INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS TERESINA CENTRAL
TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

RAFAEL DA SILVA ARAÚJO

ISA – SISTEMA PARA ANÁLISE DE VULNERABILIDADES NA WEB

TERESINA

2022

ISA – SISTEMA PARA ANÁLISE DE VULNERABILIDADES NA WEB

Trabalho de Conclusão de Curso (monografia) apresentado como exigência parcial para obtenção do diploma do Curso de Superior de Tecnologia em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central.

Orientadora: Prof. Ma^{e}. Cláutenis Carvalho Viana

TERESINA

2022

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD

Araújo, Rafael da Silva

A663i ISA : sistema para análise de vulnerabilidade na web / Rafael da Silva Araújo. - 2022.

30 f.: il. color.

Trabalho de Conclusão de Curso (Tecnologia em Análise e Desenvolvimento de Sistemas) - Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central, 2022.

Orientador : Prof Me. Cláutenis Carvalho Viana.

1. Segurança da informação. 2. Vulnerabilidades. 3. Varredura. 4. Aplicação web. I.Título.

CDD - 004.21

Elaborado por Maria Rosismar Farias CRB 3/631

RAFAEL DA SILVA ARAÚJO

ISA – SISTEMA PARA ANÁLISE DE VULNERABILIDADES NA WEB

Trabalho de Conclusão de Curso (monografia) apresentado como exigência parcial para obtenção do diploma do Curso de Superior de Tecnologia em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Teresina Central.

Aprovado pela banca examinadora em 16/12/2022 às 16 horas.

BANCA EXAMINADORA:

Prof. Ma^{e}. Cláutenis Carvalho Viana
Instituto Federal de Educação, Ciência e Tecnologia do Piauí - IFPI

Prof. M^{e}. Fernando Castelo Branco Gonçalves Santana
Instituto Federal de Educação, Ciência e Tecnologia do Piauí - IFPI

Prof. Dra. Nádia Mendes dos Santos
Instituto Federal de Educação, Ciência e Tecnologia do Piauí - IFPI

TERESINA

2022

Dedico este trabalho a todos que de alguma forma me ajudaram.

AGRADECIMENTOS

Agradeço primeiramente a Deus, sem Ele não teria chegado tão longe.

Agradeço aos meus pais, que sempre investiram e me apoiaram nos meus estudos, sou muito grato por deixá-los orgulhosos.

Agradeço aos meus amigos pelo incentivo e orações, — Em especial André Rabelo, Evenlly Cunha, Isabella Araújo, Cíntia Maria, Maria Clara. Que nunca deixaram eu me desanimar.

Agradeço muito pelos professores incríveis que eu tive durante todo curso, que compartilharam suas experiências e conhecimentos que fizeram eu me sentir mais atraído pela área da tecnologia.

Finalmente agradeço à minha orientadora, Cláutenis Viana, pela dedicação, apoio, orações, direcionamento, disponibilidade e excelentes conselhos que tornaram esse trabalho possível.

RESUMO

A segurança da informação vem ganhando cada vez mais importância e espaço nessa era digital. Tudo isso, deve-se a grande expansão dos meios de tecnologia pelo mundo, tais como, celulares, notebooks, desktops, dentre outros. Tal fenômeno influenciou diretamente no aumento do número de pessoas que estão conectadas à internet. Contudo, isso também significa dizer mais dados de usuários disponíveis na rede e que estão a todo instante sob risco de serem roubados. Sujeitos mal intencionados, usando-se do anonimato e da distância que a Internet lhe propícia, tentam a todo instante prejudicar outras pessoas que estão conectadas à rede através de ataques. Podendo esses ocorrerem de várias formas, por exemplo, conseguindo acesso a dados sigilosos, sendo capaz de destruir, modificar ou até vender a informação ali contida, invadindo sites, realizando espionagem cibernética, ou ainda, crackeando algum serviço disponível na internet. No entanto, têm-se desenvolvido muitas formas de diminuir esses ataques hackers, as quais é importante destacar: ter conhecimento de como ocorre uma invasão e quais as vulnerabilidades mais comuns encontradas nas páginas web. Dessa forma, com o intuito de saber quais práticas devem ser evitadas ao, por exemplo, informar e armazenar algum dado, mantendo assim o sigilo e a seguranças de seus dados na web. Assim sendo, este trabalho apresenta uma ferramenta que se propõe a auxiliar, principalmente, pessoas da área de TI no que se refere à segurança de dados na internet. A aplicação em questão, é uma ferramenta web que faz a varredura e descreve eventuais vulnerabilidades encontradas nas páginas web verificadas.

Palavras-chaves: segurança da informação; vulnerabilidades; varredura; aplicação web; ataques.

ABSTRACT

Information security is gaining more and more importance and space in this digital age. All of this is due to the great expansion of technology around the world, such as cell phones, notebooks, desktops, among others. This phenomenon directly influenced the increase in the number of people who are connected to the internet. However, this also means that more user data is available on the network and is at risk of being stolen at all times. Malicious individuals, using the anonymity and distance that the Internet provides, try at all times to modify other people who are connected to the network through attacks. These can occur in several ways, for example, preventing access to confidential data, being able to destroy, modify or even sell the information contained therein, invading websites, performing cyber espionage, or even cracking some service available on the internet. However, many ways have been developed to reduce these hacker attacks, which is important to highlight: having knowledge of how an invasion occurs and what are the most common vulnerabilities found on web pages. In this way, in order to know which practices should be avoided when, for example, informing and storing some data, thus maintaining the secrecy and security of your data on the web. Therefore, this work presents a tool that proposes to help, mainly, people in the IT area with regard to data security on the internet. The application in question is a web tool that scans and describes any vulnerabilities found in the verified web pages.

Key-words: information security; vulnerabilities; scan application web; attacks.

LISTA DE ABREVIATURAS E SIGLAS

ISA	Integrity Security Availability
URL	Uniform Resource Locator
IP	Internet Protocol
SQL	Standard Query Language
XSS	Cross-site scripting
NMAP	Network Mapper
TCP	Transfer Control Protocol
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure

SUMÁRIO

1 INTRODUÇÃO

Nesse capítulo será apresentado a contextualização sobre a segurança da informação na web: conceito, características, aspectos legais no Brasil e o atual cenário dos crimes cibernéticos. Depois será apresentado uma justificativa sobre a necessidade da ferramenta ISA.

1.1 CONTEXTUALIZAÇÃO

Atualmente, a procura por soluções web tem tido um crescimento considerável, devido ao enfrentamento da COVID-19, fazendo muitas empresas e instituições migrarem seus negócios para a internet. Dessa forma todo tipo de informações como identificação, autorização e outras, se transformou num elemento essencial para o mercado, para que os usuários desse meio consigam se mover, com autonomia e liberdade, nos corredores denominados sociedade da informação[1].

No entanto, com esse crescimento de aplicações web, houve um aumento de ataques hackers, no primeiro semestre de 2021 os crimes cibernéticos cresceram 220% em comparação ao mesmo período de 2020 no Brasil segundo A MZ[2]. Com a ansiedade de migrar para o remoto durante a pandemia, as empresas não tiveram a devida preocupação com a segurança de dados, notícias como:

- Um vazamento de dados do Facebook atingiu 8 milhões de brasileiros de um total de 533 milhões de contas da rede social em todo o mundo;
- Uma subsidiária da JBS nos Estados Unidos informou que pagou U$ 11 milhões em resgate após um ransomware (sequestro digital);
- O site oficial da Lojas Renner sofreu um ransomware;
- Os sistemas da agência de turismo CVC foram sequestrados e deixou passageiros com dificuldades no momento da viagem;
- O IFood sofreu um ataque que desfigurou o nome dos estabelecimentos comerciais cadastrados;
- Em uma série de ataques ocorridos a partir de dezembro segue em investigação, criminosos invadiram sites do Sistema Único de Saúde (SUS), Conecte SUS (respon- sável pelos dados de vacinação da população brasileira), Polícia Rodoviária Federal, Ministério da Economia, Controladoria Geral da União (CGU) e Instituto Federal do Paraná.

Tais crimes, que terminam por prejudicar milhões de pessoas, têm tornado-se cada vez mais frequentes no cotidiano. A Surfshark, uma empresa holandesa de cibersegurança, estima que 1 em cada 5 pessoas em todo o mundo tiveram dados vazados em 2021. Só no Brasil ainda

de acordo com a Surfshark, foram 24,2 milhões de usuários expostos entre Janeiro e Novembro de 2021[3].

Em razão disso, a Segurança da Informação torna-se vital para diminuir a incidência de crimes dessa natureza, uma vez que, ela está relacionada a toda forma de proteção de dados, sendo que esses dados podem ser considerados qualquer informação de uma pessoa ou de uma empresa.

Os principais aspectos da segurança de dados são: confidencialidade, integridade e disponibilidade. A confidencialidade está relacionada à disponibilidade da informação somente a quem esteja autorizado pelo proprietário dela. Por sua vez, a integridade é a garantia de que a informação não foi alterada por terceiros. Por fim, a disponibilidade garante para o proprietário que a informação sempre vai estar disponível para ser acessada.

Ademais, para manter esses três pilares da segurança de dados incorruptíveis, existe um profissional especializado chamado de pentester que faz o teste de intrusão *(pentest)*, fazendo uma varredura em um determinado sistema de um computador ou de uma rede a procura de bugs e falhas, e assim, reporta todas as vulnerabilidades encontradas ao contrante[4].

No Brasil temos a *Lei 12.737/2012* que dispõe a tipificação dos crimes cibernéticos, Art. 154-A. Invadir dispositivo informático alheio, conectado ou não à rede de computadores, mediante violação indevida de mecanismos de segurança a fim de obter, adulterar ou destruir dados ou informações sem autorização expressa, ou tácita do titular do dispositivo, ou instalar vulnerabilidades para obter vantagem ilícita[5]. Nesse sentido, para ser considerado crime a invasão deve ocorrer sem o consentimento do titular da informação, porém, caso haja uma autorização para o *pentester*, não serão considerados crimes nenhuma de suas ações ao tentar descobrir vulnerabilidades.

Atualmente, há muitas falhas que podem ser exploradas por criminosos, falhas essas que poderiam ser evitadas caso fossem tomadas as devidas medidas de segurança. Portanto, um dos principais motivos para tantos ataques hackers é a falta de qualificação dos desenvolvedores de software, pois não dão a atenção necessária à segurança e acabam deixando vulnerabilidades em seus sistemas desenvolvidos. Outro motivo, é a escassez de profissionais especializados, a falta de mão de obra para segurança da informação e cibersegurança ocorreu em 70% das empresas, indica uma pesquisa publicada pelo Enterprise Strategy Group (ESG) e pela Information Systems Security Association (ISSA)[6].

Nessa perspectiva, considerando todos os motivos citados, surgiu a ideia da ISA, que atuará como uma ferramenta de varredura a procura das principais e mais conhecidas vulnerabilidades em um site na web.

1.2 JUSTIFICATIVA

Na atualidade, existem disponíveis no mercado softwares que procuram falhas de forma automatizada em diversos tipos de aplicação, tais como, Nessus[1], Core Impact [2], Qualys Guard[3], entre outros. No entanto, essas ferramentas possuem um custo oneroso, dessa forma, grande parte dos proprietários de domínios de sites empresariais, seja a empresa de pequeno ou médio porte, não podem arcar com esse valor.

Ademais, cabe salientar que os softwares citados anteriormente são complexos, pois a utilização desses programas exige do usuário uma base considerável de conhecimento sobre segurança da informação. Em vista disso, a ferramenta web desenvolvida nesse trabalho auxilia os desenvolvedores de websites, com o objetivo de procurar vulnerabilidades em seus projetos desenvolvidos.

Por fim, é válido ressaltar que a ferramenta web tem características como usabilidade, baixo custo e simplicidade, pois caso tenha sido detectado alguma falha no sistema web analisado, o usuário da ferramenta terá um feedback explicando de uma forma simples do que se trata a vulnerabilidade, causas, consequências e como o sistema chegou naquele resultado. Dessa forma, incentivando os utilizadores da solução web a buscar mais conhecimento na área de segurança da informação.

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Desenvolver uma ferramenta web que faça a análise de um site x - podendo esse ser qualquer site da internet - em busca de vulnerabilidades, tais como, falhas ou dados sensíveis expostos que porventura deixem o site indicado vulnerável, possibilitando ainda, prestar auxílio a profissionais e estudantes de segurança na internet.

1.3.2 Objetivos Específicos

- Procurar vulnerabilidades em uma determinada URL;
- Indicar sites que comprometam de alguma forma a segurança da URL pesquisada, especificando a causa e as consequências das falhas encontrada;
- Incentivar usuários da ferramenta a se aprofundar na área de segurança.

[1] https://www.tenable.com/products/nessus
[2] https://www.coresecurity.com/products/core-impact/pricing-plans
[3] https://www.qualys.com/apps/vulnerability-management-detection-response/

2 FUNDAMENTAÇÃO TEÓRICA

2.1 ANÁLISE DE VULNERABILIDADE

Devido ao crescimento dos crimes cibernéticos, tomar medidas de proteção contra ataques hackers se tornou uma necessidade do mundo atual(digital). Filho, Carvalho, Melo(2013) falam sobre a importância da segurança da informação e o quanto ela estando ausente no meio pode ser prejudicial para a pessoa ou empresa[7]. Dessa forma, organizações, que não se preocupam em armazenar de forma segura suas informações, se tornam um alvo fácil para que seus dados sejam o roubados e utilizados por terceiros. Monteverde(2014) dar ênfase nas vantagens de conhecer as vulnerabilidades, pois identificar as principais vulnerabilidades Web permite estabelecer contramedidas, assim aumentando a segurança das aplicações[8].

2.2 PROCESSO DE INVASÃO

Fraga[9] define que o processo de invasão de um sistema são dividido em três etapas:

Conhecer: que resumidamente é coletar informações do alvo a ser invadido, de todas as maneiras possíveis;

Analisar: usar os dados coletados para extrair o máximo de informações possíveis sobre o alvo. Ele reforça dizendo que essa etapa é principal para uma invasão bem sucedida;

Explorar: explorar o que foi analisado tentando acessar, quebrar, atacar.

2.3 WEB

2.3.1 HTTP (*Hypertext Transfer Protocol*)

O HTTP é um protocolo de comunicação, utilizado para sistemas de informação de hipermídia, distribuídos e colaborativos. descreve um protocolo sem estado com o qual os dados podem ser transmitidos em uma rede IP[9]. A aplicação mais importante é a transferência de páginas da Internet e dados entre um servidor web e um navegador. No entanto, ele não está restrito a sites, sendo usado para aplicativos[10].

HTTP é um protocolo baseado em texto, ou seja, toda a informação transmitida está em texto, os dados do usuário e do servidor podem ser interceptados ou alterados no meio do caminho. Dessa forma, o usuário na rede pode interceptar os seus dados e lê-los e alterar a página que você recebe ou a informação que envia para o servidor. Com o uso do HTTPS, que é o HTTP seguro, adiciona-se alguns princípios de segurança, como confidencialidade, integridade e autenticação. Por confidencialidade, entende-se que a mensagem só é lida pelo destinatário

real da mensagem. A integridade representa que a mensagem não foi alterada e o princípio da autenticação prova que o servidor é realmente quem diz ser[11].

2.3.2 URL (*Uniform Resource Locator*)

A URL é o endereço web, quando se deseja acessar uma página na internet colocamos esse endereço na barra de pesquisa do navegador. Esse endereço contém informações específicas, que seguem um padrão pré-determinado para que o usuário possa sempre encontrar o serviço que procura, desde que ele digite o endereço corretamente A URL é dividida basicamente entre duas partes, para exemplificar será utilizado a seguinte URL: *https://www.google.com*. A primeira parte é o esquema que é um protocolo de rede que fica antes dos ":" na parte inicial do endereço, podendo indicar endereços na web (http, https), comunicação via e-mail (mailto), transferência de arquivos entre computadores (ftp), comunicação via chats (irc) etc.

A segunda parte é o caminho, que é dividido em pelo menos três partes.

Hostname: primeira parte antes do primeiro ponto como www;

Domínio do site: após o primeiro ponto como google;

Domínio de alto nível: que é o último termo após o último ponto e antes da primeira barra e indica o tipo do site (com, net, org. . .) [12].

2.3.3 IP

IP*(Internet Protocol)* é um endereço exclusivo que identifica um dispositivo na Internet ou em uma rede local. Atualmente, existem duas versões do protocolo IP que são: o IPv4 e o IPv6. O endereço IPv4 é uma sequência de 4 números separados por pontos, por exemplo, 192.158.1.38. Cada número do conjunto pode variar entre 0 e 255, ou seja, o intervalo de endereçamento IP vai de 0.0.0.0 a 255.255.255.255, dessa forma, a versão mais utilizada do protocolo IP é o IPv4[13].

Contudo, com aumento dos dispositivos na web foram se esgotando os endereços IPv4, pois há um limite teórico de 4,3 bilhões IP's. Assim sendo, houve a necessidade de criar o endereço IPv6, que suporta mais endereços com capacidade de atender a demanda do aumento de aparelhos conectados à Internet, tendo 340 undecilhões de limite[14].

Outrossim, por ser uma combinação bem maior que inclui não somente números, mas também letras, torna o processo de descobrir o endereço IP na base da chamada "força bruta"bem mais oneroso, e até mesmo utilizando ferramentas combinatórias o processo ainda seria árduo e custativo. Agregando assim, mais segurança a quem utiliza essa versão do endereço IP.

2.4 GOOGLE HACKING

Uma das ferramentas mais poderosas para o descobrimento de vulnerabilidades é o Google hacking. Cortes, Dantas(2015) dizem que encontrar informações sensíveis faz parte da rotina de um Google Hacker, que pode utilizar o Google na busca de servidores negligenciados, diretórios expostos, relatórios de segurança expostos e na busca de informações pessoais e documentos compartilhados por engano na Internet como: planilhas, tabelas, vídeos, documentos do Word, fotos, bancos de dados e outros arquivos que possuam alguma informação relevante[15].

2.5 VULNERABILIDADES

Segundo a ISO27000 vulnerabilidade é uma fraqueza de um ativo que poderia ser potencialmente explorada por uma ou mais ameaças, por exemplo, firewalls mal configurados, exposição de informações, senhas fracas, falha na validação de dados, entre outros [16]. Para um ativo se tornar vulnerável é necessário que ele perca algum desses atributos: confidencialidade, integridade e disponibilidade.[9]

2.5.1 SQL Injection

O SQL Injection é uma ameaça de segurança que se aproveita de falhas em sistemas que interagem com base de dados via SQL*(Standard Query Language)*, ocorre quando o atacante consegue inserir uma série de instruções SQL dentro de uma consulta através da manipulação das entradas de dados de uma aplicação[9].

2.5.2 O ataque XSS

O ataque XSS*(Cross-site scripting)* é uma falha explorada por causa de erros na validação dos parâmetros de entrada do usuário do servidor na aplicação web. Os criminosos cibernéticos injetam scripts maliciosos no conteúdo do site alvo, que é então incluído com o conteúdo dinâmico entregue ao navegador da vítima. O navegador da vítima não tem como saber que os scripts maliciosos não são confiáveis e, portanto, os executa[17].

2.5.3 Exploit

Exploit é um software malicioso(malware) bastante utilizado pelos hackers, esse software possui uma sequência de instruções ou dados, cujo objetivo específico é se aproveitar de vulnerabilidades dos sistemas. Assim sendo, esse software pode, por exemplo, entrar em um computador e provocar uma instabilidade no sistema, executando comandos pra roubar informações, ou bloquear acesso, ou enviar vírus[18].

3 METODOLOGIA

Este trabalho propõe o desenvolvimento de uma ferramenta web utilizando o método ágil Scrum Solo, que utiliza princípios utilizados no Scrum para programadores que desenvolvem aplicações web sozinhos. No desenvolvimento da aplicação será utilizado materiais de apoio como livros, cursos, e artigos.

3.1 CASOS DE USO

Caso de uso 01 - Solicitar Análise de URL: caso de uso onde o usuário irá informar a URL que ele deseja analisar;

Caso de uso 02 - Realizar Pesquisa sobre sites já verificados: caso de uso onde o usuário pode ver as vulnerabilidades em sites já pesquisado.

As especificações dos casos de uso estão disponíveis no Anexo A - Especificação de Casos de Uso.

3.2 IMPLEMENTAÇÃO

Segue abaixo a descrição das tecnologias que foram utilizada na implementação e o motivo da escolha:

Python: é uma linguagem de alto nível de uso geral. Seu uso se deu, entre outros motivos, por ser uma das linguagens mais usadas para hacking. Além do que, grande parte das aplicações de terceiros que foram consumidas na ferramenta web, são desenvolvida em python, possuir várias bibliotecas que ajudou no processo de desenvolvimento da ferramenta, tanto na estrutura do código como na localização de erros;

Django: É um framework para desenvolvimento web feito em python, com objetivo de facilitar a criação de aplicações web, com módulos de segurança já implementados, como controle de usuários. E facilita o processo de conexão com banco de dados e a integração de aplicações de terceiros. O Django teve uma grande importância na implementação da ferramenta ISA, pois foi capaz de integrar os scripts de buscas com os modelos e a lógica do sistema para gerenciar os templates e a persistência dos dados.

SQLite: Sistema Gerenciador de Bancos de Dados escolhido para manter os dados. Inicialmente seria utilizado PostgresSQL, mas por não ter a necessidade de algo tão complexo, optou-se por utilizar o SQLite por nativamente vir integrado com o Django;

HTTP for Humans: Uma biblioteca desenvolvida para Python para fazer requisições de site na web de forma simplificada. Essa biblioteca foi a principal ferramenta utilizada para desenvolver a aplicação, responsável auxiliar na automatização do uso do Google hacking;

BeautifulSoup: É uma biblioteca desenvolvida em python pra analisar páginas HTML, procurando por qualquer informação desejada pelo usuário. essa ferramenta na aplicação foi utilizada para verificar se na página tinha palavras chaves que denunciavam algum tipo de associação ao site alvo ou vulnerabilidade

Devido ao grande número de requisições feitas pela ferramenta, não foi possível integrar ferramentas como NMAP, Thehravester e Pompem, pois além da complexidade de estruturação do ambiente para se obter os insumos que essas aplicações complementares proporcionariam, cada requisição feita durante a varredura realizada em uma página web tem duração média de 15 segundos, e para cada site são realizadas, em média, 250 requisições. Assim sendo, é inviável, nesse primeiro momento, a integração com outras ferramentas.

4 ISA

O sistema ISA é um scanner de vulnerabilidades que realiza uma varredura em determinada página web, procurando por falhas que deixem o site pesquisado vulnerável, dessa forma, auxiliando profissionais da segurança impedir possíveis ataques agindo nas duas primeiras etapas de um ataque hacker. Logo após, o sistema armazenará essas informações no banco de dados, mantendo assim um histórico de todas as varreduras já realizadas. Por fim, a ferramenta retornará para o usuário informações sobre as possíveis vulnerabilidades encontradas, com uma explicação simples e direta do que se trata aquela falha, informando: as causas, as consequências e disponibilizando os sites associado a URL pesquisada

Além disso, a forma simples de explicar as falhas encontradas e a descrição de como o sistema fez para chegar nos resultados apresentados, tem por objetivo estimular o usuário a procurar desenvolver aplicações mais seguras, ou ainda, motivar aquele usuário que não é da área a se interessar por hacking e desmitificar a ideia de complexidade de um ataque hacker

4.1 APRESENTAÇÃO DA APLICAÇÃO

Nessa seção será apresentada de forma ilustrativa o sistema desenvolvido nesse trabalho. Na página inicial temos um campo de busca onde o usuário informar o endereço do site que deseja realizar a varredura, se a URL já tiver sido pesquisada, o sistema redireciona para a página que irá exibir os resultados, caso contrário o usuário terá que esperar a ferramenta fazer todas as requisições.

Figura 1 – Captura de Tela da Página Inicial do Projeto ISA

Fonte: Elaborado pelo Autor

Primeira etapa: Após o usuário solicitar a varredura do site, será feito uma análise no link separando-o em hostname, domínio do site e domínio de alto nível, para que seja adicionado a lista de parâmetros de busca que serão feitos utilizando o Google Hacking. Na lista utilizada estão presentes insumos pra pesquisar arquivos, páginas de login, senhas, usuários, e-mails, entre outros. Para a realização da busca, está sendo utilizado o HTTP for Humans para fazer uma requisição no Buscador do Google utilizando como pesquisa as consultas geradas pela URL pesquisa. Logo após, a ferramenta começa a consumir a biblioteca Beautiful Soup para extrair todos os links resultante da busca.

Segunda etapa: Depois de pesquisar por todos os itens que foram usados no Buscador, gera-se uma lista de todos os sites encontrados, deste modo a ferramenta inicia uma primeira filtragem para ignorar sites que não interessam para a varredura, diminuindo boa parte dos links.

Terceira etapa: agora, com a primeira filtragem, a aplicação realiza outra filtragem, dessa vez separando entre arquivos e sites encontrados, depois os endereços encontrados são separados entre vulnerabilidade de XSS e SQL Injector, o que não se associam com essas falhas, são separado como dados sensíveis.

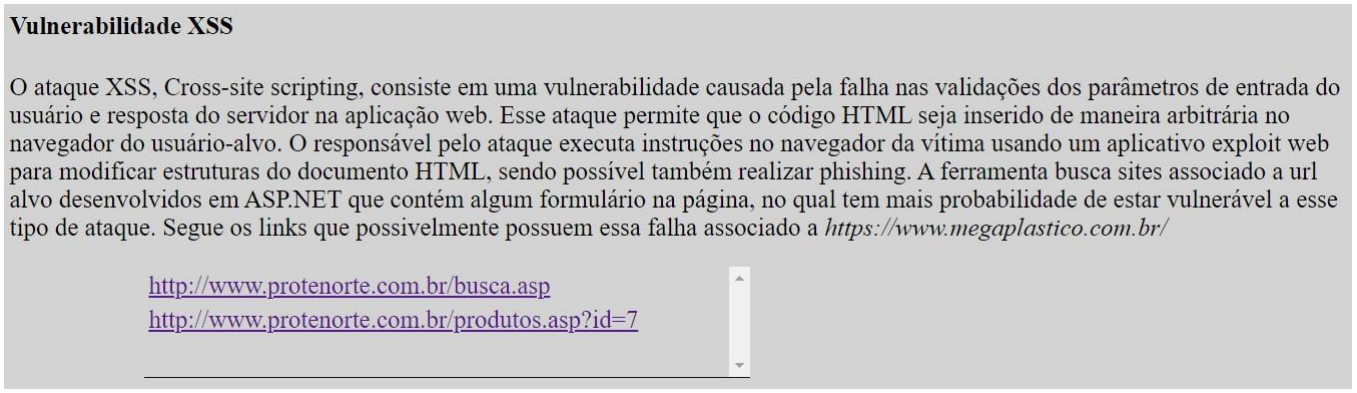
Quarta Etapa etapa: Com os dados sensíveis é feito uma requisição utilizando o HTTP for humans novamente e utilizando o Beautiful Soup pra procurar na pagina elementos que de alguma forma associa esse link ao site alvo pesquisado, caso encontre alguma ocorrência de associação, o link é adicionado ao banco de dados para no fim da varredura ser exposto para o usuário esses sites e os documentos encontrados na análise.

Figura 2 – Captura de Tela do Resultado da Varredura do Projeto ISA

Fonte: Elaborado pelo Autor

Quinta etapa: Com os links da vulnerabilidade XSS, é feito uma análise em cada item. Sites que no final da URL tem ".asp" tem mais probabilidade de ter essa falha, desse modo, a ferramenta busca por endereços como esses e verifica se tem campos de formulário pra se preencher, caso tenha, o link em questão é salvo no Banco de Dados para que ao final do processo seja informado uma lista de sites que tenham possivelmente essa falha.

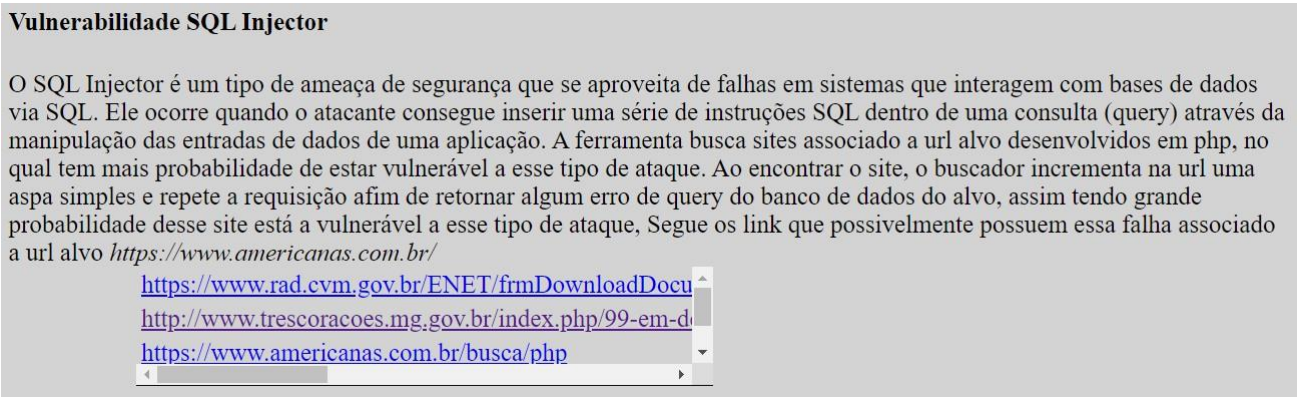
Figura 3 – Captura de Tela da Vulnerabilidade XSS do Projeto ISA



Fonte: Elaborado pelo Autor

Sexta etapa: E, por fim, com as vulnerabilidades de SQL injector, a verredura acontece em sites feitos em "php", no qual, é mais comum essa falha, no processo de análise cada site da lista é adicionado um aspa simples no final da URL, e se faz uma requisição, caso o resultado da requisição retorne um erro de SQL, provavelmente esse link está vulnerável a essa falha, assim então, adicionando ao Banco de Dados, e após a varredura é exposto para o usuário os links associados ao site alvo.

Figura 4 – Captura de Tela da Vulnerabilidade SQL Injector do Projeto ISA

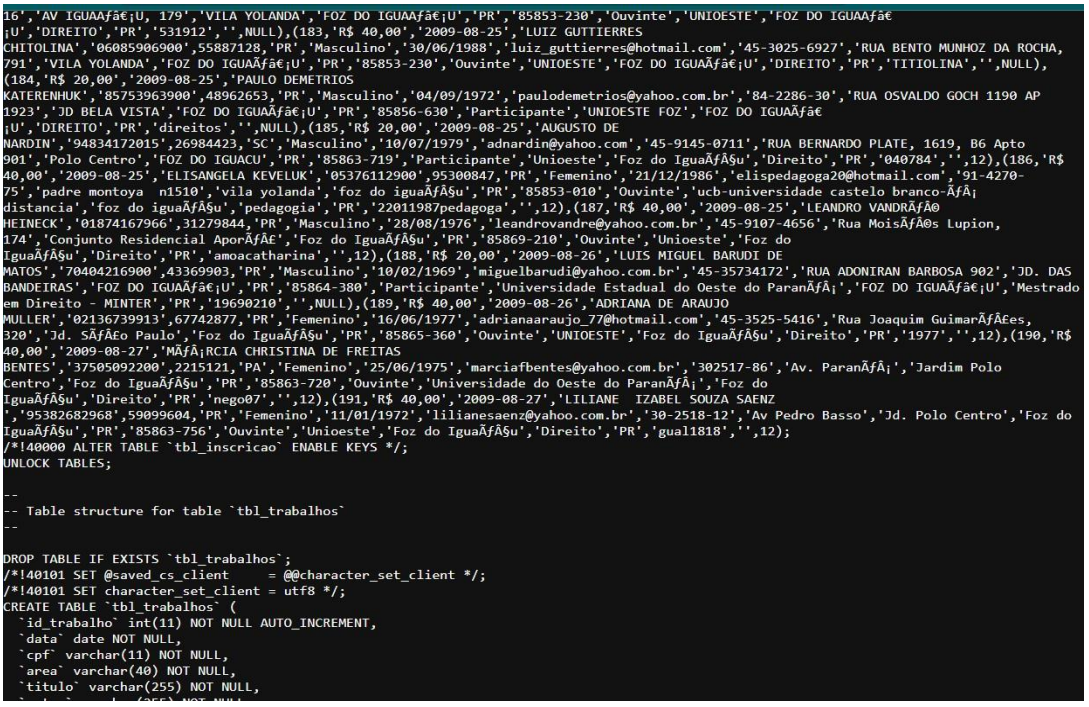


Fonte: Elaborado pelo Autor

4.2 RESULTADOS

Após utilizar a ferramenta desenvolvida neste trabalho, foi possível se obter alguns dados sensíveis que podem ser essenciais pra realizar um ataque, e conhecimento de sites que estão vulneráveis, segue abaixo alguns exemplos:

Figura 5 – Captura de Tela de Dados Sensíveis de um do resultados de um link descoberto pela Ferramenta

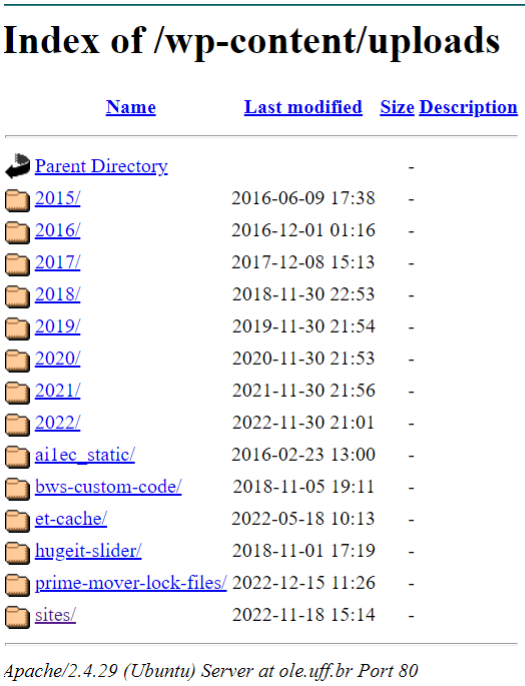


Fonte: Elaborado pelo Autor

Figura 6 – Captura de Tela de Falha de um dos links da Ferramenta

Fonte: Elaborado pelo Autor

Figura 7 – Captura de Tela de arquivos encontrado em um dos resultados da Varredura



Fonte: Elaborado pelo Autor

5 CONCLUSÃO

A aplicação web, ISA, desenvolvida neste trabalho, é uma ferramenta de nível acadêmico muito útil, que pode auxiliar os estudantes nos primeiros passos na segurança da internet, e também, interessar a todos aqueles que desejam adquirir conhecimentos da área de segurança de dados, proporcionando uma boa base para desenvolverem softwares mais seguros, e ainda, sem qualquer custo para utilizar a ferramenta.

Portanto, conclui-se que o trabalho em questão, além de abordar um tema bastante atual para as organizações de modo geral, deixa como resultado uma importante colaboração no apren- dizado dos futuros profissionais de segurança da internet: uma ferramenta que possibilita obter uma boa base de conhecimento na área e exercitar, na prática, a segurança de aplicações web. O código-fonte da ferramenta está disponível no endereço https://github.com/raffaellgates/Project_ISA.

5.1 TRABALHOS FUTUROS

Planeja-se estruturar o ambiente de desenvolvimento para que seja possível a integração de ferramentas como NMAP, Thehravester, Pompem entre outras, com a finalidade de realizar uma varredura que seja mais precisa e assertiva, e consiga expor mais vulnerabilidades. A fim de possibilitar e viabilizar essa atualização, será necessário a utilização de máquinas virtuais, objetivando assim, que elas otimizem todo processo de análise. A próxima versão da ferramenta ISA contará com uma interface mais agradável e robusta, pois o foco maior nesse primeiro momento era encontrar as falhas na web. E por fim, desenvolver uma API com a intenção de tornar mais prático a utilização dos recursos disponíveis na ferramenta por desenvolvedores intermediários e avançados, possibilitando-os usufruir desses recursos em suas aplicações.

6 TRABALHOS RELACIONADOS

Após a pesquisa por trabalhos acadêmicos similares que servissem de auxílio e inspiração para o desenvolvimento do sistema proposto nesse trabalho, foram encontrado alguns que tinham soluções que seguiam o mesmo raciocínio, porém com objetivos diferentes, entre eles os principais serão expostos a seguir.

6.1 HOW I WAS ABLE TO FIND 50+ CROSS-SITE SCRIPTING (XSS) SECURITY VULNERABILITIES ON BUGCROWD PUBLIC PROGRAM?

Trabalho feito por Takshal(2022), que desenvolveu uma ferramenta que conseguiu descobrir mais de 50 vulnerabilidade XSS, o autor conta que utilizou varias ferramentas similares ao funcionamento do Google hacking e conseguiu reunir informações de 578 subdomínios, focou em partes especificas de cada informação adquirida para encontrar todas essas falhas de XSS. [19]

6.2 UTILIZANDO GOOGLE HACKING PARA ENCONTRAR VULNERABILIDADES EM SITES

Trabalho Elaborado por Cortes(2017) que explorou de forma mais abrangente o Google Hacking, o autor também expôs como funcionava a ferramenta e demostrou os resultados que obteve após a utilização do buscador, como informações sigilosas encontradas em documentos, vale ressaltar que no ano que foi feito o trabalho, essa funcionalidade do Google não era tão conhecida, desse modo, conseguido mais informação com nível de sigilo superior. [15]

REFERÊNCIAS

1 BARBOSA, J. S. et al. Data protection and information security in the pandemic covid-19: national context. *Research, Society and Development*, v. 10, n. 2, p. e40510212557, Feb. 2021. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/12557>. Citado na página 10.

2 Cássio Rufino. *ATAQUES CIBERNÉTICOS NO 1S21*. 2021. Disponível em: <1https://blog.mzgroup.com/pt-br/ataques-ciberneticos-no-1s21/>. Acesso em: 10 junho 2022. Citado na página 10.

3 Filipe Prado. *Brasil foi 5º país com mais ataques cibernéticos no ano: relembre os principais*. 2021. Disponível em: <https://www.istoedinheiro.com.br/brasil-foi-5o-pais-com-mais-ataques-ciberneticos-no-ano-relembre-os-principais/>. Acesso em: 10 junho 2022. Citado na página 11.

4 Wikipédia, a enciclopédia livre. *Teste de intrusão*. 2021. Disponível em: <https://pt.wikipedia.org/wiki/Teste_de_intrus%C3%A3o>. Acesso em: 10 junho 2022. Citado na página 11.

5 BRASIL. Lei nº 12.737, de 30 de novembro de 2012. Dispõe sobre a tipificação criminal de delitos informáticos; altera o Decreto-Lei nº 2.848, de 7 de dezembro de 1940-Código Penal; e dá outras providências. Disponível em:<http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12737.htm>. Acesso em: 10 junho 2022. Citado na página 11.

6 Paulo Brito. *Falta de pessoal de cibersegurança afeta 70% das organizações*. Disponível em: <https://www.cisoadvisor.com.br/falta-de-pessoal-em-ciberseguranca-afeta-70-das-organizacoes/>. Acesso em: 10 junho 2022. Citado na página 11.

7 FILHO, C. G. L. et al. Análise de vulnerabilidades em redes de computadores–estudo de caso com a ferramenta nmap. Citado na página 13.

8 MONTEVERDE, W. A. *Estudo e análise de vulnerabilidades web*. Dissertação (B.S. thesis) — Universidade Tecnológica Federal do Paraná, 2014. Citado na página 13.

9 FRAGA, B. *Técnicas de Invasão: Aprenda as técnicas usadas por hackers em invasões reais*. Editora Labrador. ISBN 9786550440190. Disponível em: <https://books.google.com.br/books?id=YPXADwAAQBAJ>. Citado 2 vezes nas páginas 13 e 15.

10 PROTOCOLO HTTP: entenda o que é e para que serve! Disponível em: <https://www.hostgator.com.br/blog/o-que-e-protocolo-http/>. Acesso em: 11 junho 2022. Citado na página 13.

11 PISA, P. *Qual a diferença entre HTTP e HTTPS?* Disponível em: <https://www.techtudo.com.br/noticias/2012/07/qual-a-diferenca-entre-http-e-https.ghtml>,. Citado na página 14.

12 GOGONI, R. *O que é URL?* 2020. Disponível em: <https://tecnoblog.net/responde/o-que-e-url/>. Acesso em: 09 junho 2022. Citado na página 14.

13 O que é endereço IP – definição e explicação. Disponível em: <https://www.kaspersky.com.br/resource-center/definitions/what-is-an-ip-address>. Acesso em: 11 junho 2022. Citado na página 14.

14 IPV4 x IPv6: Qual é a diferença? Disponível em: <https://www.avast.com/pt-br/c-ipv4-vs-ipv6-addresses>. Acesso em: 11 junho 2022. Citado na página 14.

15 CORTES, B. R. B. Utilizando google hacking para encontrar vulnerabilidades em sites. *O Comunicante*, v. 7, n. 2, p. 23–27, 2017. Citado 2 vezes nas páginas 15 e 25.

16 SManager. *Vulnerabilidade na área de TI*. 2021. Disponível em: <https://smanager.com.br/vulnerabilidade-ti/>. Acesso em: 16 novembro 2022. Citado na página 15.

17 O que é um ataque de cross-site scripting? Definição e explicação. Disponível em: <https://www.kaspersky.com.br/resource-center/definitions/what-is-a-cross-site-scripting-attack>. Acesso em: 11 junho 2022. Citado na página 15.

18 MARIANO, C. d. O. M.; FELIX, T. R. Estudo do exploit spectre: funcionamento e possibilidades reais de ataque. Faculdade de Tecnologia de Americana, 2018. Citado na página 15.

19 InfoSec Write-ups. *How I was able to find 50+ Cross-site scripting (XSS) Security Vulnerabilities on Bugcrowd Public Program?* Disponível em: <https://infosecwriteups.com/how-i-was-able-to-find-50-cross-site-scripting-xss-security-vulnerabilities-on-bugcrowd-public-ba33db2b0 Acesso em: 20 novembro 2022. Citado na página 25.

ANEXO A – ESPECIFICAÇÃO DE CASOS DE USO

ESPECIFICAÇÃO DO CASO DE USO 01 — SOLICITAR ANÁLISE DE URL

Objetivo

Este documento tem por objetivo descrever todos os fluxos envolvidos no Caso de Uso 01 - Solicitar Análise de URL. São listados e detalhados todos os atores, fluxos, requisitos funcionais e não-funcionais.

Identificação dos Atores

Esta seção lista e descreve todos os atores envolvidos nos fluxos que compõem o Caso de Uso 01 - Solicitar Análise de URL.

- **Usuário**: Qualquer pessoa que utiliza o ISA;
- **Cliente**: Interface através da qual o Usuário utiliza o ISA.

Identificação dos Fluxos

Esta seção lista e descreve todos os fluxos que compõem o caso de uso.

- **Solicitar Análise**: Este fluxo descreve como se dá a solicitação de uma análise de um site.

Detalhamento dos Fluxos

- **Solicitar Análise de URL**: Este fluxo especifica a ação de efetuar uma solicitação de uma análise. O usuário fornece uma URL ao cliente o qual, junto ao sistema, faz uma varredura e retorna as vulnerabilidades encontradas.
    - **Atores**: Usuário, Cliente;
    - **Pré-Condições**: Nenhuma;
    - **Pós-Condições**: Nenhuma;
    - **Requisitos Funcionais**: O sistema deve prover uma interface para o resultado da análise.

    Fluxo Básico

    1. O ator Cliente mostra ao ator Usuário um formulário solicitando uma url;
    2. O Usuário Informa o endereço do site desejado;
    3. O Cliente envia a url ao Sistema;

4. O Sistema faz a varredura na url;
5. O Sistema registra os resultados no SGBD;
6. O Sistema retorna ao Cliente os dados das vulnerabilidades encontradas pelo Sistema;
7. O caso de uso se encerra.

Fluxo Alternativo

Não se aplica.

ESPECIFICAÇÃO DO CASO DE USO 02 — REALIZAR PESQUISA SOBRE SITES JÁ VERIFICADOS

Objetivo

Este documento tem por objetivo descrever todos os fluxos envolvidos no Caso de Uso 02 - Realizar Pesquisa sobre sites já verificados. São listados e detalhados todos os atores, fluxos, requisitos funcionais e não-funcionais.

Identificação dos Atores

Esta seção lista e descreve todos os atores envolvidos nos fluxos que compõem o Caso de Uso 02 - Realizar Pesquisa sobre sites já verificados.

- **Usuário**: Qualquer pessoa que utiliza a ISA;
- **Cliente**: Interface através da qual o usuário utiliza o ISA.

Identificação dos Fluxos

Esta seção lista e descreve todos os fluxos que compõem o caso de uso.

- **Pesquisar por *Vulnerabilidade***: Este fluxo especifica a ação de fazer uma busca de uma vulnerabilidade já existente no banco de dados.
    - **Atores**: Usuário, Cliente;
    - **Pré-Condições**: Nenhuma;
    - **Pós-Condições**: Nenhuma;
    - **Requisitos Funcionais**: O Sistema deve prover uma interface de buscar, para pesquisar um site analisado pela ferramenta.

    Fluxo Básico

    1. O ator Cliente mostra ao ator Usuário um formulário solicitando uma url qual site quer consultar no Banco de dados;
    2. O Cliente retorna ao usuário um formulário de busca;
    3. O Usuário informa;
    4. O Usuário solicita a busca;
    5. O Cliente faz uma consulta no banco de dados
    6. O Cliente retorna os resultados obtidos da pesquisa
    7. O caso de uso encerra;

    Fluxo Alternativo
    Não se aplica.