



Análise de dados administrativos de e-commerce: Uma abordagem focada no cliente

Brayner Cristian Mello Carvalho, Nilson M. Lazarin

¹Bacharelado em Sistemas de Informação – Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (Cefet/RJ) – Nova Friburgo, RJ – Brazil

brayner.carvalho@aluno.cefet-rj.br, nilson.lazarin@cefet-rj.br

Abstract. *With the increase of e-commerce, the most diverse virtual scams were created and spread throughout the Internet. Furthermore, deciding whether a store is safe for online shopping is not always easy. Then, the user must have tools at his disposal to verify the legitimacy of the e-commerce. Therefore, this work presents an approach to identify administrative data on e-commerce sites with Top-Level Domains “.BR”. Besides, a browser extension was developed that crosses the legal entity data with the domain registration. With this, this paper aims to contribute to users’ decisions about e-commerce.*

Resumo. *Com o aumento do e-commerce, os mais diversos tipos de golpes virtuais são propagados pela Internet. Nem sempre é fácil decidir se uma loja é ou não segura para efetuar compras online. Portanto, é ideal que o usuário possua ferramentas para verificar a legitimidade de um site. Diante disso, este trabalho apresenta uma abordagem de identificação de dados administrativos em sites de e-commerce com TLD (top-level domain) “.BR”. Para validar a proposta foi desenvolvida uma extensão de navegador que cruza os dados de pessoa jurídica com os dados do registro do domínio. Com isso, o trabalho espera contribuir com a decisão dos usuários sobre sites de e-commerce.*

1. Introdução

O aumento do *e-commerce* mostra ser uma tendência crescente na sociedade atual, principalmente quando comparado com as taxas e custos de manutenção das lojas em seus ambientes físicos [Chagas 2017]. A possibilidade e o conforto de se receber os produtos em seu lar, além da melhoria considerável no valor e nos prazos de entrega dos produtos, foi desenvolvendo um hábito de consumo crescente nas pessoas, justificando essa tendência de aumento [Fernandes 2020]. Aumento esse acelerado nos últimos anos, entretanto, os clientes, muitas vezes, não gozam do conhecimento e das boas práticas necessárias para proteger a si e aos seus dados na internet [Guilherme et al. 2021].

Infelizmente, também foi notado um aumento considerado na quantidade e na variedade de golpes utilizados via internet [Mota 2021]. As pessoas, que não foram ensinadas ou orientadas a se inteirar sobre formas de se prevenir desses golpes, acabaram sendo alvos fáceis das investidas de criminosos que, utilizando-se dos mais variados tipos de golpes como, por exemplo, o *phishing*, sendo a forma mais popular de se capturar os dados das vítimas, e o *keylogger*, que consiste em uma ferramenta de monitoramento do teclado, obtendo e armazenando tudo que o usuário digita, inclusive os seus dados sensíveis [CERT.br 2012]. Na maioria desses golpes, a pessoa nem sequer tem ciência

de que foi capturada. As que conseguem saber, muitas vezes, não sabem muito bem qual o procedimento correto a se tomar, já que, até existem iniciativas e medidas a serem realizadas em caso de sofrer algo do gênero, mas isso não é nada claro ao público leigo [Rocha et al. 2022].

É importante que o usuário tenha ciência da loja na qual está comprando, bem como na sua reputação. Atualmente, uma alternativa comum para tentar auxiliar o cliente neste ponto é a utilização de selos de reputação. Mas, esses selos se mostraram pouco confiáveis e muito inconsistentes em seus dados [Carvalho et al. 2022]. Dessa forma, a falta de ferramentas de uso facilitado para auxiliar o usuário na busca de informações seguras sobre sites de *e-commerce* é um agravante para a questão.

Diante disso, este trabalho apresenta uma abordagem de identificação de dados administrativos em sites de *e-commerce* com TLD (*top-level domain*) “.BR”. Para tal, é apresentado o desenvolvimento e uma prova de conceito com uma extensão de navegador que busca apoiar a decisão do usuário, confrontando informações do domínio com informações da empresa.

Este trabalho está organizado da seguinte forma: na Seção 2 são apresentados os conceitos necessários para o entendimento da proposta; a proposta e a implementação são apresentadas na Seção 3. uma prova de conceito é apresentada na Seção 4; Por fim, na Seção 5 é apresentada uma discussão e os trabalhos futuros.

2. Fundamentação Teórica

Os *Domain Name System* (DNS) é um sistema UDP (*User Datagram Protocol*) utilizado para traduzir nomes de domínio em endereços IP (*Internet Protocol*). Cada site tem um nome e um número na rede, indicando onde está hospedado. Para fins de conforto ao usuário, o DNS elimina a necessidade de memorização desses endereços números extensos (IPv4), ou até mesmo do código alfanumérico (IPv6) [Mockapetris 1987][Ksinant et al. 2003].

O WHOIS é um protocolo TCP (*Transmission Control Protocol*) utilizado para consultas de dados sobre o registro de domínios que prevê o uso de um servidor WHOIS para atender às solicitações de clientes WHOIS na porta TCP 43 [Daigle 2004]. No Código 1 é apresentada uma solicitação, realizada via terminal, a um servidor. O cliente envia o nome de um domínio e o servidor retorna as informações públicas do domínio.

Código 1. Exemplo de código de terminal

```
user@pc:~$ whois sbseg2023.ufjf.br
% Copyright (c) Nic.br
% The use of the data below is only permitted as described in full by the Use and Privacy Policy at https://registro.br/upp, being prohibited its distribution, commercialization or reproduction, in particular, to use it for advertising or any similar purpose.

domain: ufjf.br
owner: UNIVERSIDADE FEDERAL DE JUIZ DE FORA
ownerid: 21.195.755/0001-69
responsible: Infraestrutura de TI (CGCO/UFJF)
owner-c: SUCGC
tech-c: SUCGC
nsrserver: ns1.ufjf.br 200.131.19.101 2001:12f0:614:19::101
nsstat: 20230824 AA
nic-hdl-br: SUCGC
person: Suporte CGCO
e-mail: suporteinfra.cgco@ufjf.br

% Security and mail abuse issues should also be addressed to cert.br, http://www.cert.br/ , respectively to cert@cert.br and mail-abuse@cert.br
%
% whois.registro.br accepts only direct match queries. Types of queries are: domain (.br), registrant (tax ID), ticket , provider, CIDR block, IP and ASN.
```

3. Proposta

Um nome de domínio possui pelo menos um endereço IP associado. Entretanto, diferentes URLs – *Uniform Resource Locator* (formada por protocolo, domínio, porta e caminho até um recurso) podem redirecionar para um mesmo site. Essa prática é comum quando lojas desejam redirecionar o usuário para seu domínio nacional, ou quando mudam sua identidade e/ou nome, e aceitam provisoriamente o nome antigo e o nome novo.

Os domínios de TLD *.BR*, gerenciados pelo Núcleo de Informação e Coordenação do Ponto BR - NIC.br, possuem um cadastro público sobre seus proprietários, contendo CNPJ ou CPF, e-mail contatos administrativos, de cobrança, telefones e endereços físicos, entre outros. A abordagem proposta neste trabalho visa apresentar, de forma facilitada ao usuário de *e-commerce*, uma comparação entre os dados de propriedade do domínio e as informações de inscrição e situação cadastral da pessoa jurídica, contribuindo dessa forma com a tomada de decisão do usuário.

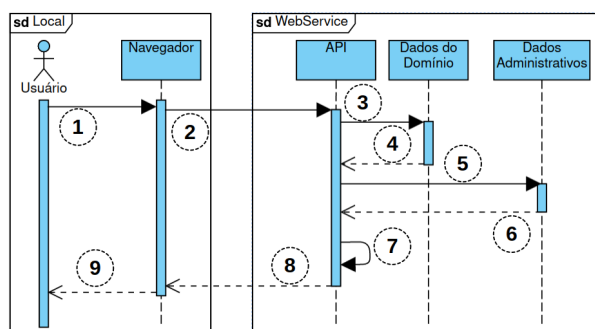


Figura 1. Diagrama de sequência resumindo o comportamento da aplicação.

Na Figura 1 são apresentadas as etapas da abordagem proposta, conforme descrito abaixo:

1. O Usuário informa a URL do *e-commerce* que deseja acessar;
2. O Navegador extrai o FQDN (*Fully Qualified Domain Name*) e faz uma requisição a um *WebService*;
3. A API (*Application Programming Interface*) informa o FQDN a um servidor WHOIS;
4. O servidor WHOIS retorna os dados de propriedade do domínio;
5. A API informa o CNPJ para um servidor de dados de pessoa jurídica;
6. O servidor as informações de inscrição e a situação cadastral da pessoa jurídicas;
7. A API compara os dados do domínio com os dados administrativos;
8. O *WebService* retorna ao navegador o resultado da comparação dos dados;
9. O Navegador exibe, de forma gráfica, o resultado ao Usuário.

3.1. Implementação

Para demonstrar a viabilidade da abordagem proposta, foi criada uma extensão¹ para o navegador Chrome, utilizando JavaScript, HTML e CSS. Foram utilizadas cores e dizeres

¹https://github.com/LabRedesCefetNF/Carvalho_2023_BrowserExtension

para informar ao usuário sobre a situação dos sites em que está navegando, de forma que, caso o cliente se depare com um *e-commerce* que possua dados inconsistentes, a extensão assumirá uma cor marcante. Logo, o usuário poderá analisar os dados e decidir se continua a compra.

Para atender às solicitações da extensão, foi desenvolvido uma API² que requisita dados dois *WebServices* especializados. Os dados dos domínios são obtidos por um servidor público³ do Registro.br, o departamento do NIC.br responsável pelas atividades de registro e manutenção dos nomes de domínios. Os dados administrativos são obtidos por um serviço gratuito⁴ da ReceitaWS, uma empresa que oferece uma API para obtenção de dados de pessoas jurídicas, advindos diretamente da Receita Federal. A implementação focou especificamente em domínios *.COM.BR*. Uma vez, de posse dos dados de CNPJ, endereço, conta responsável, telefone, criação e validade do domínio, estes são comparados com os dados de inscrição, situação, telefones de contato, e-mail e endereço do Cadastro Nacional de Pessoa Jurídica.

Na Figura 2 são apresentadas as três possibilidades de informações exibidas ao usuário. Caso não existam dados faltantes ou ocultos, a extensão assumirá cor verde. Se houver apenas uma ausência, assumirá o amarelo. E se existir pelo menos 3 ausências de informações, assumirá o vermelho.

INFORMAÇÕES RELEVANTES SOBRE O SITE .COM.BR	INFORMAÇÕES RELEVANTES SOBRE O SITE .COM.BR	INFORMAÇÕES RELEVANTES SOBRE O SITE .COM.BR
CRIAÇÃO DO DOMÍNIO 2011-07-05	CRIAÇÃO DO DOMÍNIO Não encontrado	CRIAÇÃO DO DOMÍNIO 2016-03-09
ÚLTIMA ALTERAÇÃO DE DOMÍNIO 2023-06-12	ÚLTIMA ALTERAÇÃO DE DOMÍNIO 2020-10-26	ÚLTIMA ALTERAÇÃO DE DOMÍNIO 2022-07-22
VENCIMENTO 2024-07-05	VENCIMENTO 2024-09-27	VENCIMENTO 2025-03-09
DATA - SITUAÇÃO 24/09/2009	DATA - SITUAÇÃO 02/04/2012	DATA - SITUAÇÃO Não encontrado
CNPJ 00.000.000/0001-69	CNPJ 15.000.000/0001-03	CNPJ 00.000.000/239-40
CNPJ ATIVO? Sim	CNPJ ATIVO? Sim	CNPJ ATIVO? Não
TELEFONE (11) 3000-0000	TELEFONE (11) 0000-2000	TELEFONE Não encontrado
EMAIL tributario@receita.com.br	EMAIL compliance@receita.com.br	EMAIL Não encontrado
ENDEREÇO Rua da Assembleia, 100 - Centro, São Paulo - SP, 05001-000	ENDEREÇO Rua da Assembleia, 100 - Centro, São Paulo - SP, 05001-000	ENDEREÇO Rua da Assembleia, 100 - Centro, São Paulo - SP, 05001-000

Figura 2. Informações exibidas ao usuário. Da esquerda para a direita são apresentados os casos de: sem dados faltantes; faltando um ou dois dados; faltando três ou mais dados.

4. Prova de Conceito

Para analisarmos a abordagem, propomos uma prova de conceito com o seguinte cenário: Um usuário sem conhecimento sobre como obter dados do WHOIS ou da Receita Federal, que deseja um determinado produto. Este utiliza um navegador, com a extensão previamente instalada, e por meio de um buscador, pesquisa sobre determinado produto. Definindo sua escolha baseada apenas no preço ele navega por lojas desconhecidas.

²https://github.com/LabRedesCefetNF/Carvalho_2023_WebService

³<https://registro.br/rdap/>

⁴<https://developers.receitaws.com.br/>

Neste cenário espera-se que a extensão possa ajudar o usuário a decidir sobre a confiabilidade da loja. A prova de conceito foi realizada em duas máquinas virtuais executando o Debian GNU/Linux 11 (bullseye) e ambiente gráfico LXDE com a seguinte configuração: CPU (1 socket, 2 cores), 2GB de memória RAM. As duas máquinas estão executando em uma mesma rede local. A primeira executa um servidor NodeJS (v12.22.12) e a API desenvolvida, atendendo às requisições da extensão na porta TCP 3000. A segunda executa o Google Chrome (108.0.5359.124 (Versão oficial) 64 bits) para o usuário realizar o acesso. A Figura 3 apresenta a realização da prova de conceito.

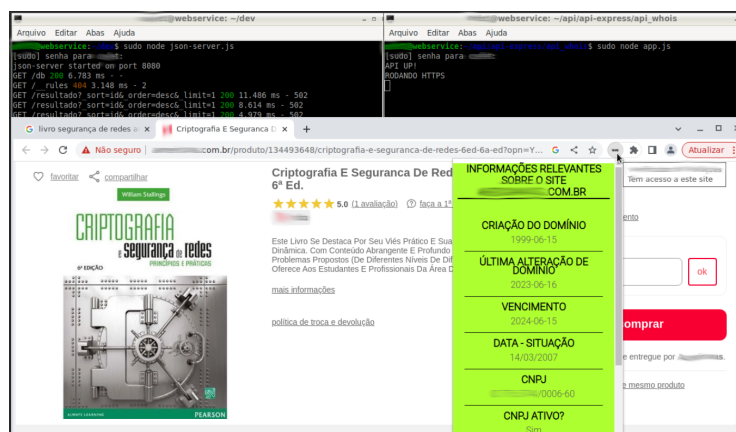


Figura 3. Execução da prova de conceito para análise da abordagem, através do cenário proposto.

5. Discussão

Vale ressaltar que mesmo com a utilização da abordagem, a decisão de continuar no *e-commerce* ainda depende do usuário. Além disso, mesmo que a extensão informe que todos os dados estão corretos, isso não garante o sucesso da compra. Entretanto, espera-se que esta abordagem possa contribuir com a redução dos riscos de falsos sites de *e-commerce*. Como o objetivo é fornecer informações para a tomada de decisão, optou-se por uma extensão sem pop-ups. Caso o usuário deseje, ele deve clicar no ícone da extensão, no canto superior direito do navegador, para verificar resultados obtidos.

Durante a implementação, foram notadas algumas limitações. Uma grande dificuldade foi a necessidade de adaptar o código para se trabalhar com recursos assíncronos. Além disso, pelo fato de ser uma extensão, todas as requisições feitas pela aplicação eram interpretadas pelo navegador como requisições efetuadas pelo site que se estava navegando no momento. Ou seja, caso a loja acessada fosse a *meusitedecompras.com.br*, o navegador considerava a requisição como sendo feita pelo próprio site, já que a extensão se torna parte do todo. Isso gerou muitos problemas de CORS (*Cross-origin Resource Sharing*) e de SSL (*Secure Sockets Layer*).

Para driblar a limitação de CORS, se fez necessária a criação da API para, ao invés da extensão requisitar diretamente para um servidor WHOIS ou um servidor da Receita Federal, a extensão solcita para a API desenvolvida. Evitando dessa forma o comportamento de que o próprio site estava requisitando informações, e, assim, evitando o CORS. Já o SSL, é geralmente obtido via certificado ao contratar algum serviço de hospedagem. Como não hospedamos a API em nenhum serviço online, foi necessário criar

um certificado autoassinado, temporário, para os servidor e o navegador se comunicarem corretamente.

Os próximos passos desta pesquisa serão uma análise de aceitação da tecnologia, através da hospedagem do serviço e publicação da extensão em um serviço de aplicação online, como o Chrome Web Store. Para uma melhor aceitação, também será necessário implementar a abordagem para outros navegadores, como o Firefox, etc. Será necessário avaliar a viabilidade da adição de informações sobre a reputação dos portais de *e-commerce*, que podem ser obtidos do Consumidor.gov e ReclameAqui. Outra necessidade é o estudo da viabilidade de expansão para domínios além do TLD “.COM.BR”.

Referências

- Carvalho, B. C. M., Corrêa, G. E., Pinto, J. P. R. d. M. B., and Lazarin, N. M. (2022). Sistemas de reputação e sua influência: Um comparativo. In *Anais do XIII Computer on the Beach - COTB'22*, pages 306–308, Itajaí - SC. Universidade do Vale do Itajaí. <https://doi.org/10.14210/cotb.v13.p306-308>.
- CERT.br (2012). *Cartilha de Segurança para Internet*. Comitê Gestor da Internet no Brasil, São Paulo, 2ª edition. <https://cartilha.cert.br/livro/>.
- Chagas, R. B. d. (2017). *O aumento da participação do e-commerce no Brasil: uma mudança na preferência na forma de consumo*. Graduação, Universidade Federal Fluminense, Campos dos Goytacazes. <http://app.uff.br/riuff/handle/1/6145>.
- Daigle, L. (2004). *WHOIS Protocol Specification*. Number 3912 in Request for Comments. RFC Editor. doi.org/10.17487/RFC3912.
- Fernandes, D. (2020). Após descobrirem e-commerce na pandemia, 94% querem manter hábito de compra. *E-Commerce Brasil - Artigos e Dicas sobre comércio eletrônico*. <https://www.ecommercebrasil.com.br/noticias/ecommerce-manter-habito-coronavirus>.
- Guilherme, L., Ferreira, M., Fonseca, G. d., and Lazarin, N. M. (2021). Uma breve noção sobre o comportamento dos internautas em relação à segurança na rede. In *Anais da VII Escola Regional de Sistemas de Informação do Rio de Janeiro*, pages 1–7, Porto Alegre, RS. SBC. <https://doi.org/10.5753/ersirj.2021.16972>.
- Ksinant, V., Huitema, C., Thomson, D. S., and Souissi, M. (2003). *DNS Extensions to Support IP Version 6*. Number 3596 in Request for Comments. RFC Editor. <https://doi.org/10.17487/RFC3596>.
- Mockapetris, P. (1987). *Domain names - concepts and facilities*. Number 1034 in Request for Comments. RFC Editor. <https://doi.org/10.17487/RFC1034>.
- Mota, M. d. O. (2021). *Estudo de caso sobre segurança em e-commerce*. Graduação, Pontifícia Universidade Católica de Goiás, Goiânia. <https://repositorio.pucgoias.edu.br/jspui/handle/123456789/3206>.
- Rocha, H., Klein, E., Rimes, L., and Lazarin, N. (2022). Iniciativas de popularização da segurança da informação: Um survey. In *Anais Estendidos do XVIII Simpósio Brasileiro de Sistemas de Informação*, pages 37–40, Porto Alegre, RS, Brasil. SBC. https://doi.org/10.5753/sbsi_estendido.2022.222787.