HTTPS: STAYING PROTECTED ON THE INTERNET

Sang Ah Kim*

CITE AS: 1 GEO. L. TECH. REV. 119 (2016)
https://perma.cc/F4US-M82B

One may occasionally see five letters displayed at the beginning of a URL: HTTPS. To understand HTTPS and its importance in internet communication, it is first necessary to understand HTTP (Hypertext Transfer Protocol.) The difference of a single letter could contribute to the invasion of your privacy on the internet and the theft of your sensitive personal information.

HTTP is, at its core, a "protocol." A protocol dictates the structure of communication between an internet user and a website by establishing exactly how the two parties will exchange information and in what format.[1]An example of an internet communication is when a user clicks the title of an article on the Washington Post website–she is requesting that the website show her the article. The website takes the request and responds by displaying the article on the user's screen.[2] Think of a protocol as a set of rules that must be satisfied before two parties begin requesting and responding. To analogize, playing soccer has rules, such as the rule that players in general cannot use their hands to pass the ball. An internet communication would be like two players passing the ball to one another by kicking.

A challenge with online communications is the interception of requested information by a third party using a "man-in-the-middle," or MITM attack.[3] Similar to wiretapping, a MITM attack allows the an adversary to intercept the content of a user's information flows before it reaches its intended recipient, often without the user's knowledge. Data such as social security numbers and credit card numbers may be intercepted and end up on black market sites, where

[*] GLTR Staff Member; Georgetown Law, J.D. expected 2018; University of Georgia, B.A. 2014. © 2016, Sang Ah Kim.

[1] *See* Victor Laurie, *Computer Protocols- TCP/IP, POP, SMTP, HTTP, FTP and More,* INTERNET TIPS AND TRICKS (Oct. 29, 2016), http://vlaurie.com/computers2/Articles/protocol.htm

[2] *See generally* Celine Otter, *World Wide Web: HTTP Request-Response Cycle,* CELINE OTTER (May 10, 2015), http://celineotter.azurewebsites.net/world-wide-web-http-request-response-cycle/.

[3] *See generally* Filip Jelic, *Man in the Middle Attacks,* DEEP.DOT.WEB (Oct. 10, 2016), https://www.deepdotweb.com/2016/10/10/man-in-the-middle-attacks/.

people freely engage in trading bulk personal data for money.[4] Personal information in bulk is a valuable commodity for consumer marketing companies who use bulk information to provide targeted advertising for businesses. This trade of "data mining" is not something new but has recently grown into a multibillion-dollar industry – often unbeknownst to the sources of the personal information.[5] Besides what a user directly types onto the screen and sends over as a request, information sent using HTTP includes browser information, website content, and other user-submitted information.[6] Browser information showing when the last update to the browser occurred can also be sensitive information, as third parties can take advantage of an outdated browser's security holes to display pop-up advertising; install spyware; and collect personal information for identity theft, among other uses.[7] By using HTTP, a user gambles with the risk of uninvited perusal and exploitation of his/her personal information by third parties.

Communicating via HTTPS, as opposed to using HTTP, makes a critical difference considering such risks. To understand the true significance of HTTPS, we must first learn about the internet's inherent vulnerability. The internet, at its inception, was designed with security assumed because it was designed for use in research, not for commercial means.[8] As third parties invented ways to exploit the information being sent and received online, people needed a way to secure the information against eavesdroppers. A process called "encryption," which hides the content of the information by scrambling the text and rendering it unreadable,[9] arose to compensate for the lack of security.

While its current uses may be cutting-edge, various forms of encryption have been in use for thousands of years.[10] One well-known use of encryption

[4] *Computer Hacking and Identity Theft*, PRIVACY MATTERS (2012), http://www.privacymatters.com/identity-theft-information/identity-theft-computer-hacking.aspx.

[5] Steve Kroft, *The Data Brokers: Selling Your Personal Information*, CBS NEWS (Aug. 24, 2014), http://www.cbsnews.com/news/data-brokers-selling-personal-information-60-minutes/.

[6] OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, MEMO. M-15-13, POLICY TO REQUIRE SECURE CONNECTIONS ACROSS FEDERAL WEBSITES AND WEB SERVICES (2015), https://www.whitehouse.gov/sites/default/files/omb/memoranda/2015/m-15-13.pdf.

[7] *See generally* Paul Cucu, *The Ultimate Guide to Secure Your Online Browsing Today [Updated]*, HEIMDAL SECURITY (Oct. 26, 2016), https://heimdalsecurity.com/blog/ultimate-guide-secure-online-browsing/.

[8] *See* Craig Timberg, *A Flaw in the Design*, WASH. POST (May. 30, 2015), http://www.washingtonpost.com/sf/business/2015/05/30/net-of-insecurity-part-1/.

[9] Eric Kangas, *The Case for Email Security*, LUXSCI FYI BLOG (Mar. 31 2015), https://luxsci.com/blog/the-case-for-email-security.html.

[10] *A Brief History of Cryptography*, RED HAT SEC. BLOG (Mar. 31, 2016), https://access.redhat.com/blogs/766093/posts/1976023.

was the Enigma cypher, used by Nazi Germany to protect the content of their military communication during World War II.[11] Encryption, by making the information unreadable, prevents an adversary employing a MITM attack from accessing the encrypted message; if MITM cannot make sense of the information, MITM cannot exploit the information.

HTTPS is HTTP fortified with encryption and more to account for these vulnerabilities. The "S" in HTTPS stands for a security protocol called "Secure Sockets Layer," which was later renamed to "Transport Layer Security" (hereinafter SSL/TLS).[12] Both names allude to a security "layer." As a security protocol, SSL/TLS dictates the structure of how the information will be secured. Among many requirements, SSL/TLS requires that the communication is encrypted.[13] Having an encrypted communication is like chatting on the phone in pig Latin to prevent an adversary wiretapping the line from understanding what is going on.

However, users quickly realized that encryption alone cannot stop a MITM attack or other forms of electronic eavesdropping. Remember that encryption prevents an adversary from making sense of the intercepted information by making the information unreadable. Meanwhile, the intended recipients – the user and the website – can put the encrypted information back into readable form by using a "key" that, in theory, only the two are supposed to have.[14] It is difficult, however, to securely share this key without creating a potential vulnerability to a MITM attack or other form of hacking.

Realizing that securely sharing the key is a problem, people avoided the question altogether by coming up with an encryption method that does not involve sharing the key. Named "public key exchange," this method uses a pair of keys instead of a single key – a public key and a private key.[15] The public key of the website is free for anyone in the public to get, and the private key of the website remains in private possession of the website. For example, the user encrypts the information with the website's public key, which anyone – including MITM – can freely obtain. The catch is that information encrypted with a public key can only be put back into readable form by using the private

[11] *See The Enigma Machine*, LEARN CRYPTOGRAPHY (Nov. 1, 2016), https://learncryptography.com/history/the-enigma-machine.

[12] *See What Is SSL (Secure Sockets Layer) and What Are SSL Certificates?*, DIGICERT (Nov. 4, 2016), https://www.digicert.com/ssl.htm.

[13] *See Verify TLS is Required*, CHECKTLS.COM (Nov. 20, 2016), http://www.checktls.com/assuretls.html.

[14] *See generally Description of Symmetric and Asymmetric Encryption*, MICROSOFT (Nov. 20, 2016), https://support.microsoft.com/en-us/kb/246071.

[15] Kangas, *supra* note 9.

key, the latter of which never leaves the website's sole possession to begin with. The public key exchange is like using a padlock. Ally buys a padlock and sends her padlock to Billy. Billy writes a note, places it in a box, and locks it with the padlock. Billy cannot unlock the padlock because only Ally has the key. When Billy sends the box with the padlock to Ally, anyone can see that something was sent using a box but cannot unlock the padlock to see its content.[16] Public key encryption appears to be an adequately secure encryption method to protect the content of communication from MITM.

Yet a more fundamental question remains to be answered: how do we know if a website is really what it says it is? Is there an outside source who can verify the website's authenticity? HTTPS makes another critical difference from HTTP in this regard. Remember that HTTPS uses SSL/TLS, the security protocol which requires encryption. In addition to encryption, SSL/TLS also requires the website to prove its identity before encryption even begins – a process called "authentication."[17]

Websites must register with a trusted organization, such as Verisign, which will investigate and vouch for the website's authenticity. Such an organization, called a Certificate Authority, makes the website prove its identity through some paperwork and a fee.[18] The organization sends the website a certificate saying that the presented public key for the website is actually the public key for the website, that this has been verified by the organization, and that the verification is valid for a certain period of time.[19] The website must present the certificate to the user, who will validate the authenticity of the website by using the certificate and asking a Certificate Authority if the certificate is valid.[20]

Certificates, like insurance, come with varying degrees of benefits that match the price tag.[21] There are numerous Certificate Authorities, some of which may not maintain high levels of data checking and reduce their costs at the expense of miss-issuing certificates or at the expense of securing phishing

[16] *See generally* Brian Proffitt, *Understanding Encryption: Here's The Key,* READWRITE (Sept. 19, 2013), http://readwrite.com/2013/09/19/keys-understanding-encryption/.

[17] *See* Eoin Keary, *Authentication Cheat Sheet*, OWASP (Mar. 2, 2016), https://www.owasp.org/index.php/Authentication_Cheat_Sheet.

[18] *See* Eric Kangas, *How Does Secure Socket Layer (SSL or TLS) Work?*, LUXSCI FYI BLOG (July 22, 2013), https://luxsci.com/blog/how-does-secure-socket-layer-ssl-or-tls-work.html.

[19] *Id.*

[20] *See* Keary, *supra* note 17.

[21] *See* Doug Beattie, *How Much Does an SSL Certificate Really Cost?*, GLOBALSIGN (Sep. 23, 2016), https://www.globalsign.com/en/blog/how-much-does-an-ssl-certificate-cost/.

or malware sites.[22] A certificate "ideal for ecommerce sites" through a more well-known Certificate Authority was priced starting at $599/year.[23]

Despite the security benefits of using HTTPS over HTTP, many websites still use HTTP, risking third-party exploitation of users' personal information.[24] Historically, HTTPS was primarily used for sensitive transactions on the internet, such as online payments and corporate information transactions.[25] Over time, however, websites such as Google and Facebook began adopting HTTPS. As of March 2016, Google stated that 77% of its online traffic is encrypted.[26] Obstacles to adoption for private actors include having to change a website's underlying code and needing time to test the access from various regions across the globe using a "diversity of devices."[27]

In a 2015 memorandum articulating recommendations for cybersecurity standards for federal websites**,** the Office of Management and Budget argued that "tangible benefits to the American public [in deploying HTTPS across federal websites] outweigh the cost to the taxpayer" in that a few malicious impersonation of official federal websites or eavesdropping on the said websites may create substantial risks to members of the public.[28] As of October 28, 2016, 58% of federal websites use HTTPS.[29] Still, various administrative and financial burdens, such as development time and the burden of maintenance, affect the rate of adoption for federal websites.[30]

A significant part of daily life on the internet will remain vulnerable to third parties if websites continue to use HTTP. HTTPS generally does not affect whether a website is vulnerable to hacking, due to the internet's inherently vulnerable design. HTTPS, however, protects the communication from impersonation by third parties; keeps potentially sensitive information secure;

[22] *Id.*

[23] *See ExtendedSSL*, GLOBALSIGN (Nov. 5, 2016), https://www.globalsign.com/en/ssl/ev-ssl/.

[24] *See* Brian Barrett, *Most Top Websites Still Don't Use a Basic Security Feature,* WIRED (Mar. 17, 2016), https://www.wired.com/2016/03/https-adoption-google-report/.

[25] Scott Gilbertson, *HTTPS is More Secure, So Why Isn't the Internet Using It?,* ARS TECHNICA (Mar. 20, 2011), http://arstechnica.com/business/2011/03/https-is-more-secure-so-why-isnt-the-web-using-it/.

[26] Michael Liedtke, *Google Reveals 77 Percent of its Online Traffic is Encrypted*, PHYS.ORG (Mar. 15, 2016), http://phys.org/news/2016-03-google-reveals-percent-online-traffic.html.

[27] *See* Owen Williams, *Wikipedia Now Uses HTTPS to Stop People Snooping on your Binge Learning*, THE NEXT WEB (Jun 12, 2015), http://thenextweb.com/insider/2015/06/12/wikipedia-now-uses-https-to-stop-people-snooping-on-your-binge-learning/#gref.

[28] OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, *supra* note 6.

[29] *Secure HTTP (HTTPS),* PULSE (Oct. 28, 2016), https://pulse.cio.gov/https/domains/.

[30] OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, *supra* note 6.

and prevents the information from being tampered with or modified.[31] It is true that not all information exchanged across the internet necessarily requires impenetrable privacy; but with the increasing commodification of user activity on the internet and government surveillance, the threshold of what information we would allow to pass unencrypted might change quicker than expected.

[31] *See Introduction to HTTPS*, THE HTTPS-ONLY STANDARD (Nov. 5, 2016), https://https.cio.gov/faq/#what-does-https-do%3f.