Patrick Hill

CSU 697: Research 1

Dr Lin

**An Investigation of Cyber Security and HTTPS Adoption on Government Websites**

INTRODUCTION:

Local government websites are a significant part of critical infrastructure. In times of emergency, citizens will look to their local governments' websites for information, guidance, and advisement. Local government websites are a crucial tool that is used to help citizens feel that their local Government is stable and in control. Many hacker groups, such as the Russian Internet Alliance, will seek to disrupt citizen trust and confidence in their governments by posting malicious or unsettling images on their websites. Recently, cyber autonomy has gained increased attention by the U.S. Government due to the impact of state-sponsored hacktivism, hardware\software shipped with spyware, compelled certificates issued for Internet interception, and exploited zero-day exploits. An investigation of public-facing, local government websites to determine HTTPS, HSTS, and a digital certificate analysis was performed. Since public-facing local government websites are the front-line for delivering official information and accessing local government services, attacks targeting these websites can cause widespread panic and citizen instability and be used to steal citizen credentials and sensitive citizen data. My investigation titled “**An Investigation of Cyber Security and HTTPS Adoption on Government Websites”** will examine the local government counties' websites for critical elements of cyber security:

* HTTPS Adoption
* HSTS Adoption
* Certificate Authorities (C.A.) usage
* C.A. Country origin
* domain name registration
* insecure site calls
* susceptibility to significant exploits

Public-facing website defacements typically involve a cyber threat actor compromising the website or its associated CMS, allowing the actor to upload images to the site's landing page. In situations where such public-facing websites relate to elections (e.g., the website of a county board of elections), defacements could cast doubt on the websites' information security and legitimacy. If cyber actors could successfully change an election-related website, the underlying data and internal systems would remain uncompromised. This study examines the public's external dependency facing local government websites across the local counties in the United States of America to probe into the concept of local government cyber autonomy and security.

For example, the Town of Hilton Head experienced a breach of their webserver on 10/4/2020, resulting in the below defacement.



In the paper "An Investigation of Cyber Autonomy on Government Websites", the authors identify four significant forces that can influence the degree of a local government's autonomy, including HTTPS adoption, website development, outsourcing, and citizens' fear of large-scale surveillance, and user confusion [2]. My study attempts to expand this research and to provide an analysis of local governments in an effort to bring awareness to local government website insecurities.

My research will measure the current security standing of local government sites and sub-sites. My study expects to raise awareness of local county government website's security status and to reduce the risk of blindly trusting external entities when using critical government services.

BACKGROUND

Many local governments have multiple local Government branded websites that are not associated with a valid .gov domain and therefore have not gone through the vetting process associated with obtaining a .gov domain. Such few County's use a .gov domain for their website that an attacker can easily spoof a local government site and convey false information to citizens. Minnesota and Texas have the largest number of county sites that do not use the .gov domain. Poorly secured County websites with a low level of autonomy give attackers a much more realistic opportunity to influence and disrupt citizen activities.

The article "County Election Websites Can Be Easily Spoofed to Spread Misinformation" brings attention to how county government websites in 20 key swing states do not use a .gov domain nor enforce the use of SSL. At the time of the article, Minnesota and Texas have the largest percentage of non-.gov county government sites, with 95% of their county sites using HTTP [1]. West Virginia, Texas, and Montana have the most significant number of county governments not using SSL, which would allow attackers to redirect website visitors to alternate, malicious sites. The article states that lack of consistency in website naming and the use of SSL certificates pose a much more realistic threat to the integrity of the election process than a physical attack on voting machines. Often, County election sites are the first-place voters go to for eligibility requirements, voting locations, registration deadlines, and hours. It is feared that simple misinformation campaigns focused on vulnerable gaps at the local level could negatively impact voting results. Poorly secured county websites give attackers a much more realistic opportunity to influence the outcome of elections. Since not all counties use a .gov domain, voters would have difficulty identifying spoofed sites from real ones [1].

With inconsistency in website naming and the lack of SSL certificate use, make county websites a high-profile target for malicious actors.

The .gov top-level domain (TLD) facilitates collaboration among government-to-government, government-to-business, and government-to-citizen entities. The TLD authorizes domain names for bona fide US-based government organizations at the federal, state, and local levels, including federally recognized Indian tribes and Alaskan Native groups, known as native sovereign nations (NSNs). .gov makes government services easy to identify on the internet.

The hacker group APT28 actively interfered with the 2016 presidential election, with many sites being created to spoof local government sites. The article "Microsoft-says-it-has-found-a-Russian-operation-targeting-us-political-institutions" states that during the 2016 election, a group affiliated with the Russian Government created fake versions of six websites with the goal of hacking people that visited these fake websites; some of which were related to public policy and the US Senate [3]. U.S. officials repeatedly warned that the November elections are a major focus of malicious interference efforts. APT28, which is sometimes called Strontium or Fancy Bear us a unit under the Russian Military intelligence agency GRU, which specializes in Misinformation. APT refers to an advanced persistent threat. Hackers will often send out fake emails, directing people to visit the sites that appear to be legitimate [3].

Every .gov domain name application is carefully examined to ensure domain names requested will not create misunderstandings about the purpose of domains and their content. The .gov vetting process's overall goal is tomaintain domain name integrity, eligibility is limited to qualified government organizations, and programs for having a managed domain name such as .gov assures citizens that they are accessing an official U.S. government site. General Services Administration (GSA) arbitrates domain name issues and reserves the right to deny domain name requests that do not adequately meet requirements. The Title 41 Public Contracts and Property Management in the [Code of Federal Regulations Chapter 102, sub-chapter 173](https://ecfr.federalregister.gov/current/title-41/subtitle-C/chapter-102) outlines Government requirements for a .gov domain. Domain names must be authorized by the Chief Information Officer (CIO) of the requesting or sponsoring governmental organization. For Federal departments and agencies, the General Services Administration (GSA) will accept authorization from the department or agency's CIO. For independent Federal government agencies, boards, and commissions, GSA will accept authorization from the highest-ranking Information Technology Official. For State and local governments, GSA will accept authorization from appropriate State or local officials [11].

Not only are local governments using non .gov domains, but they are also slow to adopt HTTPS. HTTPS [3] runs HTTP over Transport Layer Security (TLS), a fundamental security protocol that enables end-to-end encryption and authentication for HTTP connections. Research into a novel attack on TLS titled "A cross-protocol attack on the TLS protocol", authors Nikos Mavrogiannopoulos, Frederik Vercauteren, Vesselin Velichkov, and Bart Preneel show that TLS alone is not infallible [4] . The authors present a cross-platform exploit that shows an attacker can interpret signed explicit elliptic curve Diffie-Hellman (DH) key exchange parameters as valid plain parameters that enable the impersonation of a trusted server. The server has to support the explicit curve options for the attack to be successful. Proper configuration of HTTPS sites is essential to improving their cyber-security as HTTPS adoption improves. The paper describes the TLS protocol as an agile protocol that allows peers to negotiate their highest supported protocol version and use a combination of ciphers during each session [4]. The TLS cipher suite determines the symmetric encryption cipher with its operational mode, the key exchange method, and the message authentication algorithm. The TLS downgrade dance is for backward compatibility with legacy servers. During handshake negotiation, a server will attempt to use the highest level of TLS that the client supports. If the handshake fails, a retry will be initiated using the next lowest TLS version. The POODLE attack will attempt to intercept the TLS handshake negotiation to force a weak SSL 3.0 connection. In the paper, “POODLE Bites: Exploiting the SSL 3.0 Fallback”, the authors stress the need for proper TLS and webserver security configuration and the disabling of older legacy support protocols such as SSL 3.0 [9].

In the work "SSL and HTTPS: Revisiting Past Challenges and Evaluating Certificate Trust Model Enhancement", authors Clark and Oorschot show many of on-going security flaws with HTTPS such as attacks involving fraudulent certificates, SSL stripping attacks, and the lack of HTTPS support [5].

Nearly all secure web communications take place over HTTPS. HTTPS is based on TLS encrypted transport protocol and a supporting key infrastructure of thousands of certificate authorities (C.A.'s) – entities trusted by users' browsers to vouch for the identity of a webserver. TLS is one of the major secure communication protocols on the internet. It is an agile protocol that allows peers to negotiate their highest supported protocol version, as well as the combination of ciphers used in a session. Not all cipher suites within TLS are strong. The paper "Analysis of the HTTPS Certificate Ecosystem" list data is collected by performing 110 Internet-wide scans over 14 months, and identity vulnerabilities and user-facing errors that negatively impact the overall security of the internet ecosystem [8]. In the paper, the authors investigate the trust relationship between root authorities. The authors analyzed 1832 CA certificates controlled by 683 organizations and found that 80% of the organizations do not have certificates from commercial certificate authorities. The C.A.'s constraints investigated and found that only 7 C.A. certificates use name contrasts and more than 40% of the C.A.'s have no length constraint. The authors identify two sets of mis-issued C.A. certificates. The authors found many problematic security issues within their study, such as a public key compromise that would require 26% of the HTTPS websites to obtain new certificates. They found that half of the trusted leaf certificates contain an inadequately secure 1024 RSA key in their trust chains.

C.P.'s are servers that host web resources for websites. When a website loads resources from untrusted servers, various undesirable consequences can occur, such as the execution of malicious scripts, malicious content, mining cryptocurrencies, or sending bot-net attacks [8].

In the research work “Re-Architecting the Internet” [6] authors state that security in the WWW architecture is based on authenticating the source server and securing the data during transport without considering the content itself. The traditional assumption is that a page is as secure as the server hosting it. However, modern web sites often have a composite structure where different actors author components of the web page and one logical page contains components collected from disparate servers. Applying a single security policy to a whole page is inadequate.

They introduce a novel way of protecting users from web-based malware, which is a new model that uses opportunistic personas to better secure web content by adding integrity and accountability to individual elements. In this paper, the authors present the overall design of the mechanism and details derived from a prototype of the system [6].

In the study, “The Matter of Heartbleed” [10], the authors perform a comprehensive, measurement-based analysis on the vulnerability impact, including tracking the population, monitoring the patching behavior over time, and assessing the HTTPS ecosystem's impact and exposing real attacks. The authors found that 44 of the top 100 Alexa websites remain vulnerable two months after the patch was released. In addition, only 10% of analyzed websites replaced their certificates compared to 73% that patched their site, and 14% of those using the same private key [10]. HTTPS is the secure variant of the HTTP protocol on which the Web is based. HTTPS provides cryptographic security protections by carrying HTTP messages over the TLS protocol instead of directly over TCP. HTTPS websites authenticate using digital certificates as part of the TLS handshake. Web users are shown an invalid certificate warning when their browser cannot validate the identity of the websites they are visiting. While these warnings often appear in benign situations, they can also signal a man-in-the-middle attack. However, many more frequent users are connecting to a legitimate website with erroneous or self-signed certificates. The research performed by Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor in "Crying Wolf: An Empirical Study of SSL Warning Effectiveness" [7] shows that invalid certificate warnings can signal a man-in-the-middle attack or a DNS spoofing attack. The authors surveyed 400 Internet users to examine their reactions to understanding website certificate warnings and their effectiveness. Their research showed that the warnings are often by-passed and that preventing users from making connections to unsafe websites is the safer approach [7].

Browsers also enforce additional policies for HTTPS pages, for example, ensuring that HTTPS pages cannot load scripts from non-secure sources. Authors Felt, Barnes, et al. of the paper "Measuring HTTPS Adoption on the Web" attempt to measure the HTTPS adoption rate [12]. The authors state the tremendous growth in HTTPS adoption has been positively trending since 2016 with nearly half of the top 100 websites supporting https. HTTPS provides cryptographic security protections by carrying HTTP messages over TLS instead of directly over TCP. HTTPS works by utilizing public-key cryptography and third-party digital signatures to encrypt and confirm data communications between clients and servers. Communications between the browser and the webserver are not accessible in plaintext to intermediate entities. Intermediate entities cannot make modifications to content sent between the browser and the webserver. The client is assured that the other end of the channel is the one that it intends to communicate with [12]. Projects like the "Let's Encrypt", HTTPS only standard and search ranking changes to promote HTTPS are credited with pushing HTTPS adoption rates. Browsers now require HTTPS to unlock and enable certain features. HTTPS is focused on protection against network attackers but does not protect against other types of attacks.

However, without HTTPS, the job of attackers become much more accessible. Without HTTPS, traffic can be intercepted by any en-route adversary, as well as off-path adversaries who are capable of hijacking routes. This inflight manipulation of web page content is one of the significant reasons to abandon HTTP websites.

METHODOLOGY

There is no single source that maintains the list of local government websites; therefore, my study will manually create a dataset on amazon using a web-accessible Cassandra database. County websites that utilize HTTP will be contacted to determine if they develop 'in-house', outsource their website development to a 3rd party, or if a non-IT individual develops and maintains the website. I will examine and log every U.S. county website for HTTPS adoption, HSTS adoption, valid certificates, originating certificate county, and .gov domain implementation and apply a score based to the site.

HYPOTHESIS

I believe my research will show that:

1) HTTPS adoption rates are poor withing local government. My research will show that 25% of all local government websites are still not HTTPS. HTTPS adoption metrics allow us to see if local government HTTPS adoption mirrors the current state of HTTPS adoption in the shifting Web at large.

2) My research will show that non-HTTPS County websites have under skilled or third party web development and that lack of government web security mandate leads to compromise and undermines citizen confidence and security.

3) My research will show that digital certificates used will be predominately from US-based certificate authorities (C.A.). Partial inclusion of the government root C.A. in the browsers trust stores results in certificate warnings, confusing the users about whether a site is legitimate. My research will confirm that 90% of all Root CA's will be US-based organizations.

4) Research will show that HSTS adoption rate is poor, with only a 10% adoption. HTTP strict transport security is a web security policy mechanism that allows a server to declare that it accepts only HTTPS connections from browsers. HSTS is designed to prevent man in the middle attacks.

5) My research will also show that .gov domain adoption rates will be used by 75% of local government sites. In addition, I will also track the resilience to TLS attacks such as Heartbleed, Freak attack, Logjam attack, Renegotiation attacks, POODLE and Sweet32 attacks. Heartbleed is an incredibly widespread exploit, allows a remote attacker to read protected memory from vulnerable servers, potentially including cryptographic keys, login credentials, and private data from an estimated 24-55% of popular HTTPS sites. My research will attempt to rate the website based on the number of attacks it is susceptible to utilizing Metasploit. My research will also examine the highest level of TLS a site supports and if TLS will downgrade to a weaker version if requested by the client.

I will use the criteria above to provide a security confidence score to each County website.

URL collection: There is no single source of up-to-date listings of local government websites. Therefore, I will build my datasets and ratings with site scraping using Maltego, Chromium version 86.0.4240.111, Metasploit and Nmap.

REFERENCES

1. Vijayan, Jai (2018, March 23) *County Election Websites Can Be Easily Spoofed to Spread Misinformation.* DarkReading.

<https://www.darkreading.com/vulnerabilities-threats/county-election-websites-can-be-easily-spoofed-to-spread-misinformation/d/d-id/1333132>

1. Hsu-Chun Hsiao, Tiffany Hyun-Jin Kim, Yu-Ming Ku, Chun-Ming Chang, Hung-Fang Chen, Yu-Jen Chen, Chun-Wen Wang, and Wei Jeng. 2019. *An Investigation of Cyber Autonomy on Government Websites*. In The World Wide Web Conference (WWW '19). Association for Computing Machinery, New York, NY, USA, 2814–2821. DOI:https://doi.org/10.1145/3308558.3313645
2. Dwoskin, E. & Timberg, C (2018, August) *Microsoft says it has found a Russian operation targeting U.S. political institutions.* Washington Post.

<https://www.washingtonpost.com/business/economy/microsoft-says-it-has-found-a-russian-operation-targeting-us-political-institutions/2018/08/20/52273e14-a4d2-11e8-97ce-cc9042272f07_story.html?noredirect=on&utm_term=.db741731087b>

1. Nikos Mavrogiannopoulos, Frederik Vercauteren, Vesselin Velichkov, and Bart Preneel. 2012. *A cross-protocol attack on the TLS protocol.* In Proceedings of the 2012 ACM conference on Computer and communications security (CCS '12). Association for Computing Machinery, New York, NY, USA, 62–72. DOI:https://doi.org/10.1145/2382196.2382206
2. J. Clark and P. C. van Oorschot, "SoK: SSL and *HTTPS: Revisiting Past Challenges and Evaluating Certificate Trust Model Enhancements,*" 2013 IEEE Symposium on Security and Privacy, Berkeley, CA, 2013, pp. 511-525, doi: 10.1109/SP.2013.41.
3. Anja Feldmann, Luca Cittadini, Wolfgang Mühlbauer, Randy Bush, and Olaf Maennel. 2009. *HAIR: hierarchical architecture for internet routing. In Proceedings of the 2009 workshop on Re-architecting the internet* (ReArch '09). Association for Computing Machinery, New York, NY, USA, 43–48. DOI:https://doi.org/10.1145/1658978.1658990
4. Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. 2009. *Crying wolf: an empirical study of SSL warning effectiveness.* In Proceedings of the 18th conference on USENIX security symposium (SSYM'09). USENIX Association, USA, 399–416.
5. *Z*akir Durumeric, James Kasten, Michael Bailey, and J. Alex Halderman. 2013. *Analysis of the HTTPS certificate ecosystem*. In Proceedings of the 2013 conference on Internet measurement conference (IMC '13). Association for Computing Machinery, New York, NY, USA, 291–304. DOI:https://doi.org/10.1145/2504730.2504755
6. Möller, Bodo; Duong, Thai & Kotowicz, Krzysztof. *This POODLE Bites: Exploiting the SSL 3.0 Fallback*, report, September 2014; (<https://digital.library.unt.edu/ark:/67531/metadc949502/>: accessed November 12, 2020), University of North Texas Libraries, UNT Digital Library, [https://digital.library.unt.edu](https://digital.library.unt.edu/); crediting UNT Libraries Government Documents Department.
7. Zakir Durumeric, Frank Li, James Kasten, Johanna Amann, Jethro Beekman, Mathias Payer, Nicolas Weaver, David Adrian, Vern Paxson, Michael Bailey, and J. Alex Halderman. 2014. *The Matter of Heartbleed.* In Proceedings of the 2014 Conference on Internet Measurement Conference (IMC '14). Association for Computing Machinery, New York, NY, USA, 475–488. DOI:https://doi.org/10.1145/2663716.2663755
8. [Code of Federal Regulations Chapter 102, sub-chapter 173](https://ecfr.federalregister.gov/current/title-41/subtitle-C/chapter-102)
9. Adrienne Porter Felt, Richard Barnes, April King, Chris Palmer, Chris Bentzel, and Parisa Tabriz. 2017. *Measuring HTTPS adoption on the web*. In Proceedings of the 26th USENIX Conference on Security Symposium (SEC'17). USENIX Association, USA, 1323–1338.