documentado o teste de intrusão. A granularidade é uma das principais características presente no *framework* ISSAF, devido ao seu detalhamento, o que permite ao *pentester* iniciante entender de forma específica cada atividade executada.

A abrangente documentação e sua facilidade de associar as tarefas a serem executadas com as ferramentas que podem ser utilizadas, é um das principais vantagens do *framework* ISSAF, pois faz com que o *pentester* tenha domínio claro do que deve ser feito, evitando erros que venham a ocorrer caso seja montado um ataque de forma aleatória. Por essas características, este trabalho se propõe a utilizar este *framework ISSAF* (Rathore et al. 2006) como ponto base de seu desenvolvimento.

A coleta de informações descrita no *framework* ISSAF é a etapa principal para garantir que um teste de intrusão seja bem-sucedido. Essa atividade consiste em reunir o maior número de informações possíveis do cliente, e, em grande medida, encontra na Internet o meio mais propício para cumprir o seu fim. Contudo, não é trivial analisar e processar os dados presentes na Internet, pois necessitaria de uma grande quantidade de tempo para análise dos dados úteis.

O mercado oferece uma gama de ferramentas específicas para esse tipo de coleta contendo funcionalidades relevantes para a etapa de coleta de informações de um teste de intrusão de modo a garantir eficiência nesta etapa, distribuições Linux como o Kali (OFFENSIVE SECURITY, 2019) e Parrot (PARROT SEC, 2019), utilizam dessas ferramentas como padrão em suas distribuições, contudo, é necessário utilizar diferentes ferramentas para realizar determinadas técnicas de coleta de informações, o que custa tempo na hora de documentar o teste de intrusão.

Conforme Weidmam (2014) tanto o Maltego (PATERVA, 2019), quanto o Netcraft (NETCRAFT, 2019), por exemplo, são ferramenta capazes de coletar informações públicas na Internet e organizá-la de forma simples no intuito de montar um cenário de intrusão, e estão disponíveis para utilização em distribuições linux para realização de testes de intrusão. O problema é que essas ferramentas só são amplamente úteis em versões pagas. O número máximo de resultados esperados na versão cliente do Maltego sai de 64.000 na versão paga para apenas 12 na versão gratuita, por exemplo, impossibilitando que profissionais com escassos recursos financeiros consigam utilizá-la em cenários reais, obrigando-os a

construírem suas próprias soluções, ou utilizarem ferramentas disponíveis em distribuições Linux que atendem técnicas específicas para coleta de informações, mas não agrupado em uma única ferramenta uma quantidade satisfatória de técnicas, fazendo com que o *pentester* utilize diversas ferramentas para se obter um resultado satisfatório de coleta de informações, o que não é prático quando se busca eficiência e rapidez ao executar essa etapa e na elaboração do relatório da mesma.

Este trabalho busca desenvolver uma ferramenta livre para realização das etapas de coleta de informações e geração de relatórios baseado no *framework* ISSAF que seja suficiente para a etapa de coleta de informações em testes de intrusão que busque agrupar técnicas de coleta de informações em um único software, de modo a não ser necessário a utilização de outras ferramentas para técnicas específicas, diferente de ferramentas encontradas em distribuições Linux com o foco em segurança da informação.

1.1 Objetivos

A seguir, são apresentados os objetivos deste trabalho.

1.1.1 Objetivo Geral

Este trabalho teve como objetivo o desenvolvimento de uma ferramenta livre que auxilie o profissional de segurança da informação (*pentester*) nas atividades de coleta de informações e geração de relatórios de alvos, unificando as recomendações presentes no *framework* ISSAF.

1.1.2 Objetivos Específicos

- Levantar requisitos funcionais por meio da análise das ferramentas *Maltego* e *Netcraft* em suas versões pagas;

- Desenvolver o projeto arquitetural do *software* por meio de diagramas UML tendo como base os requisitos coletados;
- Desenvolver o protótipo de interface gráfica;
- Desenvolver o *software* (*Texugo Scan*) de acordo com as técnicas de coleta de informações e geração de relatórios estabelecidas no *framework* ISSAF;
- Comparar o *Texugo Scan* com o Maltego e o Netcraft em suas versões pagas;
- Publicar artigo científico disponibilizando à comunidade os progressos alcançados.

1.2 Justificativa

A execução ideal de um planejamento de testes por parte do *pentester* envolve a utilização de um amplo conjunto de técnicas e ferramentas de software, parte delas disponíveis apenas em versões pagas, o que inviabiliza o trabalho de *pentesters* iniciantes que não dispõem de recursos suficientes para obtê-los, comprometendo os resultados obtidos em tempo e qualidade. Este trabalho se propõe a conceber uma ferramenta de coleta de informações automática de modo a permitir à *pentesters* com poucos recursos disponíveis a possibilidade de executarem as etapas de coleta de informações e geração de relatórios de forma gratuita tão eficiente quanto as versões pagas disponíveis atualmente.

2 FUNDAMENTAÇÃO TEÓRICA

A seguir será descrita a fundamentação teórica deste trabalho.

2.1 Testes de Invasão

Os testes de intrusão simulam em ambiente real a avaliação de riscos provenientes de prováveis brechas de segurança. Segundo Weidmam (2014), os pentesters não só identificam possíveis falhas de segurança como também exploram essas vulnerabilidades, no intuito de avaliar o que possíveis invasores poderiam fazer, caso viessem a obter êxito na exploração da vulnerabilidade.

De acordo com Bertoglio e Zorzo (2015) tais testes são capazes de fornecer um nível de verificação de segurança adequado e com muitos detalhes a respeito das fraquezas do alvo. Contudo, não se deve confundir a análise de vulnerabilidade com o teste de intrusão.

Para Potter e McGraw (2004) o *pentester* deve utilizar uma abordagem baseada no risco, no ataque a arquitetura com a mentalidade do atacante caso fosse fazer infringir a segurança do ambiente testado. Já a análise de vulnerabilidade faz a identificação das principais vulnerabilidades presentes na rede ou sistema, logo após é elaborado uma lista com as principais vulnerabilidades encontradas levando em consideração a sua criticidade, ou seja, o intuito da análise de vulnerabilidade é identificar e classificar a vulnerabilidade e seu impacto, já o intuito do teste de intrusão é explorar a vulnerabilidade no intuito de obter acesso de forma maliciosa ao ambiente testado.

Para a realização de um teste de intrusão é essencial entender o escopo abordado pelo cliente. Neste teste é definido as principais diretrizes que o *pentester* irá seguir para a realização da intrusão, como também, o tipo de teste de intrusão que será feito. Segundo Soares (2010) os testes de intrusão são definidos em 3 tipos, sendo eles:

- **Caixa Preta:** O *pentester* tem pouco ou nenhum conhecimento sobre o ambiente a ser testado, tudo deve ser descoberto pelo profissional;
- **Caixa Branca:** O *pentester* dispõe de todas as informações possíveis do alvo, supondo que o atacante conseguiu obter todos os dados da vítima para a condução da invasão;

- **Caixa Cinza (Híbrido)**: O *pentester* dispoẽ de informações limitadas da vítima, sendo o mínimo necessário para a condução do teste de intrusão.

2.1.1 Etapas de um Teste de Intrusão

No intuito de conduzir de forma estruturada o teste de intrusão, Weidman(2014) definiu 7 fases essenciais para a realização de um teste de intrusão, são elas:

1. **Preparação (*pre-engagement*)**: Nessa fase o profissional conversa com o cliente a respeito de seus objetivos para o teste de invasão;
2. **Coleta de informações (*information-gathering*)**: Nessa fase o *pentester* procura informações disponíveis publicamente sobre o cliente, e identifica maneiras de conectar-se com seus sistemas;
3. **Modelagem de ameaças (*threat-modeling*)**: O *pentester* usa as informações coletadas na fase anterior para determinar o valor de cada descoberta no intuito de determinar o impacto sobre o cliente em caso de a descoberta permitir que alguém invada os sistemas;
4. **Análise de vulnerabilidades (*vulnerability analysis*):** Nessa fase o *pentester* procura descobrir vulnerabilidades nos sistemas;
5. **Exploração de falhas (*exploitation*):** Nessa fase o *pentester* tenta invadir o sistema do cliente;
6. **Pós exploração de falhas (*post-exploitation*):** Caso a etapa anterior seja bem-sucedida, o *pentester* tirar vantagem o resultado, de modo a descobrir informações adicionais e obter dados críticos e acessar outros sistemas;
7. **Geração de Relatório (*reporting*):** A fase final, o pentester sintetiza as descobertas tanto para os profissionais executivos, quanto para os técnicos.

Esse modelo padrão não abrange todos os cenários de invasão possíveis, para isso foram desenvolvidas metodologias, frameworks, e testes de segurança, como a ISSAF, abordada neste trabalho.

2.2 Framework ISSAF

O *framework* ISSAF tem por objetivo avaliar a segurança de redes, sistemas e aplicações, ele concentra-se em 3 áreas de concepção:

- **Planejamento e Preparação:** Tem por objetivo definir o ambiente de teste, os *softwares* a serem utilizados, contratos e aspectos legais, definição da equipe de trabalho, prazos, requisitos e estrutura dos relatórios finais;
- **Avaliação**: Nela são realizados os testes de intrusão baseados em 9 atividades: coleta de informação, mapeamento de rede, identificação de vulnerabilidades, invasão, acesso e escalação de privilégios, enumeração, comprometimento de usuários remotos, manutenção de acesso, e ofuscamento de rastros;
- **Relatório**, **Limpeza** e **Destruição de Artefatos**: A fase final do framework, logo após a execução das etapas anteriores, é confeccionada pelo profissional de segurança da informação, um relatório decorrente das falhas e vulnerabilidades encontradas e propostas de correção. Também é feita a destruição de artefatos, com finalidade de não deixar nenhum rastro que comprometa a integridade do cliente.

2.2.1 Atividade de Coleta de Informação

A atividade inicial da etapa de avaliação do framework ISSAF, é a coleta de informações. Conforme Macedo (2016) o intuito dessa atividade é obter informações sem ao menos atacar o cliente diretamente (coleta de informações passiva), também conhecida como OSINT (*Open Source Intelligence*). Na maioria dos casos, a principal fonte de informação (e possivelmente a única) é a Internet. Conforme Rathore et. al (2006) a Internet pode fornecer informações sobre o alvo (empresa e / ou pessoa) usando vários métodos, tanto técnicos (por exemplo DNS / WHOIS) como não técnicos, como por exemplo: motores de busca, grupos de notícias, listas de discussão, entre outros. Além da coleta de informação passiva, também existe a ativa, onde o *pentester* interage diretamente com o ambiente a ser testado.

3 METODOLOGIA

Para o desenvolvimento da ferramenta proposta neste trabalho, o passo inicial foi o levantamento dos requisitos de *software* que foram feitos por meio da análise das versões pagas das ferramentas Maltego (PATERVA,2019) e Netcraft (NETCRAFT, 2019), essas ferramentas, segundo a *INFOSEC INSTITUTE*(2019) são amplamente utilizadas na etapa de coleta de informações em um pentest. Ambas as ferramentas foram utilizadas para a etapa de coleta de informações de teste de intrusão com alvos reais seguindo um roteiro de coleta de informações estruturado em conformidade com o *framework* ISSAF. Com base na análise das ferramentas, os requisitos funcionais foram coletados. Logo após, foi realizada a modelagem *UML* do sistema, utilizando o *software* Astah (ASTAH,2019) de modo a contemplar todas as funcionalidades em comum levantadas. Os artefatos dessa etapa foram os diagramas de caso de uso e de classes. O diagrama de caso de uso foi utilizado por ser uma excelente alternativa para a compreensão de como a ferramenta relaciona-se com o usuário que o estão utilizando, já o diagrama de classe, apresenta características estruturais da ferramenta desenvolvida, o que facilita a compreensão de outros desenvolvedores na hora de implementação de novas funcionalidades da ferramenta, por essas vantagens, utilizou-se esses diagramas.

A prototipação das telas foi realizada por meio da ferramenta *Wireframes Balsamiq (BALSAMIQ,2019)* enquanto a implementação será desenvolvida por meio do *framework Web Bootstrap (BOOTSTRAP,2019)*. O desenvolvimento das funcionalidades levantadas foram feitas na linguagem de programação *python* utilizando especificamente o *framework* para desenvolvimento Web Flask (FLASK,2019). O ambiente de desenvolvimento foi o Visual Studio Code (VISUAL STUDIO CODE, 2019). E Todo código produzido está disponibilizado de forma gratuita no Github (GITHUB,2019) sob a licença MIT(MIT,2019).

Por fim, a validação da ferramenta desenvolvida, denominada de *Texugo Scan*, foi realizado por meio da análise das medidas de **tempo de execução** entre as ferramentas Texugo Scan e *Netcraft*, **completude** entre todas as funcionalidades avaliadas pelas ferramentas, e **análise estatística** dos resultados obtidos para as funcionalidades (S*ubdomain enumeration,e Technology Identification*), gerados por cada uma das ferramentas sendo

analisadas.

3.1 Desenvolvimento do Texugo Scan

Neste capítulo, as etapas do desenvolvimento deste trabalho são detalhadas. Inicialmente, é mostrada uma visão geral sobre como o sistema foi desenvolvido, detalhando as técnicas e recursos utilizados. Posteriormente, é apresentado o modelo arquitetural em UML utilizando o diagrama de caso de uso e diagrama de classe. Em seguida, é apresentado as interfaces da ferramenta desenvolvida, bem como o modelo de relatório gerado.

3.2 Visão Geral

Com base nas análises das ferramentas Maltego e Netcraft, os seguintes requisitos funcionais foram coletados como evidência como mostra a Tabela 2 a seguir.

Tabela 2- Requisitos Funcionais do Texugo Scan

Requisito Funcional	**Descrição**
Whois (RF1)	A ferramenta deve disponibilizar as informações do Protocolo Whois para o usuário de acordo com a URL disponibilizada.
***Port Scan* (RF2)**	A ferramenta deve efetuar uma varredura de portas no servidor de hospedagem da URL informada.
Enumeração de Subdomínios (RF3)	A ferramenta deve efetuar a enumeração de subdomínios de acordo com o domínio da URL informado.
Identificação de Bloco de IP (RF4)	A ferramenta deve efetuar a identificação de informações referente ao bloco de endereços IP's ao qual a URL informada está hospedada.

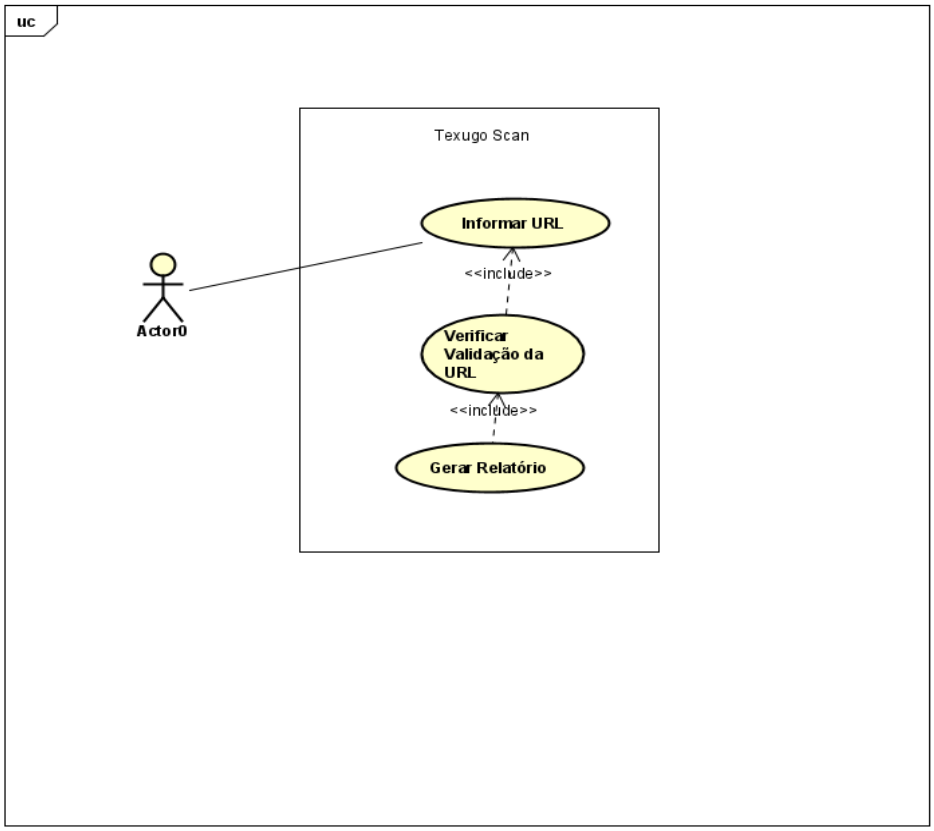
Captura de Banner [1](RF5)	A ferramenta deve capturar o Banner do servidor onde encontra-se hospedados o site.
Identificar o Servidor de DNS Reverso (RF6)	A ferramenta deve fazer a identificação de servidores do Servidor DNS Reverso onde encontra-se armazenado o site da URL informada.
Identificação de Tecnologias (RF7)	A ferramenta deve identificar as tecnologias utilizadas pelo site da URL informada.
Verificação (RF8)	Verificar se a URL informada é válida ou não.
Gerar Relatório	A ferramenta deverá gerar um relatório com base nos resultados obtidos dos requisitos: **RF1, RF2, RF3, RF4, RF5, RF6, RF7.**

Fonte: Autoria Própria.

Considerando os requisitos descritos anteriormente, foi possível modelar as funcionalidades do sistema proposto. Para isso, inicialmente, é apresentado na Figura 2 o diagrama de casos de uso do sistema.

[1] Captura de Banner, ou Banner Grabbing, trata-se de uma técnica utilizada para coleta de informações de serviços que estão sendo utilizados em determinado servidor na WEB.

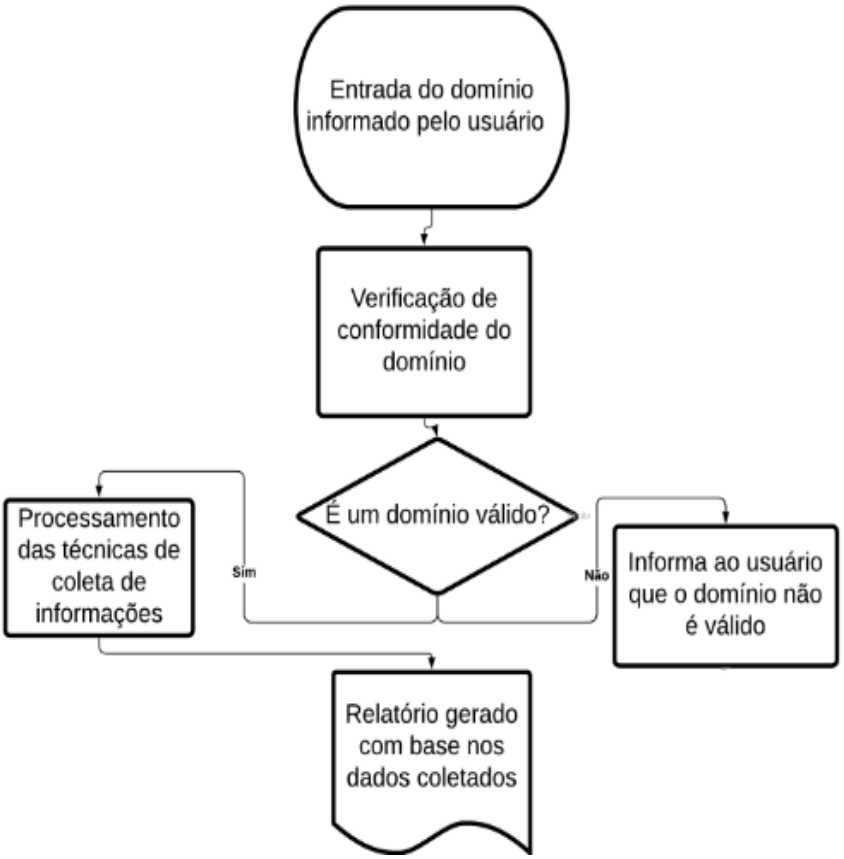
Figura 2. Diagrama de Caso de Uso



Fonte: Autoria Própria.

A partir do diagrama de caso de uso, o seguinte fluxograma foi desenvolvido, para entendimento do fluxo de execução da ferramenta Texugo Scan, mostrando as etapas lógicas a serem executadas até a emissão do relatório, como mostra a Figura 3.

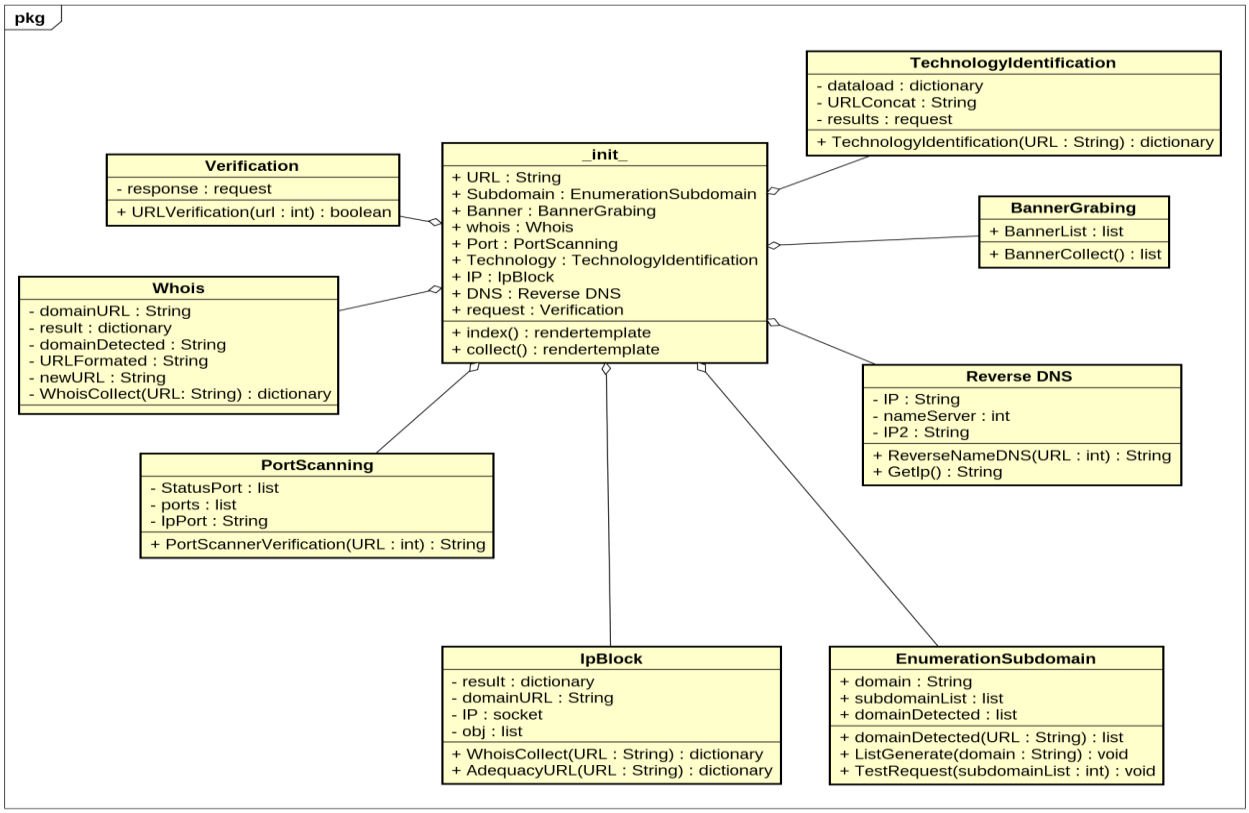
Figura 3. Fluxograma de Execução do Texugo Scan



Fonte: Autoria Própria.

Com base nos requisitos funcionais levantados, o seguinte diagrama de classe foi desenvolvido no intuito de descrever a estrutura do software do Texugo Scan como mostra a Figura 4.

Figura 4. Diagrama de Classe Texugo Scan



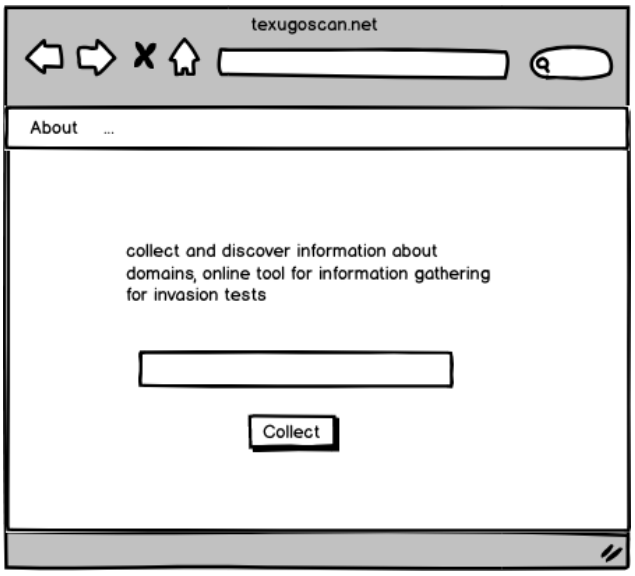
Fonte: Autoria Própria.

3.3 Prototipagem de Interfaces

Os protótipos das interfaces foram desenvolvidos no intuito de modelar a interface da ferramenta para que posteriormente fosse realizada o desenvolvimento das telas em HTML utilizando a *framework Bootstrap*. Segundo SOMMERVILLE (2007), o objetivo da prototipação de interfaces é permitir que os usuários do sistema desenvolvido ganhem experiência com a interface que será desenvolvida.

Dessa forma, foi desenvolvido um protótipo de média fidelidade, especificando como seria o layout da página inicial do sistema desenvolvido, a figura 5, mostra o protótipo desenvolvido.

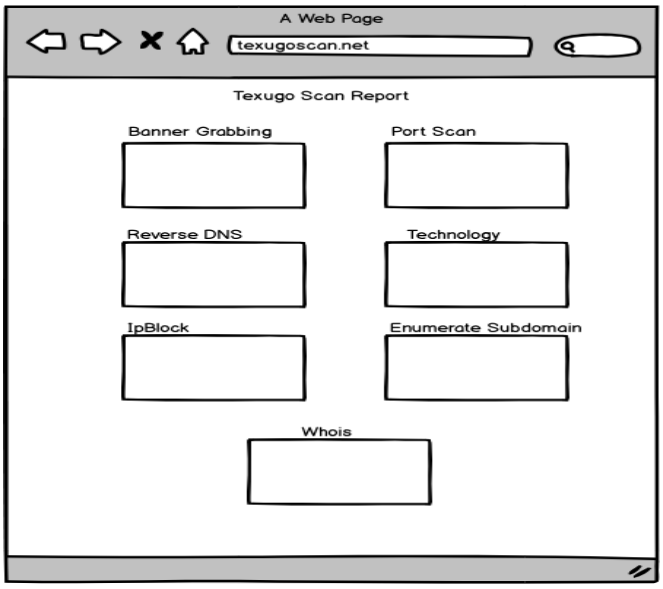
Figura 5. Protótipo de Média Fidelidade da Interface Inicial do Texugo Scan



Fonte: Autoria Própria

Após o protótipo da tela inicial do sistema, foi desenvolvido o layout da página onde é gerado o relatório da ferramenta, como mostra a Figura 6.

Figura 6. Protótipo de Média Fidelidade do Relatório Gerado pelo Texugo Scan

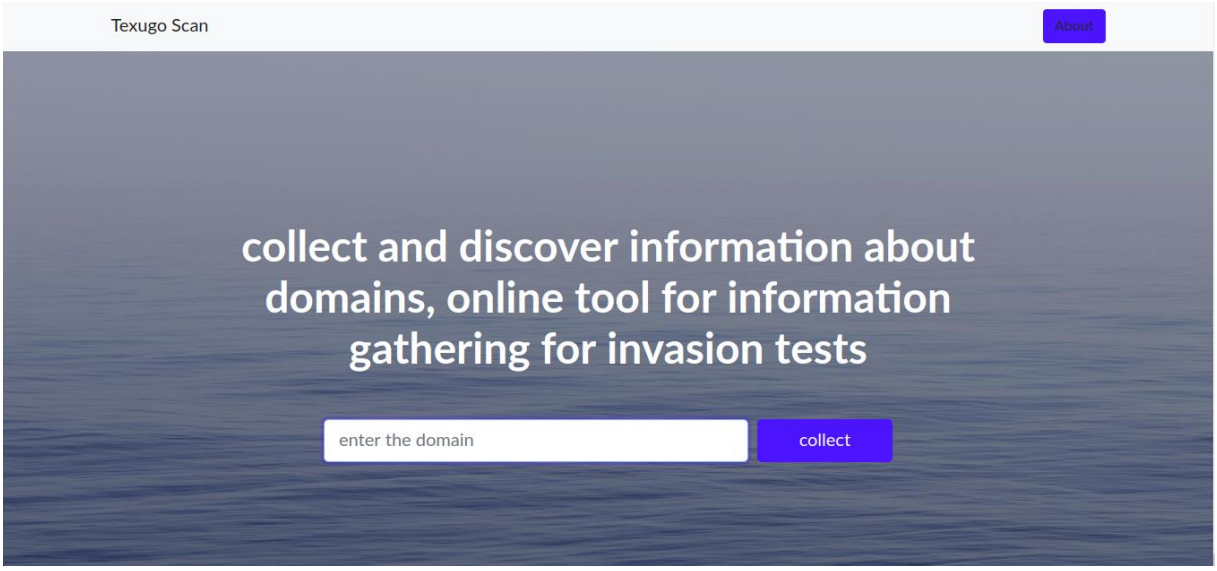


Fonte: Autoria Própria

3.4 Interface do Texugo Scan

Com base nos protótipos de interface desenvolvidos, as seguintes interfaces foram desenvolvidas bem como as suas funcionalidades comentadas anteriormente na seção 4.1. A interface inicial da ferramenta, recebe como entrada uma URL, que será informado no campo de texto, logo após, é necessário clicar no botão **collect** para iniciar a etapa de processamento e geração de relatório, a Figura 7 mostra a interface inicial do Texugo Scan.

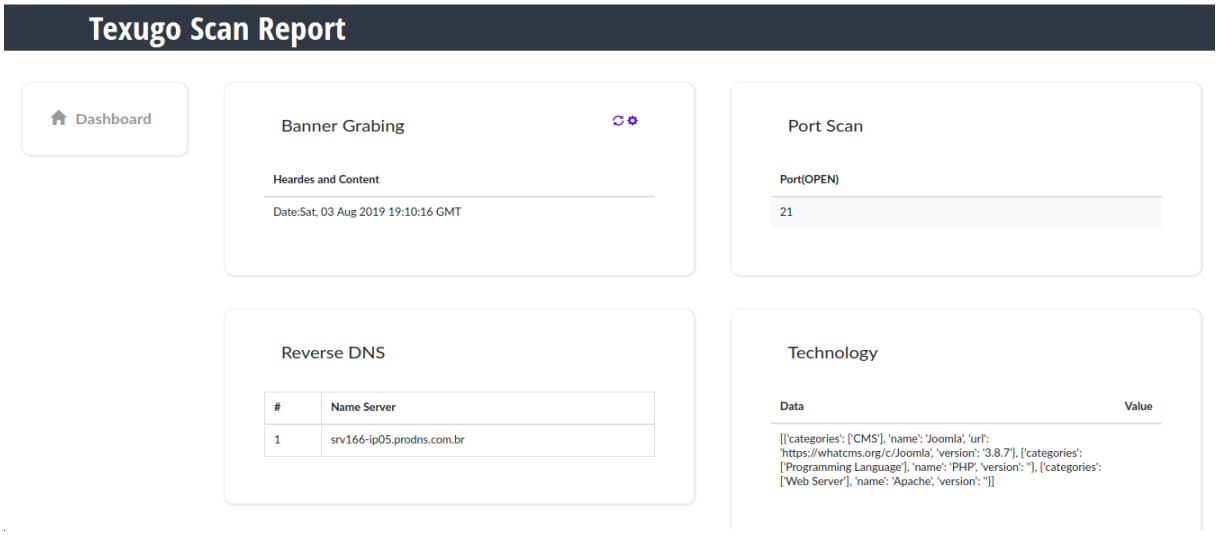
Figura 7. Interface Inicial do Texugo Scan



Fonte: Autoria Própria

Após o usuário informar a URL, é iniciada a etapa de verificação. Em caso de a URL ser válida, o processamento é iniciado. O resultado do processamento é a apresentação de um relatório contendo o resultado da aplicação das várias técnicas de coleta de informações como mostra a Figura 8.

Figura 8. Interface do Relatório do Texugo Scan



Fonte: Autoria Própria

4. VALIDAÇÃO

Na etapa de validação, buscou-se analisar os resultados obtidos pelas ferramentas, no intuito de compará-las nas funcionalidades em comum entre elas.

Para início da etapa de Validação, o *Texugo Scan*, ferramenta proposta neste trabalho, foi comparado com as ferramentas Maltego e *Netcraft*, ambas as ferramentas receberam como entrada os mesmos sites que serviram como alvos para coleta de informações, onde as 3 ferramentas testadas desenvolveram relatórios sobre os alvos. A priori, pesquisou-se sites de municípios do estado do Rio Grande do Norte, como também de órgãos da administração pública do Estado, concentrando-se em 25 alvos para análise específica. Além dos 25 alvos ligados diretamente a administração pública do estado, pesquisou-se na plataforma Alexa (ALEXA, 2019) os websites mais visitados do Brasil, ao final, desenvolvem-se uma lista com os sites devidamente organizados como mostra a Tabela 3.

Tabela 3- Lista de sites utilizado para análise de dados

http://www.carnaubais.rn.gov.br/	https://www.facebook.com/
https://assu.rn.gov.br/	https://www.globo.com/
http://angicos.rn.gov.br/	https://outlook.live.com/owa/
https://www.fernandopedroza.rn.gov.br/	https://www.mercadolivre.com.br/
https://www.afonsobezerra.rn.gov.br/	https://www.uol.com.br/
https://altodorodrigues.rn.gov.br/	https://www.otvfoco.com.br/

https://www.defensoria.rn.def.br/	https://www.blogger.com/about/?r=2
https://natal.rn.gov.br/	https://br.yahoo.com/?p=us
http://www.tce.rn.gov.br/	https://www.metropoles.com/
http://www.cotic.rn.gov.br/	https://www.netflix.com/br-en/
https://ufersa.edu.br/	https://www.instagram.com/
https://paudosferros.rn.gov.br/	https://olx.com.br/
https://caico.rn.gov.br/	https://twitter.com/
http://www.caern.rn.gov.br/	https://www.xvideos.com/
http://portal.uern.br/	https://www.americanas.com.br/
https://www.ufrn.br/	https://www.msn.com/pt-br/
http://www.caraubas.rn.gov.br/	http://www.caixa.gov.br/Paginas/home-caixa.aspx
http://www.pm.rn.gov.br/	https://www.bet365.com/
https://portodomangue.rn.gov.br/	https://www.aliexpress.com/
http://macau.rn.gov.br/	https://www.microsoft.com/pt-br/
https://apodi.rn.gov.br/	https://www.reclameaqui.com.br/
https://www.prefeiturademossoro.com.br/	http://fazenda.gov.br/
http://www.policiacivil.rn.gov.br/	https://www.bol.uol.com.br/
http://www.mprn.mp.br/portal/	https://www.amazon.com.br/
https://www.google.com.br/	https://abril.com.br/

Fonte: Autoria Própria

Após a etapa de elaboração da lista de alvos para coleta de informações, os sites foram submetidos à análise das ferramentas Maltego, Netcraft e Texugo Scan, avaliando, em cada uma delas as funcionalidades de coleta de informações do *framework* ISSAF. A Tabela 4 apresenta a lista de funcionalidades comparadas entre as ferramentas testadas.

Tabela 4- Lista de Técnicas para coleta de informações

ID	**Técnica**	**Descrição**
T1	*Whois*	Protocolo para identificar informações sobre entidades DNS na internet.
T2	*Port Scan*	O Port Scan tem o objetivo de mostrar quais portas TCP/UDP estão abertas em um determinado *Host*.

T3	*Subdomain enumeration*	Consiste na identificação de ramificações de um domínio público na Internet.
T4	*Netblock Identification*	Consiste na identificação de informações relativas ao bloco de IP's do qual o site está hospedado.
T5	*Banner Grabbing*	Consiste na obtenção de informações sobre serviços que estão sendo utilizados em determinado computador ou servidor de hospedagem.
T6	*Reverse DNS Identification*	Consiste na identificação de um nome de domínio associado a um endereço IP na Internet.
T7	*Technology Identification*	Consiste na identificação de tecnologias que estão sendo utilizadas por determinado site ou sistema.

Fonte: Autoria Própria

Para a validação da ferramenta foram realizadas duas análises. Completude e Eficiência, para análise de eficiência foi medido o tempo de execução das ferramentas Texugo Scan e *Netcraft*. Devido sua falta de portabilidade por ser uma ferramenta *Desktop*, o software Maltego não entrou no critério de validação por tempo de execução, como também a análise quantitativa dos dados encontrados das técnicas *Subdomain enumeratio n*(**T3**) e *Technology Identification* (**T7**), pois são as únicas funcionalidades onde se pode quantificar os resultados encontrados e aplicar testes estatístico. Para validação por completude, elaborou-se um *check list* dos resultados obtidos dos relatórios gerados pelas ferramentas de acordo com as técnicas de coleta de informações do framework ISSAF.

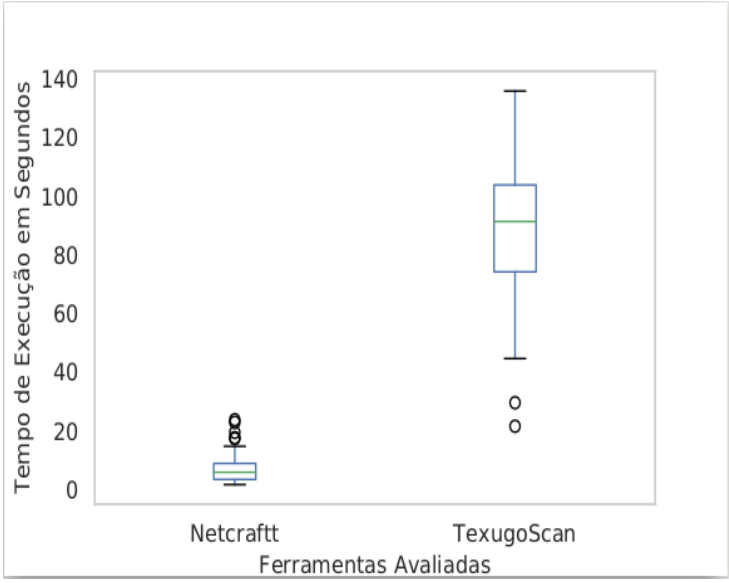
4.1 Reprodutibilidade

Os relatórios gerados pelas ferramentas podem ser encontrados no endereço (https://github.com/andersonvalentim/TexugoScan/tree/master/ProjetoTexugoScan/RelatoriosTcc), os arquivos referentes às análises realizadas podem ser encontrados no seguinte endereço(https://github.com/andersonvalentim/TexugoScan/tree/master/ProjetoTexugoScan/Valida%C3%A7%C3%A3o). Os dados dos relatórios das ferramentas estudadas, foram organizadas em planilhas, visando facilitar a análise, ocorrida no capítulo 5 desse trabalho.

5 RESULTADOS E DISCUSSÃO

Como resultados da medida de tempo de execução das ferramentas avaliadas, os seguintes dados foram encontrados ao avaliar as ferramentas *Netcraft* e *Texugo Scan*, como mostra o Gráfico 1.

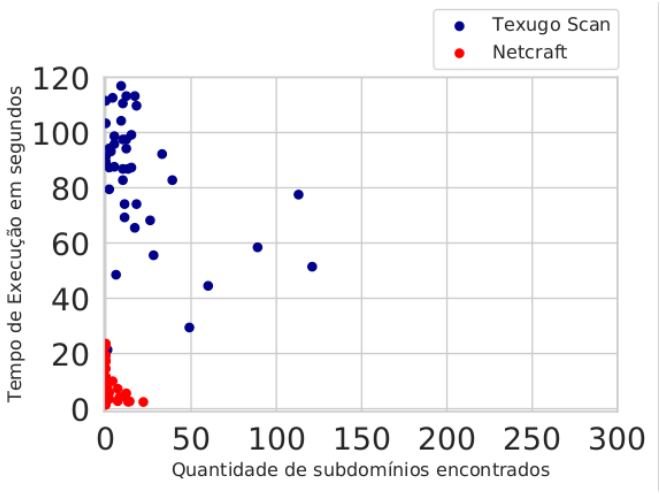
Gráfico 1- Medida de tempo de execução



Fonte: Autoria Própria

Com base nos dados do gráfico, onde que a unidade de medida é dada em segundos, observa-se que as ferramentas apresentam níveis médios de tempo de execução diferentes como mostra o segundo quartil do gráfico em ambas as ferramentas, dessa forma, nota-se que Texugo Scan apresenta um tempo maior de execução comparado ao Netcraft, o valor máximo de execução do Texugo Scan ultrapassa os 100 segundos de execução. Já o *Netcraft*, apresenta menor tempo de execução comparado ao Texugo Scan, contudo, ambas as ferramentas apresentam valores discrepantes (*outliers*), os *outliers* do Netcraft apresentam valores maiores que a média dos tempos de execução, possivelmente ocasionado por uma latência na rede que levou um maior tempo de execução, já os do Texugo Scan apresentam valores menores que a média. A explicação para isso é provavelmente que a dispersão entre quantidade de subdomínios encontrados com o tempo de execução do Texugo Scan, oferece um maior tempo de execução, e uma quantidade maior de detecção de subdomínios em relação ao Netcraft como mostra o Gráfico 2.

Gráfico 2- Dispersão entre tempo e detecção de subdomínios

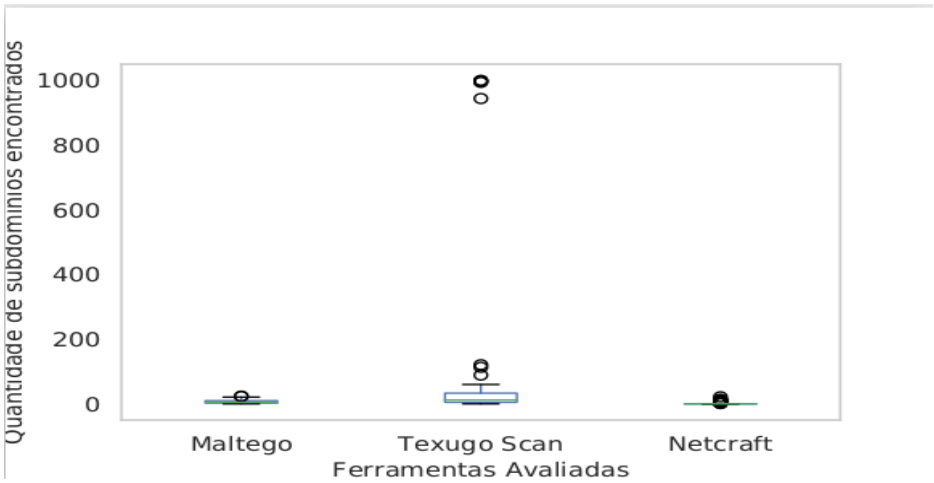


Fonte: Autoria Própria

Com base nos dados expostos no Gráfico 2, onde é mostrado a dispersão entre o tempo de execução em segundos, e a quantidade de subdomínios encontrados, conclui-se que o Netcraft é mais eficiente em tempo de execução, mostrando valores inferiores a 20 segundos, porém, aparentemente processa menos informações comparando a quantidade de domínios encontrados, em relação ao Texugo Scan, utilizando como base de comparação os alvos avaliados, o que justifica o maior tempo de execução do Texugo Scan.

Os resultados da análise quantitativa de dados da técnica Subdomain enumeration(**T3**), foram obtidos por meio da análise dos relatórios produzidos pelas 3 ferramentas comparadas, como mostra o Gráfico 3.

Gráfico 3- Detecção de Subdomínio



Fonte: Autoria Própria.

Com base nos dados mostrados no gráfico, a medida de comparação é dada pelo número de subdomínios encontrados em relação a determinado alvo avaliado, a ferramenta Texugo Scan obtém vantagem diante do Maltego e Netcraft, nota-se que ambas as ferramentas avaliadas apresentam outliers, esses outliers representam uma quantidade maior de subdomínio encontrados em determinado alvo, dessa forma, observa-se que o Texugo Scan apresenta uma detecção superior que a as demais ferramentas testadas, a Tabela 5 mostra a média de detecção de subdomínios, como também a variância de acordo com as ferramentas testadas.

Tabela 5- Média de detecção de Subdomínios

Ferramenta	**Média**
Texugo Scan	120,52
Maltego	6,88
Netcraft	2,019

Fonte: Autoria Própria.

Observando os dados da tabela, conclui-se que a ferramenta Texugo Scan possui maior capacidade de detecção de subdomínios, contudo, a média por sí só não é capaz de provar a eficiência das ferramentas em questão, para isso, aplicou-se o teste T independente (Gosset, 1904) com finalidade de comprovar a hipótese que duas amostras independentes têm valores médios (esperados) idênticos. Como base para análise e levantamento de hipótese foram os dados de subdomínios obtidos pelas ferramentas *Netcraft* e Maltego em comparação com a ferramenta Texugo Scan. Obteve-se os seguintes valores bicaudal(*p*) como mostra a Tabela 6.

Tabela 6- Aplicação do teste T independente nas ferramentas

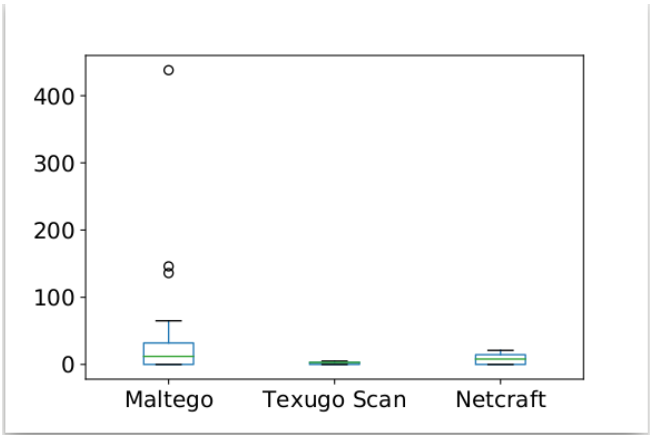
Ferramenta 1	**Ferramenta 2**	**Valor Bicaudal(*p*)**
Texugo Scan	Maltego	0.011538830946049792
Texugo Scan	Netcraft	0.008620018735277249

Fonte: Autoria Própria.

Dessa forma, pode-se analisar que o valor de *p* é menor que os valores 0,05 ou 0,1, logo descartamos a hipótese de que as médias são idênticas, portanto, as médias encontradas pelo Texugo Scan são verdadeiramente superior ao das demais ferramentas avaliadas.

Os resultados da análise quantitativa de dados da técnica *Technology Identification*(**T7**), foram obtidos através da análise dos relatórios produzidos pelas 3 ferramentas comparadas, como mostra o Gráfico 4.

Gráfico 4- Detecção de Tecnologias



Fonte: Autoria Própria.

Analisando os dados mostrados no gráfico, é notório a superioridade da ferramenta Maltego em relação aos seus concorrentes, visualmente, nota-se que sua média de detecção é bem maior que a média das demais ferramentas avaliadas, como média, tivemos os seguintes resultados como mostra a Tabela 7.

Tabela 7- Média de detecção de Tecnologias

Ferramenta	**Média**
Texugo Scan	2.0588235294117645
Maltego	28.392156862745097
Netcraft	7.803921568627451

Fonte: Autoria Própria.

Aplicando o teste T independente nos resultados obtidos na detecção de tecnologias das ferramentas avaliadas, os seguintes valores bicaudais(*p*) foram encontrados como mostra a Tabela 8.

Tabela 8- Aplicação do teste T independente nas ferramentas

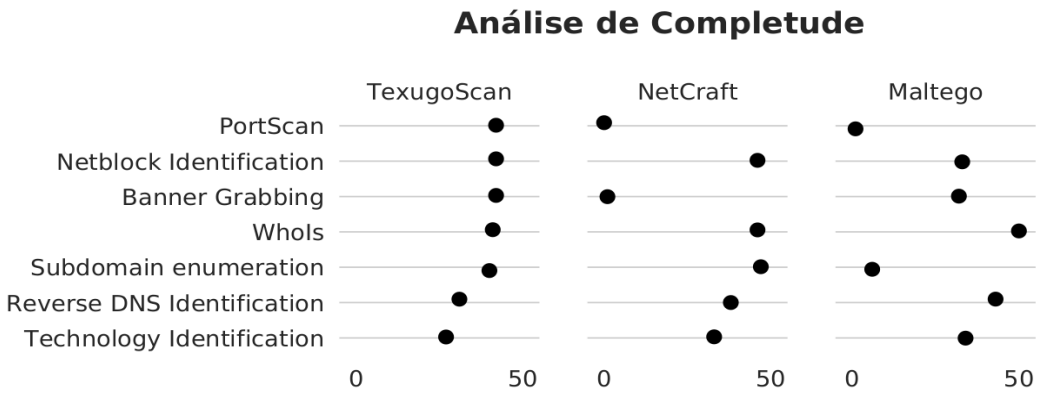
Ferramenta 1	**Ferramenta 2**	**Valor Bicaudal(*p*)**
Texugo Scan	Maltego	0.00593655999431483
Texugo Scan	Netcraft	6.912880533099614e-07

Fonte: Autoria Própria.

Com base nos valores de p encontrados, nota-se que a ferramenta Texugo Scan é verdadeiramente inferior na identificação de tecnologias comparadas as ferramentas Maltego e Netcraft.

A partir dos dados de completude coletados, onde foram avaliados os retornos das técnicas de coleta de informações, das ferramentas avaliadas em questão, o Gráfico 5 mostra a quantidade de técnicas em que houve retorno de acordo com o número de alvos avaliados.

Gráfico 5- Completude das Ferramentas



Fonte: Autoria Própria.

Com base nos dados expostos, nota-se que a ferramenta *Netcraft*, obtém maior completude de retornos nas técnicas **Whois, Reverse DNS Identification, e Technology Identification**, a ferramenta Maltego, obtém maior completude de retorno nas técnicas **Subdomain enumeration, Netblock Identification.** Já a ferramenta Texugo Scan, obteve maior retorno nas técnicas **Port Scan**, e **Banner Grabbing**.

6 CONSIDERAÇÕES FINAIS

Neste trabalho buscou-se conceber uma solução de software para coleta de informações para testes de intrusão. O intuito maior deste trabalho foi o de colaborar para a comunidade de segurança da informação, disponibilizando uma ferramenta para coleta de informações tão eficaz quanto as ferramentas disponíveis no mercado. Além disso foi publicado um artigo científico intitulado: "**TEXUGO SCAN: Coletor de Informações Livre para Testes de Intrusão**" no **Encontro de Computação do Oeste Potiguar (ECOP)** no ano de 2019, onde neste artigo foi mostrado os processos iniciais para elaboração da ferramenta proposta neste trabalho. Grande parte da ferramenta foi desenvolvida nesta etapa, mas ainda existem ajustes a serem feitos para que a ferramenta esteja aprimorada e atenda todos os requisitos de forma igualitária. Todos os códigos desenvolvidos estão disponíveis de forma livre no endereço (https://github.com/andersonvalentim/TexugoScan), para a comunidade acadêmica, e está passível a modificação e implementação de novas funcionalidades.

6.1 TRABALHOS FUTUROS

Como trabalhos futuros pretende-se:

- Incorporar uma funcionalidade para a análise de vulnerabilidades dos alvos testados;
- Realizar melhorias na interface do relatório gerado pela ferramenta;
- Analisar os relatórios emitidos pela ferramenta Texugo Scan;
- Publicar artigo referente às análises efetuadas dos relatórios em eventos científicos;
- Submeter a ferramenta para a seleção como ferramenta padrão do Parrot OS.

REFERÊNCIAS

ASTAH. Astah - Software Design Tools for Agile teams with UML, ER Diagram, Flowchart, Mindmap and More | Astah.net. [S. l.], 2019. Disponível em: http://astah.net/. Acesso em: 6 mar. 2019.

ALEXA. Alexa: Top Sites in Brazil. 2019. Disponível em: <https://www.alexa.com/topsites/countries/BR>. Acesso em: 04 ago. 2019.

BALSAMIQ. Balsamiq. Rapid, effective and fun wireframing software. | Balsamiq. [S. l.], 2019. Disponível em: https://balsamiq.com. Acesso em: 6 mar. 2019.

BOOTSTRAP. Bootstrap · The most popular HTML, CSS, and JS library in the world. [S. l.], 2019. Disponível em: https://getbootstrap.com/. Acesso em: 6 mar. 2019.

BERTOGLIO , D. D; ZORZO, A. F. Tramonto: Uma Estratégia de Recomendação para Testes de Penetração. *In*: XVI SIMPÓSIO BRASILEIRO EM SEGURANÇA DA INFORMAÇÃO E DE SISTEMAS COMPUTACIONAIS-SBSEG, 2016, Porto Alegre. Anais [...]. RS-Brasil: [*s. n.*], 2016. Disponível em: http://sbseg2016.ic.uff.br/pt/files/anais/completos/ST7-3.pdf. Acesso em: 11 fev. 2019.

BERTOGLIO, D. D ; ZORZO, A. F. Um Mapeamento Sistemático sobre Testes de Penetração. 2015. Dissertação (Mestrado) - PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL FACULDADE DE INFORMÁTICA PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO, [S. l.], 2015. Disponível em: www.pucrs.br/facin-prov/wp-content/uploads/sites/19/2016/03/tr084.pdf. Acesso em: 11 fev. 2019.

DANTAS, M. L. (2011). Segurança da informação: uma abordagem focada em gestão de riscos. Recife: Livro Rápido-Elógica.

MACÊDO, Diego. Coleta de informações em um Pentest. In: Coleta de informações em um Pentest. SL, 29 jun. 2016. Disponível em: https://www.diegomacedo.com.br/coleta-de-informacoes-em-um-pentest/. Acesso em: 6 mar. 2019.

ERNST & YOUNG. Retomada da segurança cibernética: como se preparar para enfrentar os ataques cibernéticos. [S. l.], 2017. Disponível em: https://www.ey.com/Publication/vwLUAssets/EY-GISS-2017/%24File/GISS-2017-Port.pdf.
Acesso em: 13 abr. 2019.

FLASK. Flask | The Pallets Projects. [S. l.], 2019. Disponível em: https://flask.palletsprojects.com/en/1.1.x/. Acesso em: 4 ago. 2019.

GARTNER. Gastos com segurança da informação devem ter alta de 12% em 2018. [S. l.], 20 ago. 2018. Disponível em: https://www.gartner.com/newsroom/id/3836563. Acesso em: 13 mar. 2019.

GOSSET, William S. [Student]. 1904. "The Application of the 'Law of Error' to the Work of the

Brewery." Laboratory Report, Vol. 8, Nov. 3, Arthur Guinness, Son & Co., Ltd., pp. 3–16 and unnumbered appendix.

GITHUB. GitHub is where people build software. More than 36 million people use GitHub to discover, fork, and contribute to over 100 million projects. [S. l.], 2019. Disponível em: https://github.com/. Acesso em: 6 mar. 2019.

INFOSEC INSTITUTE. Information Gathering [Updated 2019]. [S. l.], 6 maio 2019. Disponível em: https://resources.infosecinstitute.com/information-gathering/. Acesso em: 11 ago. 2019.

KISSEL, R. L., STINE, K. M., SCHOLL, M. A., ROSSMAN, H., Fahlsing, J., and Gulick, J. (2008). *Security considerations in the system development life cycle. Technical report.*

MCGRAW, G. (2006). *Software security: building security in, volume 1. Addison-Wesley Professional*.

MIT. The MIT License | Open Source Initiative. [S. l.], 2019. Disponível em: https://opensource.org/licenses/MIT. Acesso em: 6 mar. 2019.

NETCRAFT. Internet Security and Data Mining. [S. l.], 2019. Disponível em: https://www.netcraft.com/. Acesso em: 6 mar. 2019.

OFFENSIVE SECURITY. Kali Linux | Penetration Testing and Ethical Hacking Linux Distribution. [S. l.], 2019. Disponível em: https://www.kali.org/. Acesso em: 11 ago. 2019.

PARROT SEC. Parrot OS - The advanced system for security experts, developers and crypto-addicted people. [S. l.], 6 maio 2019. Disponível em: https://www.parrotsec.org/. Acesso em: 11 ago. 2019.

PATERVA. Maltego offers the user unprecedented information. Information is leverage. Information is power. Information is Maltego. [S. l.], 6 mar. 2019. Disponível em: https://www.paterva.com/buy/maltego-clients/maltego.php. Acesso em: 6 mar. 2019.

PRANDINI, M. R. (2010). Towards a practical and effective security testing methodology. The IEEE symposium on Computers and Communications.

POTTER, B.; MCGRAW, G. *"Software security testing," in IEEE Security & Privacy, vol*. 2, no. 5, pp. 81-85, Sept.-Oct. 2004, doi: 10.1109/MSP.2004.84.

RATHORE, B., BRUMER, M., DILAJ, M., HERRERA, O., BRUNATI, P., SUBRAMANIAM, R., Raman,S., and CHAVAN, U. (2006). *Information systems security assessment framework* (issaf).Draft 0.2.1, Open Information Systems Security Group.

SOARES, R. Auditoria Teste de Invasão (Pentest) – Planejamento, Preparação e Execução. [S.l.], 7 set. 2010. Disponível em: https://seginfo.com.br/2010/09/07/auditoria-teste-de-invasaopentest-planejamento-preparacao-e-execucao-2/. Acesso em: 11 fev. 2019.

SOMMERVILLE, Ian. Engenharia de Software. 8ª ed. São Paulo: Pearson Addison-Wesley, 2007.

WE ARE SOCIAL. DIGITAL IN 2018: WORLD'S INTERNET USERS PASS THE 4 BILLION MARK. [S. l.], 30 jan. 2018. Disponível em: https://wearesocial.com/blog/2018/01/global-digital-report-2018. Acesso em: 4 ago. 2019.

VISUAL STUDIO CODE. Visual Studio Code - Code Editing. Redefined. [S. l.], 2019. Disponível em: https://code.visualstudio.com/. Acesso em: 6 mar. 2019.

WEIDMAN, G. Teste de Invasão: Uma introdução prática ao hacking, 1. ed. São Paulo: Editora Novatec,2014.575 p.