Uma Avaliação das Prevenções de Phishing em Navegadores Web

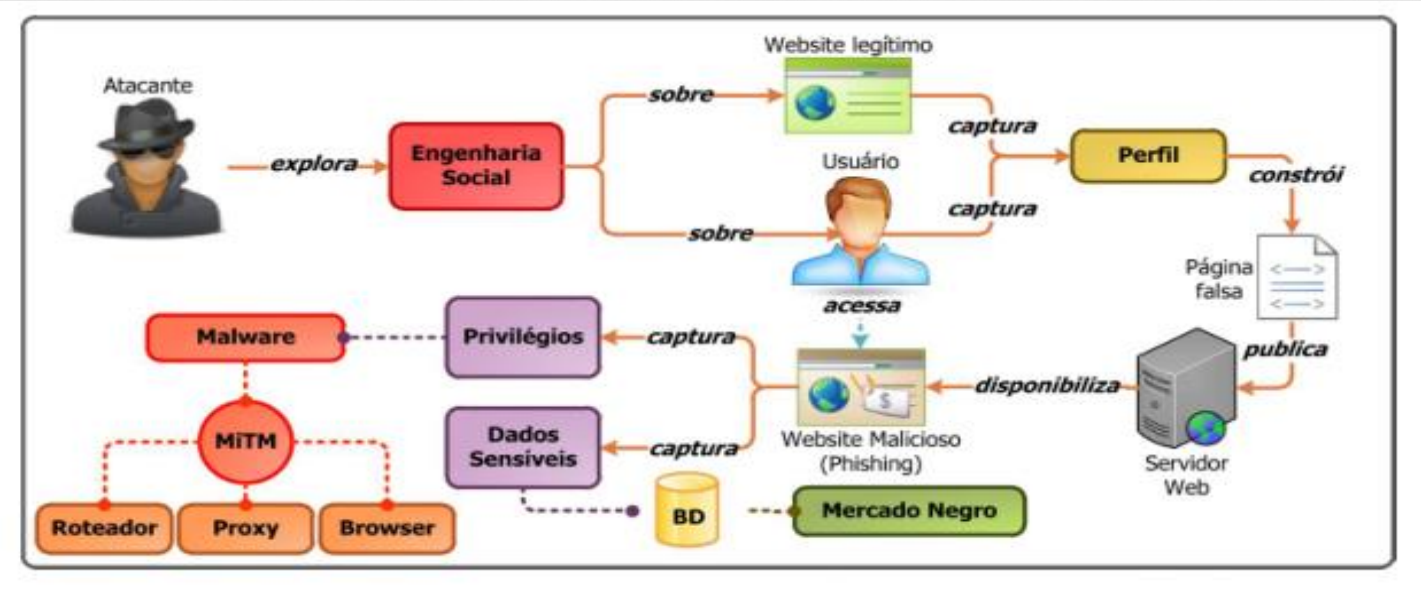
Carlo M. R. da Silva, Eduardo L. Feitosa, Vinicius C. Garcia

Juan Hermann

Matheus Tonial

Introdução

- Phishing é um ataque para obter informações usando engenharia social, o mais comum é spams via e-mail que redirecionam para sites maliciosos.



Cenário de ataque

- O tipo de ataque mais frequente é o da pagina falsa, onde o atacante desenvolve uma pagina “clone” da original, com isso ele pode capturar os dados submetidos pelo usuário.

- **Punycode** é quando usa-se caracteres de outra língua com aparência idêntica a do usuário, mascarando a URL para se assemelhar a do site original.

Cenário de ataque

- **Domínio de Topo**, há organizações que utilizam mais de um endereço (bradesco.b.br ou banco.bradesco) e com isso deixa uma oportunidade para quem resolve atacar, que pode criar uma URL semelhante a original.
- **Spear Phishing**, um phishing direcionado, ao invés de criar uma pagina popular com frequente acesso, é feito uma pagina direcionada a certo grupo/organização com um conteúdo mais convincente.

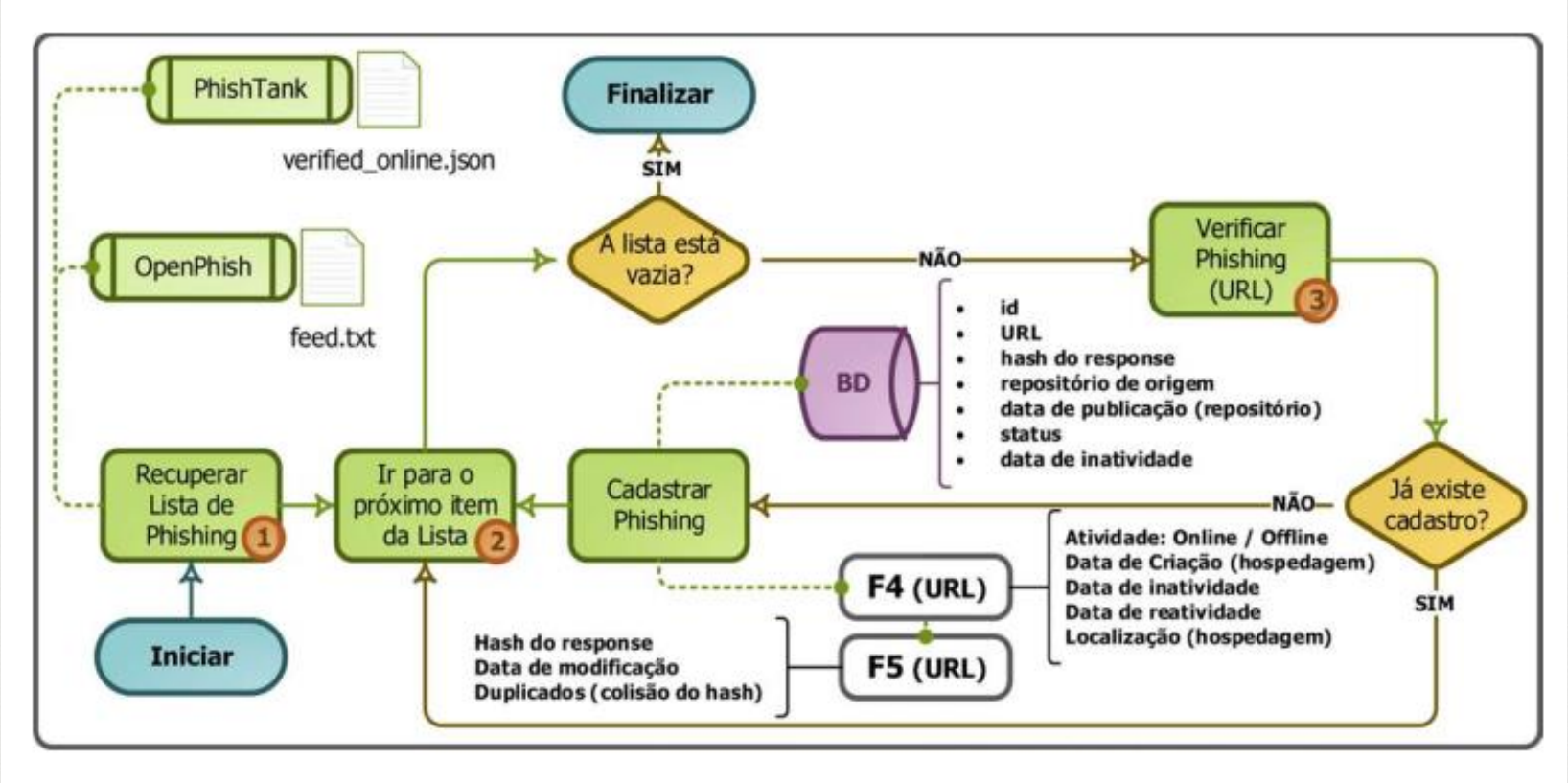
Proposta

- Para investigar sobre a proteção ao phishing fornecida pelos navegadores, não foi considerado extensões ou antivírus, somente funcionalidades originais do próprio navegador.
- A analise foi dividida em três partes, primeiro como obter uma quantidade relevante de phishing na web, seguido de expor o numero de phishing ao sistema de proteção dos navegares web e por fim analisar o padrão de cada phishing.

Como obter uma quantidade relevante de phishing na web

- Foi realizado buscas por repositórios de phishing que estão disponíveis ao publico e com atualização constante, com o objetivo de criar um cenário similar ao real.
- Foram utilizados os repositórios: PhishTank e OpenPhish (foi avaliado também os repositórios: VirusTotal e SafeBrowsing, mas descartados por não ter a base de registros abertamente).

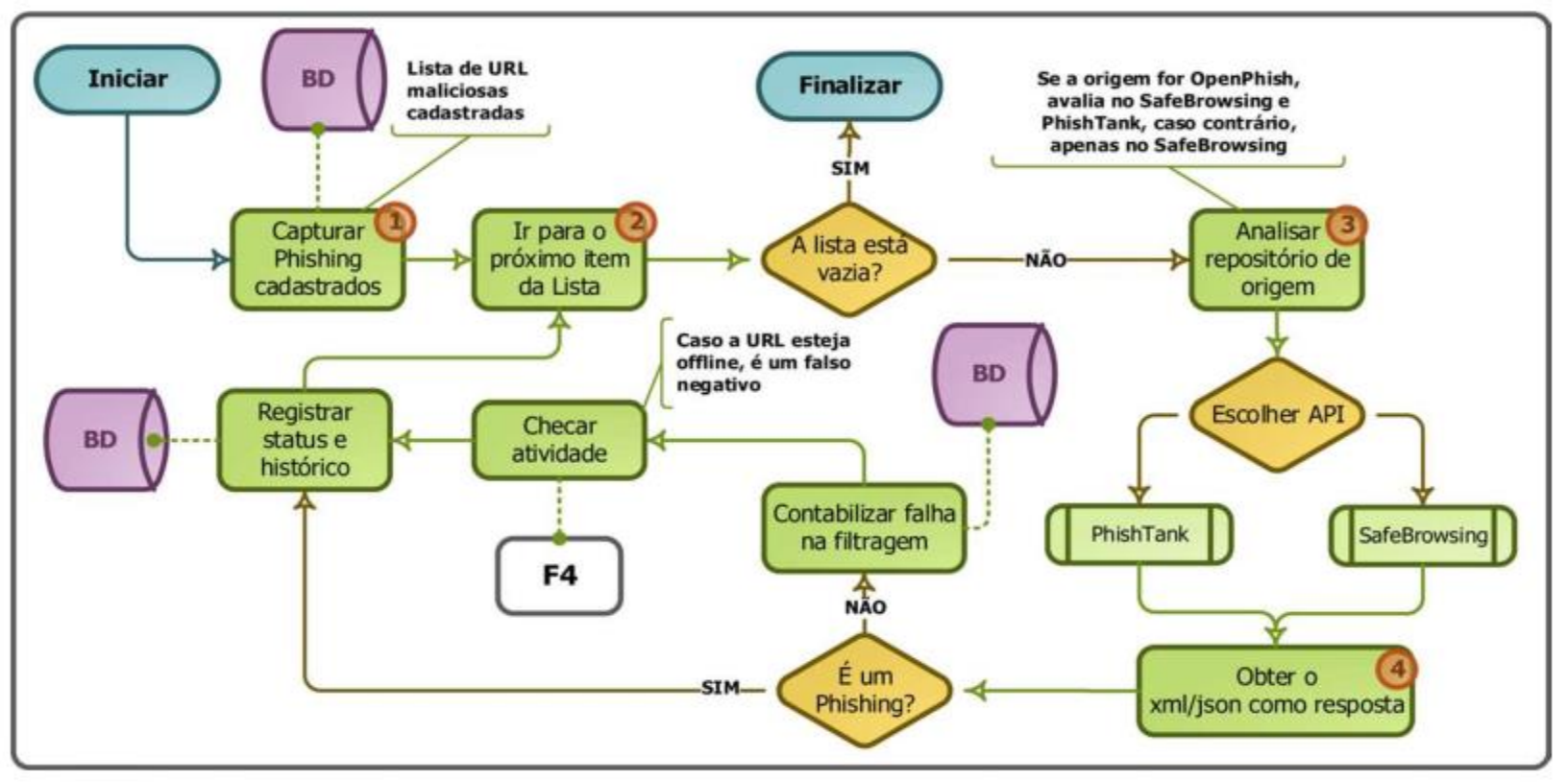
Como obter uma quantidade relevante de phishing na web



Expor o numero de phishing ao sistema de proteção dos navegares web

- Para o estudo foi considerados os navegadores: Google Chrome, Internet Explorer, Mozilla Firefox, Opera e Safari.
- Como o Chrome, Firefox e Safari tem o mesmo sistema de proteção (SafeBrowsing API) o resultado será contado em conjunto.
- O Opera utiliza o PhishTank API e o Explorer o Windows Defender SmartScreen.

Expor o numero de phishing ao sistema de proteção dos navegares web



Analisar o padrão de cada phishing

- Por fim foi efetuada uma analise para identificar os padrões de propagação, para isso as URL obtidas na primeira etapa foram armazenadas em banco de dados relacional, para que posteriormente possa ser realizadas buscas através de consultas SQL.

Resultados obtidos

Tabela 1. Phishing obtidos nos repositórios

Repositório	Total de registros	Registros de 2017	Registros *Online* de 2017	Registros *Offline* de 2017
PhishTank	25.896	12.907	12.390	517
OpenPhish	3.776	3.776	3.748	28
-	**29.672**	**16.683**	**16.138**	**545**

- Para o teste foi considerado apenas o primeiro semestre de 2017 e que estiverem online.

Resultados obtidos

Tabela 2. Avaliação das API em confronto aos repositórios

Repositório	API	Total de *phishing*	Identificados	% Acerto
PhishTank	*SafeBrowsing* API	12.390	7.824	63,15%
OpenPhish	*PhishTank* API	3.748	2.572	68,62%
OpenPhish	*SafeBrowsing* API	3.748	2.265	60,43%

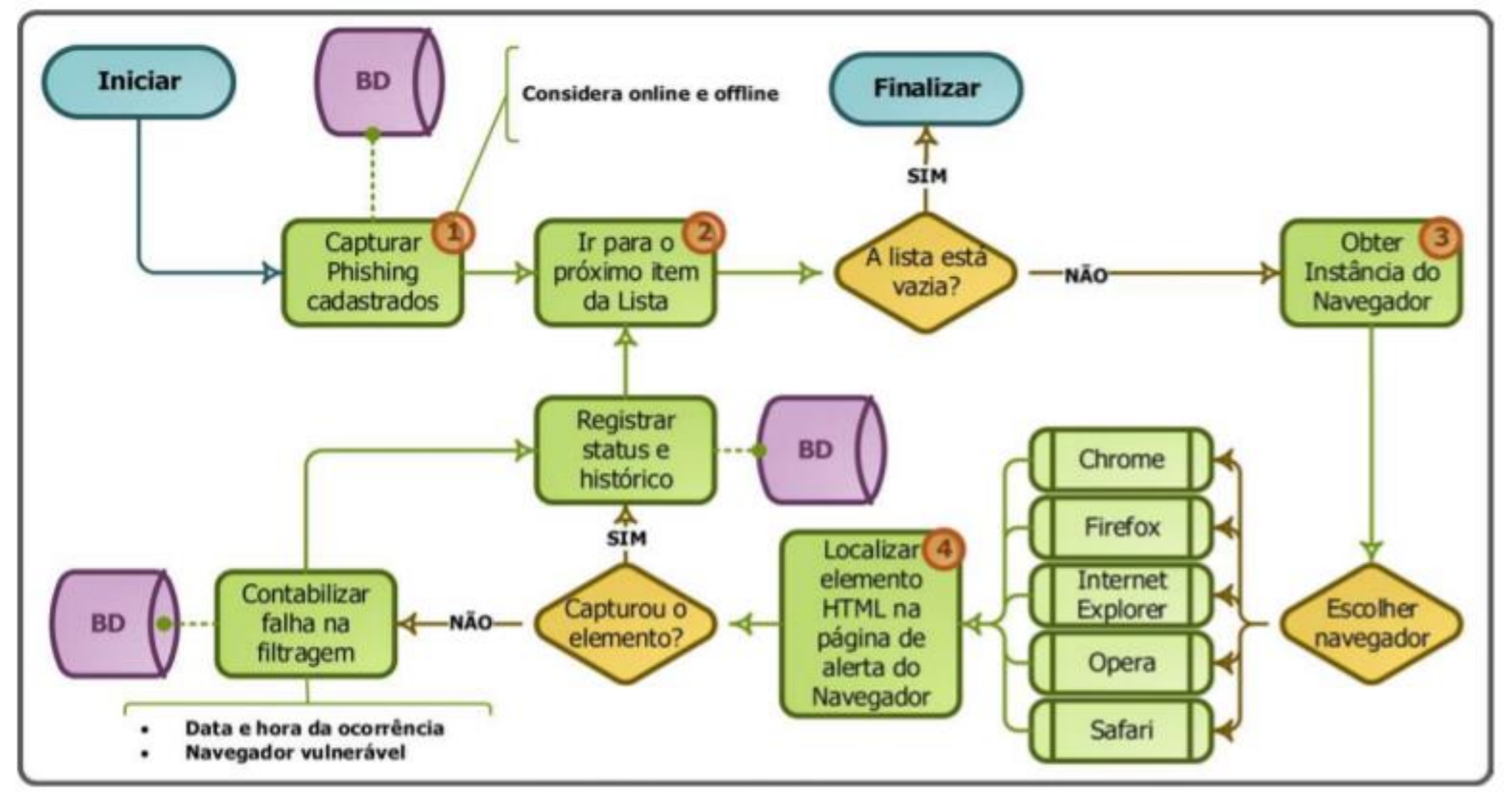
Resultados obtidos

- Foi notado diferenças entre o teste direto na API e no navegador:
- No Opera e Firefox, algumas URL não eram reconhecidas como phishing e sim como maliciosas.
- Outra diferença é que no teste direto na API não é tratado as URL que são encurtadas, já no navegador elas passariam pelo site redirecionando o usuário para o destino.
- Houve diferenças até mesmo entre o Chrome, Firefox e Safari, os quais possuem o mesmo mecanismo de proteção.

Resultados obtidos

- Um novo fluxograma foi proposto, com instancia do navegador, assim pode ser filtrado phishing independente de estar online ou offline.

Resultados obtidos



Resultados obtidos

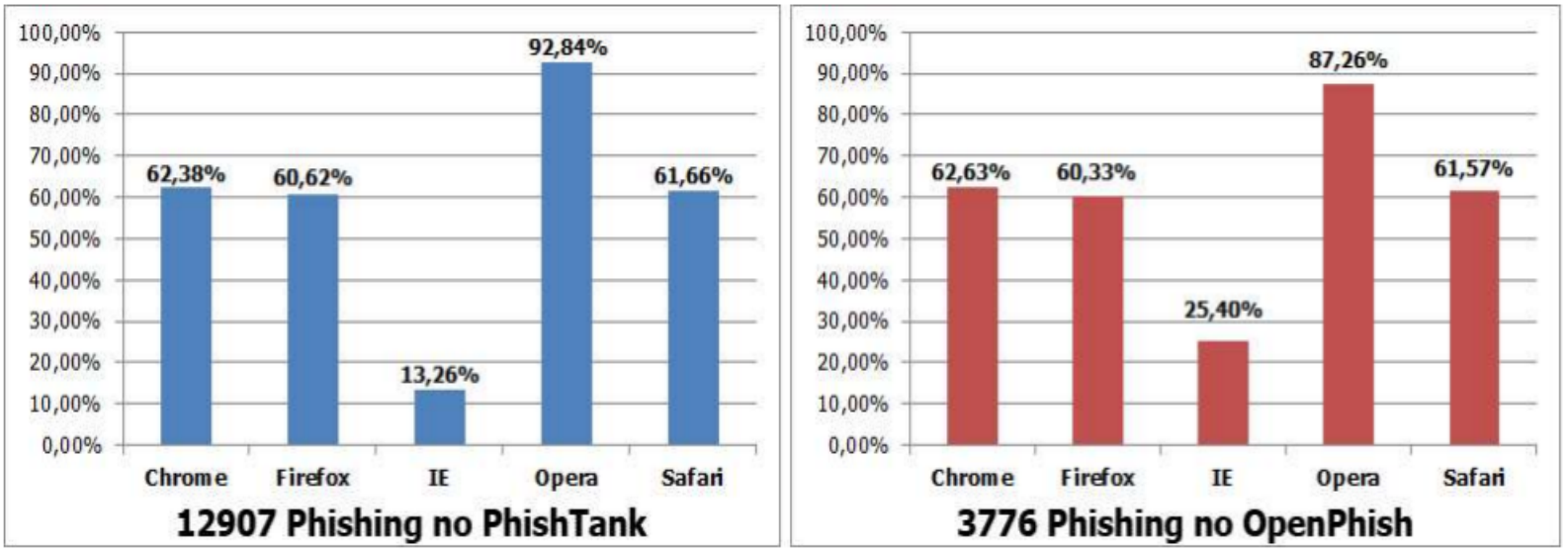


Figura 7. Resultados obtidos na filtragem dos navegadores

Resultados obtidos

- Por fim, foi analisado o padrão de palavras chave nas URL, tendo como resultado a seguinte tabela:

Tabela 3. Análise dos padrões dos *phishing* (na URL)

Palavra-chave	Ocorrências	Palavra-chave	Ocorrências	Palavra-chave	Ocorrências
login	3.596	domínios .br	1.590	domínios .org	1.345
https	1.266	google	916	paypal	596
card	529	mobile	368	boleto	302
banco	291	facebook	287	cadastr*	269
download	217	payment	204	billing	182
cert*	168	santander	154	banking	136

Conclusão

- Nenhum navegador teve mais de 90% de acerto quando comparado com um repositório diferente do seu.
- O Opera, quando avaliado com o OpenPhish, teve o desempenho de 87,26%, já a sua API, PhishTank, avaliando com o OpenPhish teve um resultado de 68,62% de acerto.
- Avaliando o Opera com seu respectivo repositório foi registrado 92,84% de acerto.
- O Internet Explorer teve a menor taxa de eficiência, com 13,26% no PhishTank.

Conclusão

- Como trabalhos futuros, pretende-se desenvolver uma analise morfológica, para acompanhar o comportamento de cada phishing e uma analise do seu ciclo de vida.
- Para a futura analise do ciclo de vida foi desenvolvido o fluxograma a seguir, com o objetivo de monitorar o tempo de criação e atividades posteriores.

Conclusão

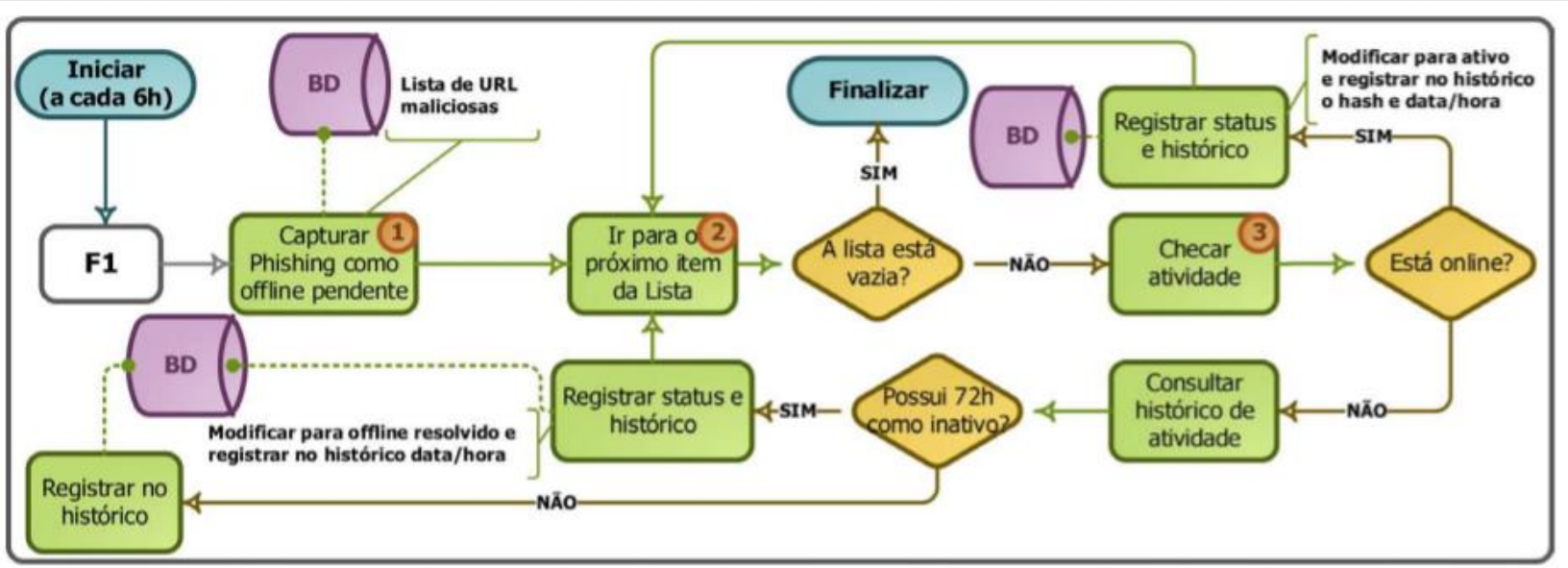


Figura 8. Fluxograma da análise de ciclo de vida do *phishing* F4

Conclusão

- Para a analise morfológica, com o intuito de monitorar as modificações sofridas pelo phishing durante um período, foi desenvolvido o seguinte:

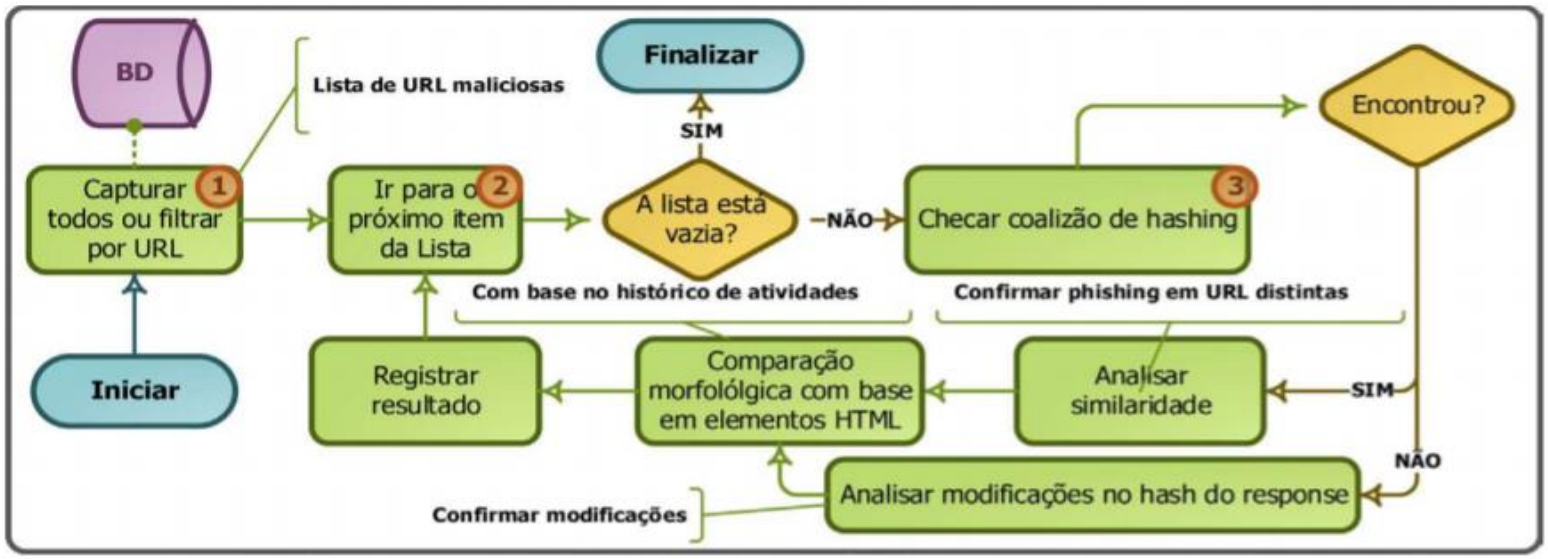


Figura 9. Fluxograma da análise de morfologia do *phishing* F5