
Sistema de Detecção de Ataques de Engenharia Social para Dispositivos Android

Bianca Messias Ataíde



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

Monte Carmelo - MG
2024

Bianca Messias Ataíde

**Sistema de Detecção de Ataques de Engenharia
Social para Dispositivos Android**

Trabalho de Conclusão de Curso apresentado
à Faculdade de Computação da Universidade
Federal de Uberlândia, Minas Gerais, como
requisito exigido parcial à obtenção do grau de
Bacharel em Sistemas de Informação.

Área de concentração: Sistemas de Informação

Orientador: Diego Nunes Molinos

Monte Carmelo - MG

2024

*Este trabalho é dedicado primeiramente aos meus pais, irmão, madrinha e padrinho por
todo amor, apoio e dedicação ao longo dos anos*

Agradecimentos

A minha família por sempre me apoiarem nas minhas decisões, me guiando pelos caminhos da vida com muito amor, carinho e sabedoria. Aos meus padrinhos por sempre estarem presentes na minha vida, me ensinando e sendo também grandes exemplos para mim. Aos meus professores, em especial ao meu orientador Diego Molinos Nunes pela paciência e dedicação no caminho da construção dessa ideia. A Deus, não somente pelas bênçãos e conquistas, mas também por me permitir conhecer e conviver com pessoas admiráveis.

“Se eu vi mais longe, foi porque estava sobre os ombros de gigantes.”
(Sir Isaac Newton)

Resumo

Os golpes são uma realidade recorrente na sociedade a um certo tempo. Com a adição de novos meios tecnológicos dentro das relações humanas, criminosos começam a encontrar um novo nicho de ação: O mundo digital. Um dos tipos de ataques mais realizados, e que com o advento da digitalização vem se tornando mais recorrente e perigoso, são os ataques de engenharia social. Cibercriminosos utilizam de meios para enganar um indivíduo, podendo extrair dele não somente dinheiro, mas informações confidenciais, passando-se por pessoas confiáveis e fontes legítimas. O presente trabalho visa trazer um estudo acerca de ataques de engenharia social, apresentando uma solução para tentar identificar a possibilidade de um determinado texto poder ser uma tentativa de ataque de engenharia social ou não, utilizando entropia para realizar a classificação dos textos a partir de um *SetList* contendo palavras e expressões utilizados em ataques do tipo *phishing*. São realizadas medições para atestar acerca da eficiência do método de classificação utilizado. Ao final do estudo através das comparações entre os valores de entropia utilizados foi possível perceber que conforme o número máximo de palavras por conjunto aumentavam, há uma melhora na classificação dos textos que são *phishing* e uma piora na classificação dos que não são. Apesar disso os resultados são considerados insatisfatórios e ao final são adicionadas sugestões que podem melhorar a classificação dos textos.

Palavras-chave: Segurança, Android, Engenharia Social.

Lista de ilustrações

Figura 1 – Arquitetura do Sistema Operativo Android	20
Figura 2 – Fluxograma de Atividade	26
Figura 3 – Interface Aplicação Usuário	30
Figura 4 – Retorno da Análise de um dos Conjuntos	34

Lista de tabelas

Tabela 1 – Trabalhos Correlatos	24
Tabela 2 – Especificação da Solução	27
Tabela 3 – Tabela Experimentos Frases 50 palavras Entropia 1	33
Tabela 4 – Tabela Experimentos Frases 50 palavras Entropia 1.5	33
Tabela 5 – Tabela Experimentos Frases 100 palavras Entropia 1	34
Tabela 6 – Tabela Experimentos Frases 100 palavras Entropia 1.5	35
Tabela 7 – Tabela Experimentos Frases 150 palavras Entropia 1	35
Tabela 8 – Tabela Experimentos Frases 150 palavras Entropia 1.5	35
Tabela 9 – Tabela Experimentos Frases 200 palavras Entropia 1	35
Tabela 10 – Tabela Experimentos Frases 200 palavras Entropia 1.5	36
Tabela 11 – Tabela Resultados Entropia 1	36
Tabela 12 – Tabela Resultados Entropia 1.5	36

Lista de siglas

ART Android Runtime

API Application Programming Interface

CPU Central Processing Unit

HAL Hardware Abstraction Layer

IA Inteligência Artificial

SMS Short Message Service

USB Universal Serial Bus

URL Uniform Resource Locator

Sumário

1	INTRODUÇÃO	11
1.1	Motivação	12
1.2	Problema	12
1.3	Hipótese	13
1.4	Objetivos	13
1.5	Organização da Monografia	13
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	Segurança Cibernética	15
2.1.1	Segurança da Informação	16
2.2	Engenharia Social no contexto de Ataques Cibernéticos	16
2.2.1	Táticas e Técnicas de Ataques Baseados em Engenharia Social	17
2.2.2	Principais Ataques Baseados em Engenharia Social	18
2.3	Segurança Cibernética em Dispositivos Móveis	19
2.3.1	Android	19
2.4	Síntese da Fundamentação Teórica	21
3	TRABALHOS CORRELATOS	23
4	METODOLOGIA	25
4.1	Análise e Avaliação Acerca dos Tipos de Ataques de Engenharia Social	25
4.2	Sistema Operacional Android	26
4.3	Definição e Refinamento do Artefato de Requisitos da Solução	26
4.4	Especificação da Solução	27
4.5	Desenvolvimento da Solução	27
4.6	Ambiente de Teste	28
4.7	Análise de Resultados	28

5	DESENVOLVIMENTO	29
5.1	Ferramentas Utilizadas	29
5.1.1	Desenvolvimento da Interface	29
5.2	Definição do Conjunto de Palavras	29
5.3	Processamento e Classificação do texto	30
5.4	Síntese da Etapa de Desenvolvimento	31
6	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	32
6.1	Detalhes do Experimento	32
6.2	Experimentação	33
6.3	Análise dos Resultados	34
7	CONCLUSÃO	37
7.1	Principais Contribuições	37
7.2	Trabalhos Futuros	37
	REFERÊNCIAS	38

Introdução

Atualmente, os ataques cibernéticos têm se tornado uma questão de grande relevância para a sociedade (TIESO; SANTO, 2020). Devido à pandemia, muitos modelos de negócios precisaram se reestruturar para acomodar um nível elevado de comunicação e integração remota. Em paralelo à tendência de digitalização da sociedade, os ataques cibernéticos ganharam destaque, devido aos significativos prejuízos que causaram a empresas e indivíduos. Dentre os mais notáveis, os ataques de engenharia social estão em ascensão, principalmente porque não requerem necessariamente um extenso conhecimento tecnológico e exploram o elo mais frágil da segurança: o humano.

De acordo com Wang et al. (2022), Alkahtani e Aldhyani (2022), Tieso e Santo (2020), a engenharia social em dispositivos Android é potencializada por diversos fatores, entre eles a ampla oferta de aplicativos, a diversidade de dispositivos e sistemas operacionais, além da própria dinâmica de interação humana com esses aparelhos.

As formas de ataque mais comuns incluem *phishing*, *baiting* e *pretexting*. Nesses cenários, o invasor utiliza táticas persuasivas para enganar o usuário, induzindo-o a realizar ações que comprometem sua segurança. Nesses tipos de ataques, é usual que o usuário mal-intencionado envie mensagens via Short Message Service (SMS) ou e-mail, se passando por uma fonte confiável, tal como uma entidade bancária. Informações como conteúdo exclusivo e prêmios são frequentemente utilizadas como iscas para atrair o usuário (TIESO; SANTO, 2020).

A fragilidade dos atuais mecanismos de segurança, principalmente os antivírus frente aos ataques de engenharia social se deve, sobretudo, à exploração do elemento mais difícil de corrigir: o fator humano. Diferentemente dos *malwares*, que podem ser identificados por meio de assinaturas digitais ou comportamentos atípicos no sistema, os ataques de engenharia social são planejados para enganar indivíduos, não sistemas, introduzindo uma camada de complexidade na atuação de ferramentas de segurança.

Ante o exposto, este trabalho propõe a criação de uma sistema específico para detecção de ataques de engenharia social, principalmente ataques baseados em *phishing*, *baiting* e *pretexting* em dispositivos Android. O objetivo é através da inserção de mensagens

oriundas de comunicadores, e-mails, e ferramentas de mensagens instantâneas, o sistema seja capaz de identificar atividades suspeitas e notificar o usuário.

1.1 Motivação

Atualmente os ataques de engenharia social têm se desenvolvido, constituindo 40% dos roubos de credenciais, com pesquisas globais apontando que 73% das organizações sofreram vazamento de dados entre 2018 e 2023 (CARPES, 2023).

Não raro chegam histórias de pessoas que receberam mensagens através do *WhatsApp* de filhos ou outros parentes (tendo neste perfil a foto do usuário ou não) pedindo dinheiro ou algo do gênero, sendo também uma forma muito comum o uso de e-mails acompanhado de links ou anexos.

Pesquisas divulgadas pela *welivesecurity* apontam que no ano de 2020, no Brasil, houve um aumento em 200% no número de ataques de engenharia social, sendo o Brasil o segundo país da América Latina com mais detecções desse tipo de ataque (WELIVESECURITY, 2021). De acordo com Volpini (2021), os ataques de engenharia social são responsáveis por 70% dos golpes aplicados no mundo digital.

Neste contexto, ataques cibernéticos baseados em engenharia social tem se tornado comuns nos dias atuais e, apesar de ferramentas que permitem o usuário se proteger dos mais diversos tipos de ameaças digitais (e.g. antivírus), estas soluções ainda carecem de efetivos mecanismos de análise e detecção de ataques maliciosos em tempo real baseados em engenharia social.

Reis et al. (2018) em seu trabalho discorre sobre a insegurança que os usuários afirmam sentir ao fazer uso do *Internet Banking*, que nada mais é do que o acesso às informações bancárias através da Internet. Um dos ataques citados é o *Cross-Site Scripting*, que ocorre a partir da validação inadequada dos sites. Como citado, esse ataque pode ser performedo em conjunto a técnicas de ataques de engenharia social, se espalhando através das redes sociais.

1.2 Problema

No cenário contemporâneo, uma variedade de sistemas de detecção de spam é implementada, tanto para e-mails quanto para mensagens SMS. Algoritmos complexos são utilizados para treinar esses softwares na identificação de padrões em textos enviados aos usuários. O objetivo destes sistemas é eliminar ou redirecionar esses conteúdos potencialmente perigosos para locais seguros, protegendo, assim, a integridade das informações dos usuários.

Além disso, são realizadas campanhas de conscientização e treinamentos com a finalidade de instruir as pessoas sobre práticas seguras. Isso inclui evitar o acesso a links de

fontes não confiáveis e verificar a legitimidade dos remetentes.

Estratégias adicionais englobam a autenticação multifatorial, que proporciona uma camada extra de segurança ao requerer múltiplas formas de verificação de identidade. Também é recomendada a utilização contínua de programas antivírus, a manutenção e atualização regulares de software, reforçando as defesas contra invasões mal-intencionadas.

Apesar dos esforços significativos, os ataques de engenharia social evoluem e aprimoram-se. Isso representa uma preocupação crescente e evidência vulnerabilidades nas técnicas de segurança atuais.

1.3 Hipótese

A eficácia dos sistemas tradicionais de segurança em dispositivos móveis é insuficiente para mitigar ataques de engenharia social baseados em *phishing*, *baiting* e *pretexting* devido à sua incapacidade de interceptar as manipulações que visam diretamente os comportamentos humanos.

1.4 Objetivos

O objetivo deste trabalho é desenvolver uma ferramenta para a detecção de ataques cibernéticos baseados em engenharia social (*phishing*, *baiting* e *pretexting*) para dispositivos móveis (i.e. *Android Based*) através do processamento de mensagens maliciosas.

Todo o processo é guiado por objetivos específicos que conduzem a estratégia de desenvolvimento e estão elencados como se segue:

- ❑ Investigar o conjunto de palavras comumente utilizadas em ataques de engenharia social, com a finalidade de criar um conjunto de consulta para a aplicação.
- ❑ Aplicar os conceitos e técnicas de processamento de texto.
- ❑ Desenvolver um sistema capaz de receber informações, provenientes de mensagens, emails ou comunicadores em dispositivos móveis e emitir uma notificação alertando sobre uma possível tentativa de ataque cibernético.
- ❑ Apresentar, por meio de análises estatísticas, a precisão e o comportamento do usuário decorrentes da utilização do sistema de notificações.

1.5 Organização da Monografia

O presente trabalho encontra-se organizado da seguinte forma:

- ❑ O capítulo 2 apresenta a fundamentação teórica, onde serão abordados os principais conceitos acerca da segurança da informação. Ainda neste capítulo será explanado a Engenharia Social aplicada no contexto cibernético e a vulnerabilidade digital aplicada a dispositivos móveis.
- ❑ O capítulo 3 apresenta os trabalhos correlatos, ou seja, trabalhos que preconizam as soluções acerca de ataques cibernéticos em dispositivos móveis com ênfase em ataques baseados em engenharia social.
- ❑ O capítulo 4 apresenta a metodologia deste trabalho, discorrendo detalhadamente todos os passos utilizados para construção da solução.
- ❑ O capítulo 5 apresenta o desenvolvimento, discorrendo detalhadamente todos os passos utilizados para construção da aplicação e nas dificuldades encontradas.
- ❑ O capítulo 6 apresenta a experimentação, contendo os dados referentes aos testes realizados e os resultados obtidos a partir disso.
- ❑ O capítulo 7 apresenta a conclusão do estudo.

Fundamentação Teórica

O presente capítulo visa elucidar algumas questões que tangenciam a definição dos conceitos que irão apoiar o trabalho. Torna-se importante subdividir em tópicos este capítulo. Na primeira seção será abordada a definição de segurança da informação sob uma ótica mais generalista, trazendo alguns conceitos iniciais dentro da área. A seguir será abordado sobre o cerne da Engenharia Social, que trata do alicerce conceitual deste trabalho. Ao final será apresentado um pouco sobre o sistema android no qual a proposta de desenvolvimento se sustenta.

2.1 Segurança Cibernética

A Cisco Systems, Inc, uma das referências globais na produção de soluções baseadas em *hardware/software* para a área de infraestrutura de redes e segurança da cibernética, define a segurança cibernética como, "Os processos e às ferramentas projetados e implantados para proteger informações corporativas confidenciais contra modificações, interrupções, destruições e inspeções".

Estes ataques realizados de maneira virtual objetivam acessar, alterar ou destruir alguma informação confidencial ou causar algum tipo de dano monetário e patrimonial aos usuários (seja através de extorsão ou de interrupção na continuidade de algum negócio).

Diante da quantidade de dispositivos tecnológicos conectados atualmente, a criação de um plano consistente que contemple todas as camadas de segurança se torna um desafio.

Do ponto de vista do usuário, é essencial promover a conscientização sobre práticas seguras, que incluem não apenas o uso de senhas robustas, a manutenção de backups e a atualização de sistemas de segurança, mas também o treinamento para reconhecer ameaças, como o recebimento de e-mails com links e arquivos maliciosos.

2.1.1 Segurança da Informação

A segurança da informação, definida pela Cisco Systems, Inc e referida como InfoSec (*Information Security*), trata-se da área da computação que visa agrupar os processos e ferramentas usadas com a finalidade de proteger as informações e dados confidenciais corporativos.

A segurança da informação tem como pilares básicos os conceitos de (a) confidencialidade, que visa garantir a privacidade dos dados, (b) integridade, que visa garantir que os dados não possam ser corrompidos ou alterados indevidamente e, (c) disponibilidade, que visa garantir que os dados estejam disponíveis para serem acessados por pessoal devidamente credenciado respeitando as regras de negócio (MACHADO, 2014).

A segurança da informação é um campo em constante progresso, cuja importância é crítica, especialmente na atualidade, onde a dependência de sistemas digitais é cada vez mais intensa.

De acordo com Aslan et al. (2023), o panorama das ameaças cibernéticas é dinâmico e complexo, com o surgimento contínuo de novas vulnerabilidades e o desenvolvimento de ataques para explorar estas fragilidades antes da implementação de medidas corretivas.

As vulnerabilidades cibernéticas são falhas ou fraquezas em um sistema que podem ser exploradas por um atacante para comprometer a integridade, disponibilidade ou confidencialidade de informações. Elas podem ser resultantes de erros de software, configurações inadequadas, práticas deficientes de segurança ou tecnologias obsoletas. A identificação e correção dessas vulnerabilidades são essenciais para a manutenção da segurança (ASLAN et al., 2023; MACHADO, 2014).

Os ataques cibernéticos podem se manifestar de diversas maneiras, tais como malware, phishing, ataques de negação de serviço (DoS) e a exploração de vulnerabilidades de dia zero e também através da engenharia social. Cada categoria de ataque visa explorar falhas específicas, seja em software, hardware ou no comportamento humano, com o intuito de atingir objetivos mal-intencionados (ASLAN et al., 2023; TIESO; SANTO, 2020).

Torna-se importante reforçar que, no contexto deste trabalho será explorado o conceito de engenharia social como mecanismo de ataque cibernético, tipos de ataques e as formas de mitigação também serão detalhadas nas seções abaixo.

2.2 Engenharia Social no contexto de Ataques Cibernéticos

Engenharia social no contexto de ataques cibernéticos refere-se aos métodos prejudiciais usados para enganar pessoas a tomar ações ou compartilhar informações privadas (TIESO; SANTO, 2020). Diferente de outros ataques que usam *software* ou *hardware*, a

engenharia social se concentra em explorar aspectos humanos, como confiança e medo, para causar erros.

Segundo Mitnick e Simon (2003),

“A engenharia social usa a influência e a persuasão para enganar as pessoas e convencê-las de que o engenheiro social é alguém que na verdade ele não é, ou pela manipulação. Como resultado, o engenheiro social pode aproveitar-se das pessoas para obter as informações com ou sem o uso da tecnologia. ”

Trata-se de ações pautadas na manipulação psicológica do usuário que podem ou não incluir o uso de tecnologia e pode ser combinada a qualquer outro tipo de ameaça/ataque, como o de *phishing* por exemplo, levando o usuário a acessar ou baixar conteúdos de uma fonte mal-intencionada.

Ataques cibernéticos tratam, essencialmente, da manipulação emocional e técnicas que visam garantir a persuasão do usuário a tomar atitudes que não são necessariamente de seu interesse, por vezes também é referida como *hacking humano* (TIESO; SANTO, 2020).

Ataques comuns de engenharia social do contexto cibernético incluem *phishing*, que usa emails ou mensagens falsas para obter dados sensíveis; *pretexting*, onde o atacante cria uma história falsa para conseguir informações; e *baiting*, que oferece algo atraente para instalar um software malicioso no sistema da vítima (TIESO; SANTO, 2020; BARBOSA et al., 2021).

Estes tipos de ataques serão detalhados na seção abaixo.

2.2.1 Táticas e Técnicas de Ataques Baseados em Engenharia Social

Para o atingimento do objetivo, usuários maliciosos que utilizam do ataque de engenharia social¹ devem se passar por uma fonte legítima e confiável para o usuário atacado. O objetivo é ter acesso a informação sigilosa, e.g. senhas de bancos, cartões de créditos, senhas de e-mails, dentre outras informações restritas. Fato é que, esses mecanismos podem ser materializados de diversas formas, abaixo seguem algumas maneiras:

- ❑ Os engenheiros sociais se passam por empresas conhecidas, confiáveis ou interligadas à vítima.
- ❑ Os engenheiros sociais se passam por autoridades políticas ou governamentais respeitadas.

¹ doravante, o termo “engenheiro social” e “usuário malicioso” que empregam conceitos de engenharia social para causar algum prejuízo possuem o mesmo significado.

- ❑ Os engenheiros sociais usam de técnicas psicológicas para criar uma situação de urgência que visa atraparalhar o senso de decisão do usuário pela carga emocional da situação, criando um gatilho de resposta rápida não consciente.
- ❑ Golpes que visam oferecer uma recompensa financeira gigante baseada no envio de algum dado ou informação (normalmente relacionados a transações bancárias, como informações de conta ou envio de dinheiro).
- ❑ Muitas vezes usada em conjunto com ataques de *phishing*, o engenheiro social realiza o ataque a partir do envio de alguma mensagem contendo um link ou arquivo que o usuário acessa por curiosidade para saber do que se trata aquilo e acaba infectado.

2.2.2 Principais Ataques Baseados em Engenharia Social

Conforme contextualizado anteriormente, a engenharia social pode decorrer em detrimento a outros tipos de ataques (TIESO; SANTO, 2020).

Abaixo seguem os principais ataques que convergem para esse estudo e possuem impacto direto neste trabalho:

2.2.2.1 *Phishing*

São primordialmente mensagens de texto ou de voz (*Vishing*) com a finalidade de manipular o usuário a compartilhar informações confidenciais, fazer transferências de dinheiro ou ativos ao atacante ou baixar softwares mal-intencionados.

Os atacantes podem se passar por pessoas (podendo tentar performar alguém do ciclo social do usuário) e também por organizações confiáveis. Neste tipo de ataque é comum ocorrer o envio de e-mails com links e programas anexos, uso parcial de credenciais ou informações vazadas para atingir alvos específicos (*spear phishing*), envio de SMS; contas falsas em mídias sociais (*angler phishing*) ou chamadas telefônicas gravadas (*vishing*) (TIESO; SANTO, 2020; SILVA; SOUSA; JÚNIOR, 2023).

2.2.2.2 *Baiting*

É a forma de ataque no qual o usuário compartilha informações confidenciais ou instala software mal-intencionado através de uma isca com uma oferta de retorno valiosa. Pode ser performado visando o retorno financeiro ou o retorno material (um Universal Serial Bus (USB) ou um software pago com uma versão pirata disponibilizada que possui código malicioso) (TIESO; SANTO, 2020; SILVA; SOUSA; JÚNIOR, 2023).

2.2.2.3 *Pretexting*

O atacante cria uma situação na qual informa que o usuário tem um problema grave (como alguma informação envolvida em um vazamento ou violação de dados) e diz ao

usuário que ele pode resolver essa situação para ele, pedindo alguma informação que irá permitir que ele execute um ataque a contas ou objetos de *hardware* do usuário (TIESO; SANTO, 2020; SILVA; SOUSA; JÚNIOR, 2023).

2.2.2.4 *Scareware*

Usa o medo para manipular o usuário a baixar software malicioso ou divulgar alguma informação confidencial, geralmente sendo performado como um aviso falso (alguma atualização de software ou alguma organização pública enviando um acusando o usuário de algum crime) (SILVA; SOUSA; JÚNIOR, 2023).

2.3 Segurança Cibernética em Dispositivos Móveis

Segundo a pesquisa desenvolvida por Amado (2002), é seguro admitir que os dispositivos móveis são equipamentos desenvolvidos a partir do avanço das comunicações sem fio, permitindo assim, o transporte de uma máquina e fomentando discussões acerca de métodos de preservar a sua capacidade computacional.

Os dispositivos móveis de acordo com o autor supracitado, faz uso do conceito de portabilidade, que trata da característica que permite o transporte do objeto, levando em conta a sua estrutura física. Neste contexto, o transporte permite a ausência de cabos e/ou presença de fontes de energia como baterias), comunicação (que visa garantir que ele consiga se comunicar através de redes sem fio) e a mobilidade (a característica que permite a mudança de localidade garantindo a permanência da conexão com a rede, o que causa um aumento na volatilidade da informação).

Esses dispositivos contam com *softwares* que visam facilitar a comunicação entre o usuário e o sistema (os sistemas operacionais).

Não obstante a adoção de práticas de segurança robustas, a segurança da informação em dispositivos móveis é uma área em constante evolução, com novas ameaças surgindo regularmente. Com o advento de novos métodos, combinado com tecnologias de proteção atualizadas, a segurança em dispositivos móveis é essencial para proteger tanto os usuários individuais quanto as organizações contra riscos.

2.3.1 Android

Trata-se de uma plataforma de desenvolvimento para dispositivos móveis criada pela Google a partir do kernel do sistema operacional Linux, possuindo funcionalidades pré-instaladas e desenvolvidas que se adequam ao fabricante (Motorola, Samsung, LG, entre outras) (PEREIRA; SILVA, 2009; LECHETA, 2013).

O Android foi oficialmente lançado em 2008 e possui grande adesão da grande maioria dos dispositivos móveis em função da flexibilidade e código aberto, o que contribui cada

vez para a adoção desta tecnologia. Atualmente, a plataforma continua a evoluir com atualizações regulares e novas funcionalidades (SHABTAI et al., 2010).

A Figura 1, disponibilizada no site oficial da plataforma Android para desenvolvedores, ilustra a arquitetura do sistema operativo.

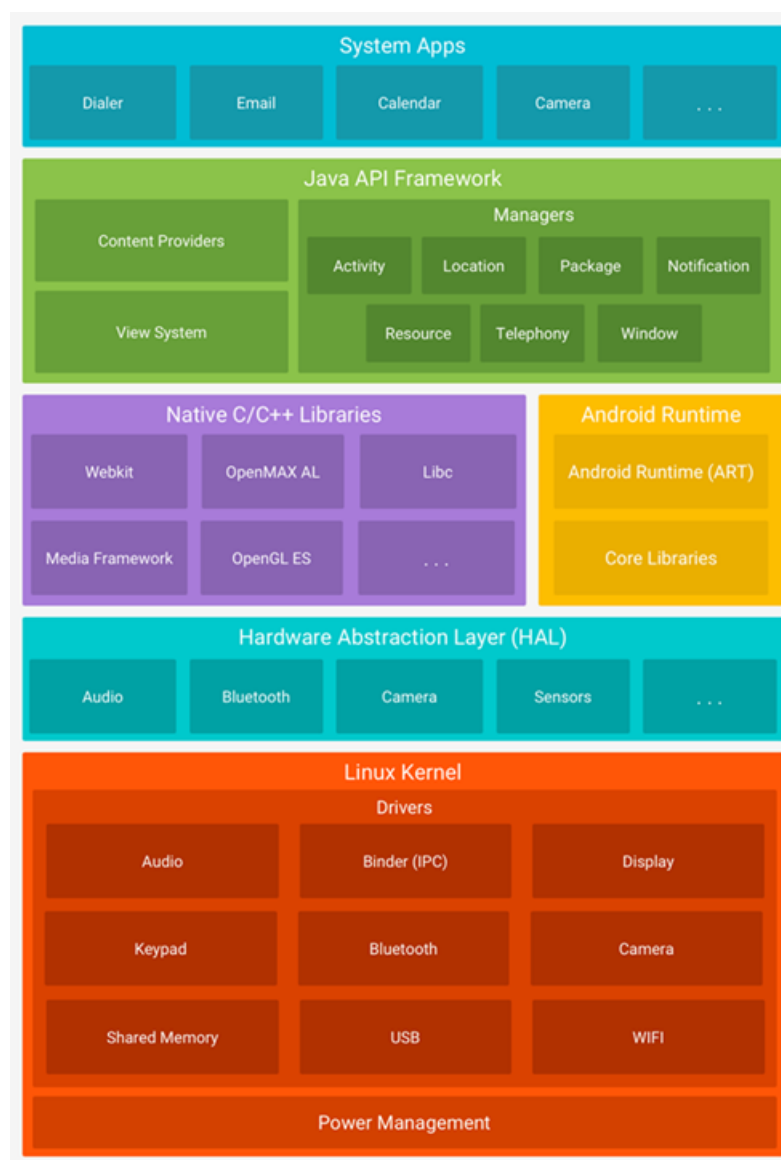


Figura 1 – Arquitetura do Sistema Operativo Android

Fonte: (<<https://developer.android.com/guide/platform?hl=pt-br/>>)

Conforme a Figura 1 é possível observar a subdivisão do Kernel, onde estão concentrados todos os drivers do sistema; sendo seguido pelo Hardware Abstraction Layer (HAL), que é uma camada que permite a comunicação dos dispositivos e sensores com os aplicativos; a seguir temos o Android Runtime (ART), que cria uma máquina virtual para permitir a execução das aplicações dentro do sistema Android.

O sistema Android conta com bibliotecas nativas em escritas em C e C++ e a Application Programming Interface (API) Java que visam permitir o desenvolvimento de aplicações no sistema, possibilitando que aplicações se comuniquem e troquem informações. Por último, se concentram os aplicativos que já vem por padrão dentro do sistema, como o aplicativo da câmera, o calendário, o email, entre outros.

Como de praxe, o desenvolvimento nativo para ambiente android é sempre dotado de desafios, ora havendo necessidade de interagir com o Kernel ou modificar as estruturas já predefinidas.

Diante disto, ferramentas tais como o Kotlin, são concebidas para auxiliar o desenvolvimento de aplicativos Android. Desenvolvida pela JetBrains, Kotlin foi projetada para ser interoperável com Java, facilitando o uso combinado das duas linguagens em projetos. Kotlin traz várias melhorias sobre o Java, o que resulta em código mais enxuto, legível e seguro (JEMEROV; ISAKOVA, 2017).

2.3.1.1 Kotlin

Inspirada em linguagens como Swift, Scala, Groovy e C#, o Kotlin é uma linguagem concisa, amigável e multiplataforma, permitindo o reuso de código em diferentes sistemas operacionais. Faz parte do projeto de código aberto e é usado por mais de 60% dos desenvolvedores segundo o site oficial do Android (Moskala e Wojda (2017)).

É descrita como uma linguagem de programação segura, concisa, moderna, expressiva e versátil que ajuda a aumentar a produtividade e a satisfação dos desenvolvedores e corrige alguns dos problemas do Java como a verbosidade e a nulidade de ponteiro (permitindo o desenvolvimento de código mais seguro a partir da implementação de um tipo nulidade através do `NullPointerException`).

Teve como base de desenvolvimento a busca pelas melhores práticas (usando como guia livros como “*Clean Code*” e “*Effective Java*”), buscando não repetir erros de outras linguagens e aproveitar as *features* mais utilizadas.

2.4 Síntese da Fundamentação Teórica

Este capítulo teórico esclareceu conceitos-chave para compreender a segurança da informação e a engenharia social no âmbito dos dispositivos móveis, ambos campos vitais na batalha contra os ataques cibernéticos contemporâneos. Foi possível compreender que a segurança cibernética é crucial para resguardar dados empresariais contra acessos não permitidos, modificações ou destruições, uma demanda que se amplia devido à crescente interligação dos dispositivos tecnológicos.

O estudo bibliográfico desta seção fez surgir importantes destaques para o trabalho, entre os principais cita-se:

- ❑ As práticas enfatizadas, como a conscientização sobre o uso de senhas fortes, a realização de backups regulares e a implementação de atualizações de segurança, são essenciais.
- ❑ Importância da educação dos usuários para identificar e-mails e links mal-intencionados. Este aspecto desempenha um papel crucial, visto que a engenharia social explora as vulnerabilidades humanas ao invés de falhas tecnológicas.
- ❑ A engenharia social, que se destaca por manipular os aspectos psicológicos das vítimas, como confiança e medo, para induzir erros, é especialmente pertinente no contexto dos dispositivos móveis.
- ❑ Técnicas como *phishing*, *pretexting* e *baiting* são comumente empregadas para extrair informações sensíveis ou instalar software mal-intencionado.
- ❑ Impacto da segurança em dispositivos móveis (i.e. sistema operacional Android) devido à sua prevalência e flexibilidade que atrai tanto usuários regulares quanto usuário maliciosos.

Trabalhos Correlatos

Nesse capítulo serão detalhados os trabalhos correlatos que se possuem relação com os objetivos definidos no capítulo 1 desta monografia.

No estudo de Takeuchi (2023), o autor explora o conceito de “fator humano” na tecnologia, que se refere aos funcionários e colaboradores. O autor argumenta que esses indivíduos podem representar um risco à segurança de um ativo. O trabalho se dedica a realizar uma revisão bibliográfica, abordando tópicos que explicam e contextualizam alguns dos ataques cibernéticos mais comuns, como *phishing*, *ransomware* e engenharia social. Além disso, o estudo descreve como a teoria das personalidades *Big Five* pode se conectar com o cibercrime e a segurança. Neste trabalho também são mencionados alguns métodos empregados para assegurar a segurança dos ativos.

Em Rezende (2023) o autor investiga sobre a taxonomia de *malwares* que afetam o sistema Android, fazendo uso da técnica de engenharia reversa que consiste em traduzir o “binário” da aplicação para outra linguagem¹ de forma a permitir que seja feito um estudo em cima do comportamento e das requisições que aquele programa faz ao sistema. Um dos estudos de caso proposto no trabalho demonstra que a aplicação transformava o dispositivo em um *bot* de envio de mensagem SMS.

Em Reis et al. (2018) é desenvolvido um estudo acerca das vulnerabilidades mais comuns em sistemas de *Internet Banking*, dentre as coisas uma que é performada através de ataques de engenharia social.

No estudo de Souza e Fernandes (2016), os autores investigam a relação entre divulgação de informações em redes sociais e ataques de engenharia social. Neste trabalho os autores realizam uma análise de vulnerabilidade a partir de uma análise de redes sociais com o viés de debater sobre os pormenores de confiança tanto pessoal quanto organizacional.

Em BIAZZOTTO, ANTICOLI e BOCCIA (2020) os autores trazem um comparativo com três softwares que visam prevenir ataques de engenharia social às corporações ocorridos a partir do comprometimento de email corporativo. Os softwares comparados são

¹ processo de (re)montagem do assembly, técnica é comumente chamada de “disassemble”

Trend Micro Advanced Threat Protection; *Gafety Secure Email Gateway* e *Microsoft Advanced Threat Protection*, que fazem a análise do conteúdo das mensagens enviadas a um email a fim de detectar alguns tipos de tentativas de ataque, tais como *phishing*. Dentro do estudo é relatado que além do conteúdo das mensagens, alguns dos *softwares* fazem análise de anexos enviados, detecção de remetentes e Uniform Resource Locator (URL) fazem uso de Inteligência Artificial e o *Trend Micro Advanced Threat Protection* defende criar mensagens de *phishing* personalizadas para os usuários afim de identificar e criar propostas de treinamento a eles.

Sob a ótica da segurança cibernética, a compreensão e a implementação de conceitos como engenharia social, segurança do sistema operacional e acessibilidade são fundamentais para proteger usuários e organizações contra uma variedade de ameaças. Cada um desses conceitos desempenha um papel crítico na criação de um ambiente móvel seguro.

Na Tabela 1 ilustra a relação das soluções apresentadas nos trabalhos correlatos e os conceitos de segurança cibernética, uso ou não de técnicas para detecção e mitigação de ataques por engenharia social, suporte ao sistema Android e se as soluções descritas apresentam suporte para acessibilidade.

Tabela 1 – Trabalhos Correlatos e sua relação com os tópicos principais do trabalho proposto.

Trabalhos	Segurança Cibernética	Suporte à Engenharia Social	Suporte ao Android
Takeuchi (2023)	X	X	X
Rezende (2023)	X	X	
Reis et al. (2018)	X	X	
Souza e Fernandes (2016)	X	X	
BIAZZOTTO, ANTICOLI e BOCCIA (2020)	X	X	

Fonte: Autoria Própria

Metodologia

O objetivo deste capítulo concentra-se em detalhar o método de pesquisa adotado, sendo idealizado para:

- ❑ Trazer a descrição e a sequência de passos que se planeja adotar na construção do projeto e,
- ❑ Definir e delimitar o escopo da proposta. O método escolhido é o hipotético-dedutivo, cujo início se dá a partir da percepção de um problema, a formulação de uma hipótese de pesquisa e culmina na proposição de uma solução para testar a ocorrência da hipótese.

Uma das máximas deste trabalho é contribuir para a segurança de usuários através de um sistema de detecção de ataques de engenharia social ocorridos em dispositivos móveis baseados em Android. Diante deste contexto, o método contemplará as etapas de caracterização, desenvolvimento e validação da proposta através de um ambiente de testes.

Através da Figura 2 é possível observar o diagrama de atividade que visa contemplar um fluxograma e ilustrar os passos adotados no desenvolvimento dessa pesquisa.

4.1 Análise e Avaliação Acerca dos Tipos de Ataques de Engenharia Social

Nesta etapa do método uma revisão foi realizada da literatura correlata com o objetivo de identificar ataques de engenharia social contendo suas palavras-chave, métodos, plataformas em que ocorrem e principais afetados.

A partir dessa pesquisa foi possível perceber os métodos mais utilizados para se passar credibilidade e entender os sentimentos que se deseja gerar a partir deles.

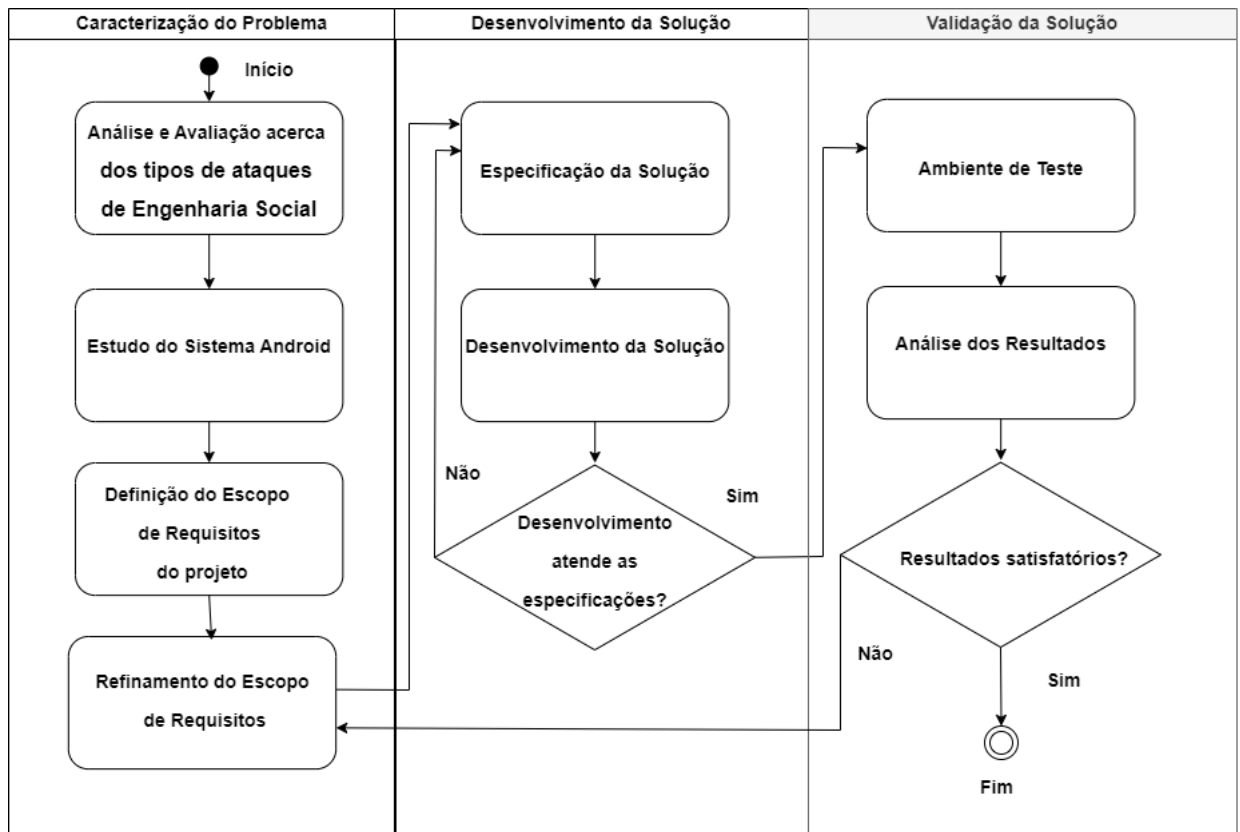


Figura 2 – Fluxograma de Atividade

Fonte: Autoria Própria

Também foi possível perceber que os ataques de engenharia social também podem se correlacionar a outros tipos de ataque em conjunto com arquivos ou links, o que será utilizado posteriormente na construção da solução.

4.2 Sistema Operacional Android

Nesta etapa do método, foi conduzido uma análise acerca do sistema operacional Android para identificar quais as bibliotecas e estruturas que o sistema já possui nativamente que são ofertadas através de API. Esta etapa é importante pois permitiu refinar a especificação da solução.

4.3 Definição e Refinamento do Artefato de Requisitos da Solução

Reforçando que a premissa principal deste trabalho é contemplar ataques derivados de mensagens de texto em plataformas digitais, tendo como principal enfoque aquelas

que desejam causar dano patrimonial/financeiro ao usuário, foi construído o artefato de requisitos.

Através da revisão da literatura, trabalhos correlatos e, estudo de soluções similares, foi possível elencar os requisitos e limitações deste trabalho. Não obstante o levantamento de requisitos, o refinamento dos mesmos foi realizado para otimizar o processo de desenvolvimento.

4.4 Especificação da Solução

Torna-se importante reforçar que, este projeto possui foco para sistema operacional Android.

No que tange a especificação da solução, o objetivo do sistema é receber uma entrada de textos e realizar o processamento de análise de palavras chaves que permite identificar contextos maliciosos. Para mensurar a probabilidade de uma mensagem ser maliciosa, aplica-se a fórmula de Shannon para calcular a entropia dos termos presentes, com alertas emitidos ao usuário caso o valor supere um limiar definido. Este processo visa ser leve e eficiente, com testes conduzidos em dispositivos Android 10 para garantir que a solução mantenha um equilíbrio entre precisão na detecção e um baixo consumo de recursos do sistema. A Tabela 2 apresenta todas as ferramentas utilizadas no desenvolvimento.

Tabela 2 – Especificações da Solução contendo as tecnologias utilizadas

Tecnologias	Especificações
Plataforma	Android
Linguagens de Programação	Kotlin
Plataforma de Desenvolvimento (IDE)	Android Studio
Plataforma de Teste	Dispositivo Móvel com Sistema Operacional Android 10

Fonte: Autoria Própria

4.5 Desenvolvimento da Solução

Para o desenvolvimento desta solução com foco em detecção de ataques de engenharia social em dispositivos Android, foi adotado um conjunto de ferramentas e metodologias que garantem eficiência e precisão no reconhecimento de possíveis tentativas de *phishing* ou manipulação. A linguagem de programação escolhida foi o Kotlin, que oferece suporte nativo para Android e é conhecida por sua compatibilidade com Java, além de ser mais concisa e segura. O ambiente de desenvolvimento Android Studio foi utilizado para construir e testar o aplicativo, aproveitando suas funcionalidades avançadas para emular o sistema operacional Android, configurar permissões específicas e garantir a integração com APIs de interesse para o projeto. A solução foi projetada para ser executada em

dispositivos com Android 10, garantindo compatibilidade e estabilidade com um sistema amplamente utilizado, o que facilita a replicação dos testes em diversos dispositivos.

No processamento dos dados, a aplicação utiliza uma lista de palavras-chave gerada a partir de bases de dados de *phishing* e *spam*, expandida com sinônimos para melhorar a precisão. O texto capturado passa por um pré-processamento, onde é transformado em uma lista de palavras em letras minúsculas e sem acentuação. A fórmula de Shannon é aplicada ao texto com o objetivo de calcular a entropia e identificar padrões que possam indicar uma tentativa de ataque. Este processo visa reconhecer a complexidade do texto, calculando a probabilidade de ocorrência de termos associados a ataques de engenharia social. Caso o valor da entropia ultrapasse um limiar previamente definido, o sistema emite um alerta ao usuário, promovendo a conscientização e incentivando a verificação do remetente ou da mensagem.

4.6 Ambiente de Teste

Para a etapa de testes, no que tange os testes de precisão e a eficiência do sistema, foi configurado um ambiente de testes com um banco de dados contendo milhares de mensagens classificadas como *phishing* e não *phishing*. Este processo de avaliação permitiu a identificação de pontos de melhoria e a otimização da configuração de alertas, com o objetivo de reduzir a ocorrência de falsos positivos e falsos negativos. Além disso, o uso de dispositivos físicos e emuladores no Android Studio permitiu simular cenários reais de uso, garantindo que o sistema funcione em diversas condições e que a aplicação esteja pronta para atender às necessidades de segurança e acessibilidade de usuários vulneráveis a ataques de engenharia social.

4.7 Análise de Resultados

Nesta etapa do método, com base nos testes experimentados, avaliou a eficiência da solução na detecção de ataques, ou seja na inferência, considerando métricas como acurácia, falsos positivos e falsos negativos para entender sua precisão.

Ainda foi compreendido na etapa de análise, comparações com soluções correlatas, identificando vantagens e limitações da solução.

Desenvolvimento

O objetivo deste capítulo é apresentar de forma detalhada o processo de desenvolvimento da aplicação, desde sua concepção até sua implementação final, destacando as etapas realizadas, as escolhas técnicas e os desafios enfrentados ao longo do projeto.

5.1 Ferramentas Utilizadas

Para a construção da solução, foram empregadas tecnologias amplamente utilizadas no ecossistema Android, incluindo a linguagem de programação Kotlin e a plataforma de desenvolvimento Android Studio. A escolha do Kotlin foi motivada por sua interoperabilidade com Java, simplicidade sintática e segurança aprimorada contra erros comuns, como nulidade de ponteiros. O ambiente de teste foi configurado em dispositivos físicos e emulados com Android 10, garantindo compatibilidade e estabilidade.

5.1.1 Desenvolvimento da Interface

Com o estudo sobre a entropia desenvolvido foi criado um aplicativo Android na linguagem Kotlin que recebe um texto de entrada do usuário. Ao clicar no botão submit o usuário envia o texto para análise. Caso a entropia seja maior que 1.5, é retornado ao usuário uma caixa de texto do tipo *Dialog* contendo o texto “A mensagem foi classificada como uma possível tentativa de ataque de Engenharia Social!! Fique atento a autenticidade do conteúdo da mensagem e do autor.”, caso contrário a caixa de texto deve retornar “A mensagem não foi classificada como uma possível tentativa de ataque de Engenharia Social.”. A Figura 3 ilustra a interface do usuário.

5.2 Definição do Conjunto de Palavras

Foi realizada a busca por conjunto de dados contendo as palavras mais utilizadas dentro de textos classificados como *phishing* ou *spam* para construção de um *dataset* para

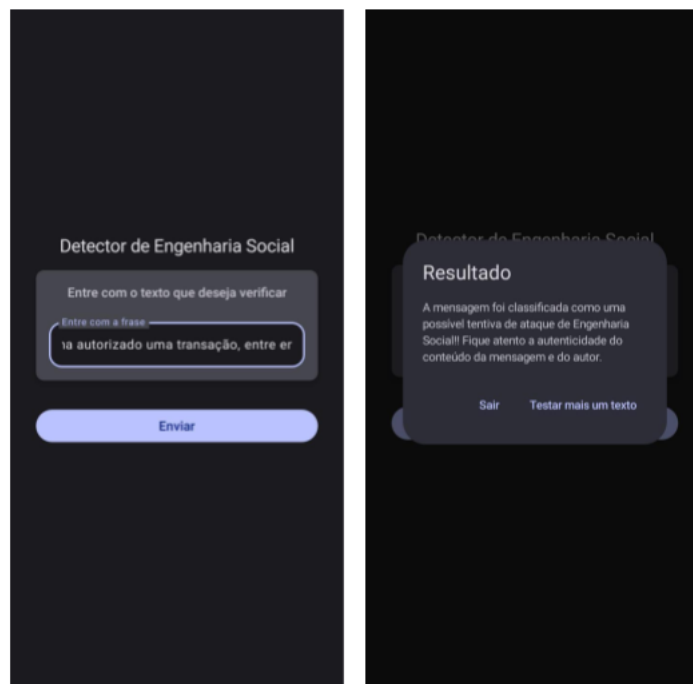


Figura 3 – Interface Aplicação Usuário

Fonte: Autoria Própria

suportar o processo de análise.

Foram observados vários estudos que utilizavam LLM (*Large Language Models*, modelos de Inteligência Artificial (IA) que utilizam bases de dados para processamento de texto) utilizavam alguns bancos de dados como Nasario e AssassinSpam para realizar a classificação dos seus dados. Essas bases correspondem a um conjunto de emails com informações como destinatário, remetente, data e conteúdo de emails classificados como phishing e spam, respectivamente.

Para gerar então o conjunto de dados de busca desse estudo foi gerada uma lista contendo palavras-chave e expressões utilizadas em ataques de *phishing* e a partir do conjunto inicial foi realizada uma expansão com sinônimos da lista inicial.

5.3 Processamento e Classificação do texto

Foi criada uma interface contendo uma tela que deve receber um texto de entrada do usuário, que é convertido em lista contendo as palavras em letras minúsculas e sem acentuação e enviado a uma função para calcular a entropia.

O cálculo se dá através da medição da frequência dos termos presentes no conjunto gerado na seção anterior e a partir desses valores se é utilizada a fórmula de Shannon

(MAGOSSİ; BARROS, 2021), para calcular a entropia dada por

$$H(X) = - \sum_{i=1}^n p_i * \log_2(p_i) \quad (1)$$

Essa fórmula visa calcular e retornar a probabilidade (em *bits*) de um valor aparecer. Para realizar a classificação foram realizados testes empíricos baseados nos valores de entropia de textos classificados como phishing com o conjunto dos dados e chegou-se ao valor 1.5. Portanto, caso a entropia seja maior do que 1.5 o texto pode ser uma tentativa de ataque de engenharia social.

5.4 Síntese da Etapa de Desenvolvimento

Durante o desenvolvimento da solução foram enfrentados desafios significativos. Dentre eles cita-se:

- ❑ **Ajuste do conjunto de palavras-chave:** Diversos conjuntos de textos são encontrados em inglês, optou-se por traduzir para o idioma português devido a aplicabilidade da solução apresentada. Neste contexto o desafio foi traduzir os textos sem perder o conteúdo semântico.
- ❑ **Calibração da Entropia:** Determinar um limiar ideal para classificar mensagens exigiu vários experimentos, já que valores muito baixos ou altos geravam falsos positivos ou negativos.
- ❑ **Desempenho em Dispositivos com Recursos Limitados:** A aplicação foi projetada para ser leve, garantindo eficiência em dispositivos com hardware limitado.

Experimentos e Análise dos Resultados

O objetivo deste capítulo é apresentar detalhadamente os experimentos realizados e os resultados obtidos, com foco na análise da precisão, eficiência e robustez do método proposto para a classificação de textos.

A avaliação busca identificar a eficácia da solução desenvolvida, considerando métricas como acurácia, *recall*, precisão e taxa de falsos positivos e negativos, além de apontar possíveis melhorias.

6.1 Detalhes do Experimento

Foi montado um banco de teste contendo 2.004 textos extraídos de bancos contendo emails e mensagens classificados como *phishing* e 2.004 textos que não são classificados como *phishing*. Eles foram agrupados em 4 conjuntos seguindo o critério de contagem de palavras: O primeiro conjunto apresenta frases com no máximo 50 palavras, o segundo conjunto apresenta frases que respeitam o intervalo 51 até 100 palavras, o terceiro conjunto vai de 101 até 150 e o último conjunto apresenta intervalo de 151 até 200.

Cada conjunto contém 501 frases com conteúdo malicioso e 501 frases sem conteúdo malicioso, cada uma apresentada sua classificação associada. Esses conjuntos foram convertidos em uma estrutura *Map* e a aplicação fez a avaliação um a um.

Foram utilizados dois valores de entropia para testes: 1 e 1.5. Se a entropia fosse superior ou igual ao valor de teste (1 e 1.5) e a frase possuísse classificação como *Phishing*, então a variável de contagem para verdadeiros positivos era incrementada. Se o valor fosse menor do que 1.5 e a frase possuísse classificação como *Phishing*, então a variável de contagem para falsos positivos era incrementada. Se a entropia fosse superior ou igual ao valor de teste (1 e 1.5) e a frase possuísse classificação como *Safe*, então a variável de contagem para verdadeiros negativos era incrementada. Se o valor fosse menor do que 1.5 e a frase possuísse classificação como *Safe*, então a variável de contagem para falsos negativos era incrementada.

6.2 Experimentação

Os experimentos foram conduzidos em um dispositivo móvel com as seguintes especificações de hardware:

- ❑ Processador: 4x 2.0 GHz Cortex-A53 + 4x 1.5 GHz Cortex-A53
- ❑ Chipset :Helio G25 MediaTek
- ❑ 64-bit
- ❑ GPU: PowerVR GE8320
- ❑ RAM: 2 GB
- ❑ Sistema Operacional: Android 10
- ❑ Marca: Motorola
- ❑ Modelo: XT2095-1
- ❑ Linha: Moto E
- ❑ Série: E7

Foi criada uma classe que irá receber cada um dos conjuntos em formato de estrutura *Map*, sendo armazenados os textos e sua classificação. Em seguida a aplicação foi modificada para receber um dos conjuntos e processar ele no lugar da frase do usuário. Ao final a aplicação retorna um *Dialog* contendo os valores da matriz de confusão respeitando o que foi determinado na seção 6.1 . A Figura 4 ilustra a saída recebida pela análise de um dos conjuntos.

Tabela 3 – Tabela contendo o resultado dos experimentos com frases contendo até 50 palavras e entropia 1

	Verdadeiro	Falso
Positivo	58	55
Negativo	441	418

Fonte: Autoria Própria

Tabela 4 – Tabela contendo o resultado dos experimentos com frases contendo até 50 palavras e entropia 1.5

	Verdadeiro	Falso
Positivo	17	6
Negativo	490	459

Fonte: Autoria Própria



Figura 4 – Retorno da Análise de um dos Conjuntos

Fonte: Autoria Própria

Tabela 5 – Tabela contendo o resultado dos experimentos com frases contendo até 100 palavras e entropia 1

	Verdadeiro	Falso
Positivo	156	96
Negativo	402	328

Fonte: Autoria Própria

6.3 Análise dos Resultados

Quando comparamos as tabelas 3 e 4 é possível perceber que a entropia 1 conseguiu uma melhor classificação de verdadeiros positivos que a entropia 1.5, porém ambos os valores são extremamente baixos. Também é possível notar que em ambas as tabelas a classificação das mensagens que não são *phishing* foi alta (com a entropia de 1.5 classificando melhor do que a entropia de 1).

Quando comparamos as tabelas 5 e 3 é possível notar uma melhora na classificação

Tabela 6 – Tabela contendo o resultado dos experimentos com frases contendo até 100 palavras e entropia 1.5

	Verdadeiro	Falso
Positivo	79	33
Negativo	465	405

Fonte: Autoria Própria

Tabela 7 – Tabela contendo o resultado dos experimentos com frases contendo até 150 palavras e entropia 1

	Verdadeiro	Falso
Positivo	297	164
Negativo	336	190

Fonte: Autoria Própria

Tabela 8 – Tabela contendo o resultado dos experimentos com frases contendo até 150 palavras e entropia 1.5

	Verdadeiro	Falso
Positivo	220	77
Negativo	423	267

Fonte: Autoria Própria

das mensagens que são *phishing* e uma piora na classificação daquelas que não são. O mesmo ocorre com as tabelas 6 e 4.

O padrão de melhoria na classificação das mensagens que são *phishing* e uma piora na classificação das que não são *phishing* se mantem entre as tabelas 7 e 9 e 8 e 10.

É possível observar pelas tabelas 11 e 12 que os melhores resultados das medições estão presentes dentro dos conjuntos com uma quantidade de palavras maior. Dentre os valores de entropia sugeridos, a entropia 1 possui o melhor desempenho no cálculo de *recall* (que mede a quantidade de classificações positivos que o modelo acertou), possuindo também uma precisão maior apesar dos valores baixo. Apesar desses resultados, a acurácia da entropia 1.5 é maior, sugerindo que essa entropia consegue classificar melhor os valores que não são ataques de *phishing*. Dentro do contexto do trabalho a classificação correta desses textos também é importante, uma vez que classificar como *phishing* algo que não é *phishing* tende a trazer ao usuário um sentimento de alerta e desconfiança desnecessários, o que

Tabela 9 – Tabela contendo o resultado dos experimentos com frases contendo até 200 palavras e entropia 1

	Verdadeiro	Falso
Positivo	333	230
Negativo	270	138

Fonte: Autoria Própria

Tabela 10 – Tabela contendo o resultado dos experimentos com frases contendo até 200 palavras e entropia 1.5

	Verdadeiro	Falso
Positivo	275	124
Negativo	376	196

Fonte: Autoria Própria

Tabela 11 – Tabela contendo o resultado dos experimentos com entropia 1

Conjunto	Acurácia	F1-Score	Recall	Precisão
50	0,5133744856	0,118974359	0,1218487395	0,1162
100	0,5682281059	0,2994241843	0,3223140496	0,2796
150	0,641337386	0,5303571429	0,6098562628	0,4692
200	0,6210092688	0,6201117318	0,7070063694	0,5522

Fonte: Autoria Própria

Tabela 12 – Tabela contendo o resultado dos experimentos com entropia 1.5

Conjunto	Acurácia	F1-Score	Recall	Precisão
50	0,5216049383	0,03458799593	0,1632231405	0,0335
100	0,5539714868	0,1536964981	0,3223140496	0,1452
150	0,6514690983	0,389380531	0,4517453799	0,3421
200	0,6704428424	0,4901960784	0,5838641189	0,4224

Fonte: Autoria Própria

não só coloca em risco a percepção e confiança por parte dos usuários na ferramenta como pode causar confusão quando o usuário se deparar com situação análoga a da classificação.

O resultado do experimento retornou uma acurácia bem baixa e alguns dos fatores que podem ter influenciado esse resultado são:

- ❑ O *SetList* utilizado para realizar a Busca: Por se tratar de um conjunto de dados que foi montado pela autora do trabalho através de observações em textos de exemplo ele carece de alguns termos e rigores que tendem a trazer uma maior segurança e completude na hora de avaliar os dados de teste. Algumas palavras e tópicos podem ficar fora do escopo de classificação e isso tende a impactar na hora de calcular a entropia da frase
- ❑ O valor de entropia utilizado: Uma vez que o *SetList* utilizado apresenta incompletude, isso tende a influenciar no cálculo da entropia e pode trazer ao estudo de inferência valores equivocados para a métrica utilizada

Conclusão

Este trabalho apresentou o desenvolvimento de uma solução para classificação de mensagens como tentativa de ataque de engenharia social. Foi possível construir um conjunto de palavras utilizadas dentro do contexto de ataques de *phishing*, *baiting* e *pretexting* e utilizar técnicas para receber uma entrada de texto e processar esse texto buscando cada palavra do conjunto e a sua frequência. Os resultados obtidos através das análises estatísticas permitiram observar melhorias propostas dentro da criação do conjunto, visam melhorar a ferramenta de classificação e trazer um valor melhor para a separação entre o que é considerado ataque de engenharia social e o que não é.

7.1 Principais Contribuições

As principais contribuições desse trabalho são a criação de um aplicativo para dispositivos móveis baseados em Android que recebe um texto e o classifica como possibilidade de ataque de engenharia social ou não e a criação de um *SetList* contendo um conjunto inicial de palavras de interesse para identificar um ataque de engenharia social do tipo *phishing* disponibilizado em: <<https://github.com/Bianca-ataide/SetList-de-Palavras-Chave>>.

7.2 Trabalhos Futuros

Se mostra necessário refazer o cálculo de inferência para encontrar um valor melhor para a entropia. Também se mostra interessante a realização de medições de desempenho da Central Processing Unit (CPU) durante a realização da classificação dos textos, afim de observar as diferenças acerca da classificação de cada conjunto. Outra abordagem interessante seria a utilização de similaridade de cossenos como forma de classificação no lugar da entropia.

Outra proposta de melhoria é o desenvolvimento de um modelo que faz a captura dos textos na tela do usuário em tempo real e analisa a possibilidade de conter conteúdo malicioso.

Referências

- ALKAHTANI, H.; ALDHYANI, T. H. Artificial intelligence algorithms for malware detection in android-operated mobile devices. **Sensors**, MDPI, v. 22, n. 6, p. 2268, 2022. Citado na página 11.
- AMADO, P. G. F. Bancos de dados móveis: visão geral, desafios e soluções atuais. **Centro de Informática. Universidade Federal de Pernambuco**, 2002. Citado na página 19.
- ASLAN, Ö. et al. A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. **Electronics**, MDPI, v. 12, n. 6, p. 1333, 2023. Citado na página 16.
- BARBOSA, J. S. et al. A proteção de dados e segurança da informação na pandemia covid-19: contexto nacional. **Research, Society and Development**, v. 10, n. 2, p. e40510212557–e40510212557, 2021. Citado na página 17.
- BLAZZOTTO, F.; ANTICOLI, F. A.; BOCCIA, G. V. Uso de tecnologias para minimizar ataques de engenharia social em ambientes corporativos. 168, 2020. Citado 2 vezes nas páginas 23 e 24.
- CARPES, R. **Engenharia social responde por 40% dos roubos de credenciais de identidade**. Convergência Digital, 2023. Acessado em: 13 de novembro de 2024. Disponível em: <<https://convergenciadigital.com.br/especial/febraban-tech-2023/engenharia-social-responde-por-40-dos-roubos-de-credenciais-de-identidade/>>. Citado na página 12.
- JEMEROV, D.; ISAKOVA, S. **Kotlin in action**. New York: Simon and Schuster, 2017. Citado na página 21.
- LECHETA, R. R. **Google Android-3ª Edição: Aprenda a criar aplicações para dispositivos móveis com o Android SDK**. São Paulo: Novatec Editora, 2013. Citado na página 19.
- MACHADO, F. N. R. **Segurança da informação: princípios e controle de ameaças**. São José dos Campos: Saraiva Educação SA, 2014. Citado na página 16.
- MAGOSSI, J. C.; BARROS, A. C. da C. A entropia de shannon: uma abordagem axiomática. **REMAT: Revista Eletrônica da Matemática**, v. 7, n. 1, p. e3013–e3013, 2021. Citado na página 31.

- MITNICK, K. D.; SIMON, W. L. **The art of deception: Controlling the human element of security**. Indianapolis: John Wiley & Sons, 2003. Citado na página 17.
- MOSKALA, M.; WOJDA, I. **Android Development with Kotlin**. Birmingham: Packt Publishing Ltd, 2017. Citado na página 21.
- PEREIRA, L. C. O.; SILVA, M. L. D. **Android para desenvolvedores**. Rio de Janeiro: Brasport, 2009. Citado na página 19.
- REIS, A. P. d. et al. Análise de vulnerabilidades de segurança em sistemas de internet banking utilizando ferramentas de código aberto. Universidade Federal de Uberlândia, 2018. Citado 3 vezes nas páginas 12, 23 e 24.
- REZENDE, F. N. Taxonomia de malwares de dispositivos móveis. Universidade Federal de Uberlândia, 2023. Citado 2 vezes nas páginas 23 e 24.
- SHABTAI, A. et al. Google android: A comprehensive security assessment. **IEEE Security & Privacy**, IEEE, v. 8, n. 2, p. 35–44, 2010. Citado na página 20.
- SILVA, E. V. F.; SOUSA, L. C.; JÚNIOR, I. G. Principais ataques em empresas utilizando a engenharia social. In: **FatecSeg-Congresso de Segurança da Informação**. Ourinhos: FatecSeg, 2023. Citado 2 vezes nas páginas 18 e 19.
- SOUZA, R. C. de; FERNANDES, J. H. C. Um estudo sobre a confiança em segurança da informação focado na prevenção a ataques de engenharia social nas comunicações digitais. **Brazilian Journal of Information Science**, Universidade Estadual Paulista, v. 10, n. 1, p. 63–75, 2016. Citado 2 vezes nas páginas 23 e 24.
- TAKEUCHI, E. C. Fatores humanos em cibersegurança uma revisão sistemática da literatura. Universidade Federal de Uberlândia, 2023. Citado 2 vezes nas páginas 23 e 24.
- TIESO, I. H. de S.; SANTO, F. do E. Ataques de engenharia social. **Revista Interface Tecnológica**, v. 17, n. 2, p. 206–218, 2020. Citado 5 vezes nas páginas 11, 16, 17, 18 e 19.
- VOLPINI, A. **70% dos golpes digitais estão vinculados à captura de dados pessoais por engenharia social**. FEBRABAN, 2021. Acessado em: 13 de novembro de 2024. Disponível em: <<https://portal.febraban.org.br/noticia/3367/pt-br>>. Citado na página 12.
- WANG, Z. et al. Threat detection for general social engineering attack using machine learning techniques. **arXiv preprint arXiv:2203.07933**, 2022. Citado na página 11.
- WELIVESECURITY. **Brasil é o segundo país da América Latina com mais detecções de ataques de engenharia social**. WeLiveSecurity, 2021. Acessado em: 13 de novembro de 2024. Disponível em: <<https://www.welivesecurity.com/br/2021/01/07/brasil-e-o-segundo-pais-da-america-latina-com-mais-deteccoes-de-ataques-de-engenharia-social/>>. Citado na página 12.