Uma Primeira Análise do Ecossistema HTTPS no Brasil

Thiago Paim Escarrone[1,3], **Diego Kreutz**[1,2,3], **Maurício M. Fiorenza**[2,3]

[1]Laboratório de Estudos Avançados (LEA)
[2] Mestrado Profissional em Engenharia de Software (MPES)
[3] Universidade Federal do Pampa (UNIPAMPA)
thiago.escarrone@gmail.com,{Nome.Sobrenome}@unipampa.edu.br

Resumo. *O HTTPS é essencial para garantir a segurança (e.g., confidencialidade dos dados) das comunicações que utilizam o protocolo HTTP na Internet. Entretanto, apesar da crescente adoção do HTTPS, muitos sites ainda não implementam da maneira correta os certificados digitais e não suportam a versão mais atual do TLS, a 1.3. Neste trabalho é apresentado um primeiro levantamento sobre o estado do ecossistema HTTPS no Brasil. Para a análise, foram selecionados 44 sites de instituições financeiras e governamentais, incluindo o governo federal e governos estaduais. Os resultados apontam que mais de 90% dos sites utilizam ou suportam apenas versões antigas do TLS/SSL, que contém vulnerabilidades conhecidas e passíveis de exploração por agentes maliciosos.*

1. Navegação Segura (?) na Internet com HTTPS

A maioria dos usuários acredita que está utilizando um canal de comunicação seguro quando navega na Internet utilizando o HTTPS (*Hyper Text Transfer Protocol Secure*), isto é, HTTP (*HyperText Transfer Protocol*) sobre SSL (*Secure Sockets Layer*)/TLS (*Transport Layer Security*). Na verdade, os próprios navegadores contribuem para esta sensação de segurança quando apresentam um cadeado verde ao lado do endereço eletrônico, mais conhecido como URL (*Uniform Resource Locator*), que o usuário está acessando. Entretanto, pesquisas demonstram que muitas vezes o ecossistema do HTTPS é, diferentemente do imaginado (ou do senso comum), inseguro [Bokslag 2016, Frost et al. 2019, Samarasinghe and Mannan 2019]. Os desafios e problemas de segurança são muitos e podem ocorrer em diferentes partes do ecossistema, incluindo falhas na especificação ou implementação dos protocolos, falhas na configuração dos certificados digitais nos servidores Web, falhas na geração dos certificados digitais, vulnerabilidades na Infraestrutura de Chaves Públicas (ICP), entre outras vulnerabilidades [Bokslag 2016, Frost et al. 2019, Samarasinghe and Mannan 2019, Matsumoto and Reischuk 2015, Merzdovnik et al. 2016].

Quando o navegador inicia uma conexão HTTPS, é realizado o *handshake* do TLS entre navegador e o servidor Web. Em seguida, o servidor apresenta o seu certificado digital X.509[1], que é utilizado para atestar que uma organização é realmente quem ela se diz ser através de uma chave pública. Um certificado é atestado e emitido por uma Autoridade Certificadora (AC)[2], que gera a assinatura digital. As ACs fazem parte das ICPs [Durumeric et al. 2013]. No Brasil, há infraestruturas de chaves públicas mantidas por instituições como o SERPRO (`https://www.serpro.gov.br/`) e a RNP (`https://www.rnp.br`).

[1]`https://tools.ietf.org/html/rfc2459`
[2]`https://tools.ietf.org/html/rfc5280.html`

Considerando um serviço online que utiliza HTTPS, como é possível obter mais informações sobre a segurança do certificado digital e das futuras conexões com o site? Que versões do SSL/TLS o site suporta? As versões suportadas pelo site possuem vulnerabilidades conhecidas? Existem estudos que tentam responder a este tipo de questões em contextos específicos, como sites da China, sites do Alexa Top 1 milhão e aplicativos bancários no Reino Unido [Samarasinghe and Mannan 2019, Vratonjic et al. 2013, Huang et al. 2019, Chothia et al. 2017]. Resultados das pesquisas apontam que ainda há um número significativo de sites e sistemas com vulnerabilidades relacionadas aos certificados digitais e versões de protocolos suportados e utilizados na prática.

O objetivo deste trabalho é realizar estudos similares sobre o ecossistema HTTPS no contexto do Brasil. Além disso, identificar as principais ferramentas disponíveis livremente para realizar as análises dos certificados digitais e protocolos suportados por sites governamentais e da iniciativa privada. Por exemplo, diferentemente dos trabalhos relacionados citados anteriormente, neste trabalho foram utilizadas ferramentas que proporcionam um diagnóstico mais detalhado, como a Geekflare TLS Scanner e a testssl.sh, que permitem identificar e catalogar falhas dos protocolos.

Existem diferentes ferramentas projetadas especificamente para analisar problemas na instalação de certificados digitais em sites e detectar falhas de protocolos[3]. Todas elas apresentam informações gerais a respeito da AC que emitiu o certificado, a sua validade e o algoritmo utilizado para assinar o certificado. Algumas ferramentas analisam também as propriedades específicas dos certificados, como a validade da cadeia do certificado e falhas conhecidas dos protocolos. Neste trabalho foram selecionadas seis ferramentas, cujas análises automatizadas e resultados são complementares, incluindo: SSL Labs[4], SSL Checker[5], Digicert[6], Wormly[7], Geekflare TLS Scanner[8] e testssl.sh[9]. A Tabela 1 resume o modo de operação e os parâmetros que são analisados por cada uma das ferramentas.

Tabela 1. Resumo das ferramentas utilizadas

	Modo de execução		**Parâmetros analisados**		
Ferramenta	**Navegador**	**Terminal**	**Informações gerais**	**Cadeia do certificado**	**Falhas no protocolo**
SSL Labs	✓		✓	✓	✓
SSL Checker	✓		✓	✓	
Digicert	✓		✓	✓	✓
Wormly	✓		✓		
Geekflare TLS	✓		✓		✓
testssl.sh		✓	✓	✓	✓

O restante do trabalho está organizado como segue. Os resultados das análises são apresentados na Seção 2. Na sequência, as considerações finais e os trabalhos futuros são apresentados na Seção 3

[3] https://geekflare.com/ssl-test-certificate/
[4] https://www.ssllabs.com/ssltest/
[5] https://www.sslshopper.com/ssl-checker.html
[6] https://www.digicert.com
[7] https://www.wormly.com/
[8] https://geekflare.com/ssl-test-certificate/
[9] https://testssl.sh

2. Análise do Ecossistema HTTPS no Brasil

As ferramentas relacionadas na Tabela 1 foram executadas em 44 (quarenta e quatro) sites, selecionados para esta primeira análise do ecossistema HTTPS no Brasil, divididos em quatro conjuntos: (a) sites importantes do governo federal (e.g., presidência da república); (b) site oficial dos 26 estados da federação mais o Distrito Federal; (c) principais sites da Prefeitura Municipal de Alegrete/RS; e (d) sites dos principais bancos e algumas *fintechs* brasileiras. Na primeira etapa, foram identificados que aproximadamente 33% dos sites (todos eles governamentais) não utilizam HTTPS. Para os 34 sites remanescentes, foram realizadas diferentes análises, conforme discriminado a seguir.

2.1. Análise dos certificados

Um certificado digital somente é considerado como confiável quando ele apresenta: (a) o nome da AC que o assinou; (b) o domínio ao qual pertence; e (c) a data de validade [Vratonjic et al. 2013]. Obviamente, as três informações devem estar corretas. As análises realizadas mostram que 26,47% dos sites quebram a cadeia de certificado, ou seja, a AC que assinou o certificado não é reconhecida pelo navegador como uma entidade confiável. Este é o caso dos sites governamentais que utilizam ACs de ICPs não reconhecidas pelos navegadores, como é o caso das ICPs do SERPRO e da RNP.

A cadeia do certificado também é quebrada quando a própria empresa assina o certificado, caracterizando-o como um certificado auto-assinado. Em 14,70% dos sites não existe uma AC confiável que assine o certificado, o que impede os navegadores de validar de forma automática e transparente a autenticidade do certificado.

Quando o nome do domínio do certificado difere do nome do domínio do site, o certificado também é identificado como não-confiável. Este caso ocorre em 20,58% dos certificados analisados. Já a data de validade do certificado também está contida no certificado e é apresentada da seguinte forma: *Not before and Not after*. Se for identificada qualquer data fora deste intervalo explícito, o certificado é considerado como expirado, como foi o caso de 8,82% dos certificados analisados.

Em resumo, 70,57% dos sites analisados, todos eles governamentais, apresentam algum problema relacionado ao certificado digital. Considerando a nova Lei Geral de Proteção de Dados Pessoais (LGPD), nº 13.709[10], que entra em vigor em 2020, fica evidente a urgência no processo de revisão e atualização dos certificados digitais utilizados nas instituições governamentais.

2.2. Versões dos protocolos SSL/TLS

A ferramenta testssl.sh foi utilizada para identificar as versões dos protocolos SSL/TLS suportadas por cada site. Como pode ser observado na Tabela 2, apesar de a versão mais recente do TLS (1.3) não apresentar nenhum tipo de ataque conhecido, 95,5% dos sites analisados ainda aceitam conexões utilizando as versões mais antigas e vulneráveis do protocolo. Novamente, o cenário apresentado é crítico pelo fato de a maioria dos sites (82,35%) ainda suportar versões bastante antigas do SSL/TLS, como a TLS 1.0.

[10] http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/Lei/L13709.htm

Tabela 2. Versões dos protocolos e respectivas vulnerabilidades

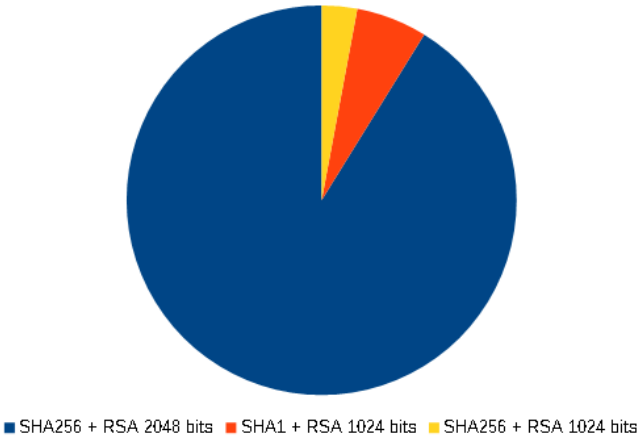
Versão	Sites que suportam	Ataques conhecidos
SSLv2	1	DROWN
SSLv3	6	POODLE, BEAST
TLS 1.0	28	BEAST
TLS 1.1	25	POODLE
TLS 1.2	29	Logjam
TLS 1.3	2	-

2.3. Algoritmos e funções criptográficas

A segunda parte do estudo consistiu em analisar os algoritmos e funções criptográficas utilizadas pelas unidades certificadoras para emissão dos certificados, incluindo algoritmos de assinatura, tamanho das chaves RSA (*Rivest, Shamir, and Adelman*) e funções de *hash* criptográfica.

A Figura 1 resume os algoritmos *hash* e os tamanhos das chaves RSA que as ACs utilizam para assinar os certificados analisados. Como pode ser observado, 91,17% certificados utilizam a combinação do algoritmo de *hash* SHA-256 (SHA2[11]) com o tamanho da chave RSA recomendada de 2048 bits[12]. Os demais certificados utilizam chaves RSA de 1024 bits, i.e., fora do tamanho recomendado. Além disso, 5,88% dos certificados utilizam ainda SHA1, o que é fortemente desaconselhado pela comunidade de segurança desde 2010.

Figura 1. Algoritmos de assinatura



2.4. Sites governamentais *versus* iniciativa privada

Tanto sites de órgãos governamentais como de empresas privadas permitem conexões utilizando versões dos protocolos SSL/TLS com vulnerabilidades conhecidas, expondo os usuários (e seus respectivos dados) a incidentes de segurança e privacidade.

[11] https://csrc.nist.gov/projects/hash-functions/nist-policy-on-hash-functions

[12] https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-57pt3r1.pdf

Sites do Governo Federal. De um total de 44 sites analisados, 6 deles são sites importantes do Governo Federal, como o site da presidência da república. As análises demonstram que 50% dos sites quebra a cadeia de certificado e em 16,66% deles o domínio do certificado não corresponde com o domínio acessado. Além disso, 83,33% dos sites são vulneráveis aos ataques BEAST e POODLE e 100% deles são suscetíveis ao ataque Logjam.

Sites dos governos estaduais. No contexto estadual, o cenário é ainda pior. Primeiro, 37,03% dos sites sequer utiliza HTTPS. Segundo, 22,22% possuem problemas relacionados ao domínio indicado no certificado, que não corresponde ao domínio acessado. Outros 22,22% dos sites possuem a cadeia de certificados quebrada e a maioria deles são certificados auto-assinados. Não bastante isso, mais de 50% dos sites suporta TLS 1.0, ou seja, são vulneráveis a ataques como o BEAST.

Sites da iniciativa privada. O cenário na iniciativa privada, considerando que foram analisados apenas sites de bancos e *fintechs*, também é bastante preocupante. Dos 9 sites analisados, 1 permite conexões utilizando o protocolo SSLv3, que possui várias vulnerabilidades conhecidas. Em 55,55% dos sites, os usuários podem estar em risco devido ao fato de suportar versões do SSL/TLS vulneráveis a ataques como o BEAST e o POODLE. De forma análoga ao cenário do governo federal, 100% dos sites analisados são suscetíveis ao ataque Logjam. Finalmente, apenas 22,22% dos sites suportam TLS 1.3, o que aumenta a segurança das comunicações para os usuários e instituições.

2.5. Impactos dos resultados apresentados

Os profissionais de tecnologia responsáveis pela manutenção dos certificados digitais devem compreender os riscos de segurança de um certificado mal configurado ou de baixa qualidade. Resumidamente, na Seção 2.1 foram discutidos os diferentes problemas de certificados de baixa qualidade ou mal configurados, o que impacta severamente a segurança do ecosistema HTTPS. Além disso, como apresentado na Tabela 2, há diferentes ataques conhecidos relacionados às diferentes versões do SSL/TLS, por exemplo.

A maioria absoluta dos sites analisados são vulneráveis ao ataque POODLE. Este ataque permite aos agentes maliciosos roubar *cookies*[13] de autenticação e obter informações sensíveis de usuários autenticados [Möller et al. 2014]. Um atacante pode também decifrar *cookies* do navegador do usuário e, consequentemente, roubar informações privadas através do ataque BEAST [Sarkar and Fitzgerald 2013], o que ocorre em grande parte dos sites analisados. Além disso, 65,90% dos sites investigados suportam o TLS 1.2, o que os torna suscetíveis a ataques que comprometem a integridade da comunicação [Bokslag 2016], isto é, ataques que permitem ao agente malicioso ler e modificar dados em trânsito.

3. Discussão e trabalhos futuros

A utilização de certificados digitais, para viabilizar conexões HTTPS, por si só, não garante segurança e privacidade aos usuários e instituições. Como visto nos resultados apresentados, a maioria absoluta dos sites analisados são vulneráveis a ataques conhecidos. Os problemas vão desde falhas de configuração dos sistemas até problemas mais específicos, inerentes à geração dos certificados digitais.

[13] `https://tools.ietf.org/html/rfc6265`

Os profissionais de tecnologia, responsáveis pela instalação dos certificados e manutenção dos sites, precisam estar minimamente cientes da importância e criticidade das suas tarefas. Por exemplo, um certificado digital qualquer, instalado sem o mínimo de cuidados técnicos, não traz segurança aos usuários do sistema. Como ocorre frequentemente, o uso de HTTPS leva a uma falsa sensação de segurança aos usuários e instituições. As análise realizadas neste trabalho, observando a utilização de certificados digitais e protocolos suportados nas conexões HTTPS do ecossistema da Internet do Brasil, mostram que ainda há um longo caminho pela frente para oferecer, efetivamente, segurança e privacidade aos usuários e instituições. Alternativas gratuitas para geração de certificados válidos e de qualidade, como o Let's Encrypt (`https://letsencrypt.org/`), podem representar uma alternativa atrativa para entidades que não possuem o conhecimento técnico ou os recursos financeiros para adquirir e gerenciar certificados de ACs tradicionais, reconhecidas pelos navegadores. Certificados emitidos pelo Let's Encrypt são reconhecidos e autenticados por quase todos os navegadores.

Como trabalhos futuros, podem ser destacados: (a) aumentar a abrangência do estudo, investigando um número maior de sites e setores da sociedade; (b) avaliar e comparar tecnicamente a qualidade das ferramentas existentes, como as listadas na Tabela 1, entre outras; e (c) analisar os principais sites do comércio eletrônico do Brasil.

Referências

Bokslag, W. (2016). The problem of popular primes: Logjam. *arXiv preprint arXiv:1602.02396*.

Chothia, T., Garcia, F. D., Heppel, C., and Stone, C. M. (2017). Why Banker Bob (still) Can't Get TLS Right: A Security Analysis of TLS in Leading UK Banking Apps. In *Int. Conf. on Financial Cryptography and Data Security*, pages 579–597. Springer.

Durumeric, Z., Kasten, J., Bailey, M., and Halderman, J. A. (2013). Analysis of the https certificate ecosystem. In *ACM IMC*, pages 291–304. ACM.

Frost, V., Tian, D. J., Ruales, C., Prakash, V., Traynor, P., and Butler, K. R. B. (2019). Examining DES-based Cipher Suite Support Within the TLS Ecosystem. In *2019 ACM AsiaCCS*, pages 539–546. ACM.

Huang, J., Zhang, Z., Li, W., and Xin, Y. (2019). Assessment of the impacts of TLS vulnerabilities in the HTTPS ecosystem of China. *Procedia computer science*, 147:512–518.

Matsumoto, S. and Reischuk, R. M. (2015). Certificates-as-an-insurance: Incentivizing accountability in ssl/tls. In *USENIX NDSS WSENT*.

Merzdovnik, G., Falb, K., Schmiedecker, M., Voyiatzis, A. G., and Weippl, E. (2016). Whom You Gonna Trust? A Longitudinal Study on TLS Notary Services. In *Data and Applications Security and Privacy*, pages 331–346. Springer.

Möller, B., Duong, T., and Kotowicz, K. (2014). This poodle bites: exploiting the ssl 3.0 fallback. *Security Advisory*.

Samarasinghe, N. and Mannan, M. (2019). Another look at TLS ecosystems in networked devices vs. Web servers. *Computers & Security*, 80:1 – 13.

Sarkar, P. G. and Fitzgerald, S. (2013). Attacks on SSL a comprehensive study of beast, crime, time, breach, lucky 13 & RC4 biases. `http://bit.do/ssl_survey`.

Vratonjic, N., Freudiger, J., Bindschaedler, V., and Hubaux, J.-P. (2013). The inconvenient truth about web certificates. In *Economics of information security and privacy iii*, pages 79–117. Springer.