

# Poster: On the (In)Security of Government Web and Mail Infrastructure

Evan Lam\*, Richard Anderson\*, Kurtis Heimerl\*, Yurie Ito†, Jonathan Joseph de Koning†, Adam M. Lange†, Jarrod O’Malley†, Adam Shostack\*†, Arastoo Taslim†, Sudheesh Singanamalla\*†

\* University of Washington, † CyberGreen Institute

**Abstract**—Government web infrastructure is a critical part of today’s Internet and the functioning of society. Citizens’ interactions with digital government infrastructure needs to be secure since they might contain important and sensitive information. These interactions can be through various web applications providing digital public services, or through communication mechanisms such as email. Government websites and mail servers typically form the long tail of today’s Internet and do not appear on large top million Internet datasets making them very understudied. DNS infrastructure forms the center piece for citizens to interact with government services allowing resolution of IP addresses, and enabling email communication and sender policy enforcement between mail service providers. In this poster, due to their inter-dependent nature, we present a comprehensive security evaluation of government web infrastructure covering both web and mail services in addition to understanding the security of the DNS services they rely on. We open source our implementation of the security scanner to the community, invite collaborators to engage with the data periodically scanned, and release the largest public dataset of government hostnames.

## I. INTRODUCTION

The usage of Transport Layer Security (TLS) enables secure Internet interactions. Due to its application-neutral nature, it has been adopted to not just secure communications to websites but also to secure communications between mail servers on the Internet through the SMTP STARTTLS extension. Authoritative DNS name servers responsible for the government domains provide the necessary information to client requests. Plain text DNS responses sent from these authoritative nameservers are vulnerable to tampering by network adversaries. However, such attacks on the integrity and authenticity of the responses can be addressed through the usage of DNS security extensions (DNSSEC). In this work, we curate multiple large datasets, filter for government hostnames and release the largest known dataset focused on governments. We build high performance network scanners to measure the adoption and usage of DNSSEC by various domains, in addition to understanding their usage of TLS on web and mail servers and draw insights about their similarity.

## II. DATASET & SCANNER TOOLS

We begin our efforts by creating a large dataset of government hostnames. We first synthesize and merge 17 publicly available top million datasets and databases published in previous research efforts [2], [1], and perform a one time extension of the dataset by identifying Subject Alternative Names (SANs) from their TLS certificates. To ensure we only include government hostnames, we crosscheck the eTLD of each hostname against a list of pre-generated and well known government eTLDs. This list was compiled by combining

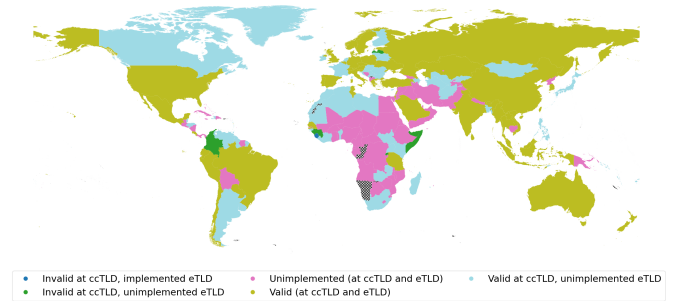


Fig. 1. The categorization of DNSSEC implementation at ccTLD (e.g. au. and .br) and eTLD (e.g. gov.au. and gov.br) levels per country. For countries with multiple government eTLDs (eg. .gov.us and .gov) most popular is chosen.

country domains (such as .ly, .us, .au) with common government extensions (such as .gov, .gob, and .go). Since a few governments such as the Canadian provincial governments do not follow this exact format, we manually add them along with special federal (.fed), and military (.mil) ccTLDs – resulting in the final dataset containing 401216 hostnames.

In collaboration with a non-profit partner CyberGreen, we built a custom high performance Internet security scanner that identifies the usage and validity of DNSSEC for a given hostname, performs TLS scans on websites and mail servers returned from the authoritative DNS nameservers. These scans are performed periodically from AWS cloud instances in Singapore with a reserved static IP address and associated reverse DNS PTR records. Due to the sensitive and probing nature of these scans, we duly notify AWS trust and safety teams about the intent to run the scans, provide appropriate opt-out and informational pages about the scanner. As a part of this poster, we open source our scanner implementation written in Golang, the extended dataset, and invite future collaborations from Internet measurement and security communities to leverage the results for longitudinal studies.

## III. RESULTS

Of the 401216 hostnames in our final dataset, we successfully resolve the A (IPv4) records for 304774 (75.96%) hostnames in our DNS scans, and are able to resolve 308241 hostnames for either IPv4 and IPv6 records, and obtain 59533 unique mail exchange MX records before attempting to make TCP and TLS handshakes. 251819 (75.91%) of these hostnames are served through HTTPS enabled webpages using port 443. Our results indicate that approximately 25% of the government websites across the world still do not use TLS.

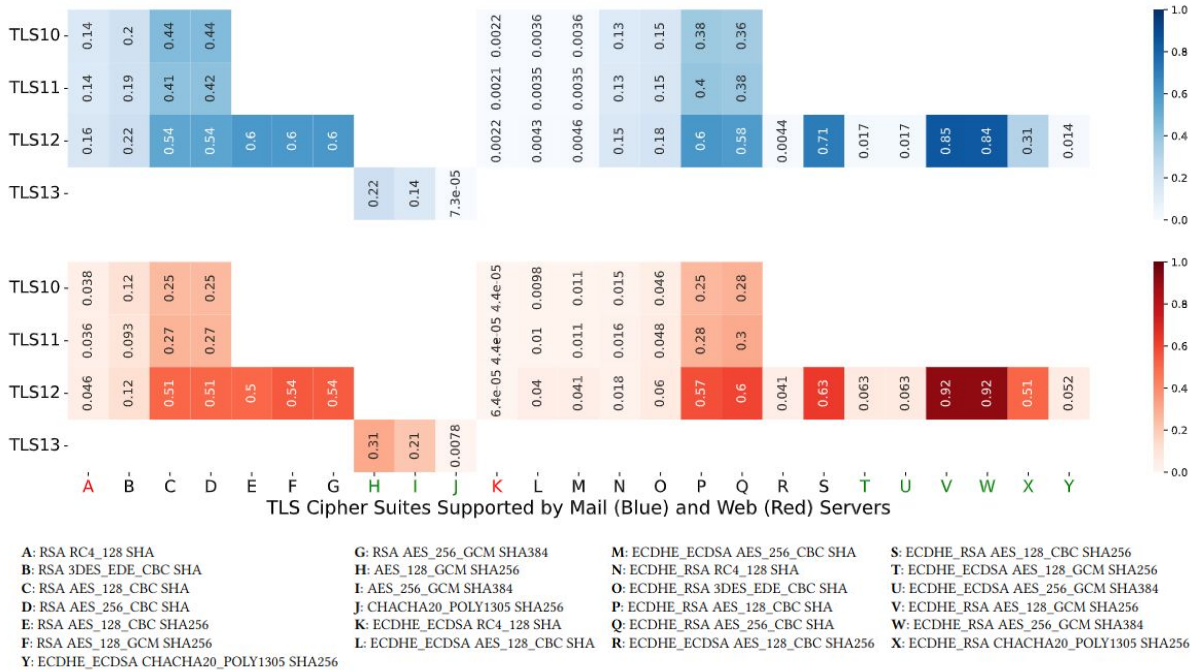


Fig. 2. Heatmaps indicating the percentage of government hostnames supporting various TLS cipher suites grouped by TLS protocol version. The blue heat map on top indicates the protocol usage for 16830 mail servers, and red heat maps indicate the protocol and cipher suite support for 251819 web servers. The cipher suites with red labels (A, K) are classified as insecure, black labels are weak, and the green labels as the secure cipher suites. The key is provided below and includes the key exchange, encryption, authentication, and MAC algorithms used in the cipher suites.

### A. DNSSEC Adoption and Usage Needs Improvement

We compare the adoption of DNSSEC at the ccTLD level (eg. *.uk*), with that of the respective country government eg. *.gov.uk* which we term as the *eTLD* for this section and present a map as shown in Figure 1 to understand the adoption of DNSSEC at each country and their respective governments’ hostnames. We observe 57 countries (light green) have valid DNSSEC enrollments, 68 (light blue) have DNSSEC enabled at the TLD but not at the government eTLD, 68 others (magenta) with the majority in Africa and the Arabian region have not enabled DNSSEC at either level, and 5 countries (dark green) have invalid DNSSEC entries at the ccTLD. Sierra Leone is the only country which has invalid DNSSEC ccTLD entries but attempts to use DNSSEC for their eTLDs. Of the 304774 hostnames, 269171 (67.09%) hostnames do not have DNSSEC enabled. Of the remaining 35603 hostnames that have DNSSEC enabled, we identify that 23676 hostnames are invalid, reducing the number of hostnames with valid DNSSEC records to 11927, a mere 2.97% of the dataset and much below the current DNSSEC adoption rate of  $\leq 7\%$  for *.com*.

### B. TLS Protocol Usage Differs Between Mail and Web Servers

Of the HTTPS supporting hostnames, 32.67% of hostnames support TLS 1.0, 35.08% support TLS 1.1, 99.62% support TLS 1.2, and 59.76% support TLS 1.3. 500 hostnames solely support TLS 1.0, and 23 support only TLS 1.1 – both of which have been deprecated. TLS 1.2 is supported by 89.09% of the mail hosts, followed by 39.91% and 37.3% of the mail hosts supporting TLS 1.0 and TLS 1.1 respectively. Only 25.8% of the mail servers support TLS 1.3. We observe that despite being a smaller set of hosts compared to the web, 924 mail

hosts only support TLS 1.0 accounting for 5.49% of the mail hosts indicating a stark contrast with web server infrastructure where only 0.19% of the hosts support TLS 1.0. A larger proportion of mail servers configured by governments tend to use insecure and deprecated TLS protocols compared to their web counterparts and are also shown in Figure 2. A higher percentage of mail hosts compared to web servers, use weak TLS cipher suites based on RSA and cipher block chaining (C, D) possibly because they are supported by TLS 1.0, TLS 1.1, and TLS 1.2. The continuing support for TLS 1.0 and TLS 1.1 puts these web servers at risk of various downgrade attacks such as BEAST, CRIME, and SLOTH.

## IV. CONCLUSION

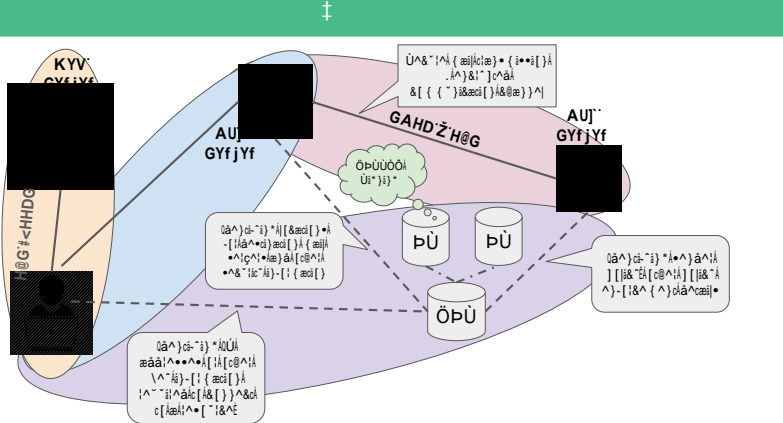
When understanding the security of web infrastructure, we posit that it is valuable to consider the entire ecosystem – especially for the long tail of the Internet. In this poster, we present a preliminary attempt to reason about and understand security by considering website, mail, and DNS infrastructure simultaneously. Our measurements reveal insights about the global extent of insecurity and inconsistency in the use of secure protocols dependent on the application. We contribute an open-source scanner and the largest curated dataset focused on government hostnames.

## REFERENCES

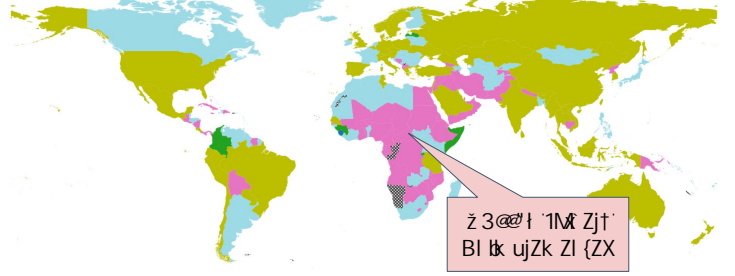
- [1] N. Samarasinghe, A. Adhikari, M. Mannan, and A. Youssef, “Et tu, brute? privacy analysis of government websites and mobile apps,” in *Proceedings of the ACM Web Conference 2022*, 2022, pp. 564–575.
- [2] S. Singanamalla, E. H. B. Jang, R. Anderson, T. Kohno, and K. Heimerl, “Accept the risk and continue: Measuring the long tail of government https adoption,” in *Proceedings of the ACM Internet Measurement Conference*, 2020, pp. 577–597.

# Cb'h\Y'fl=bŁGYW i f]hm'cZ' ; c jYfb a Ybh'KYV'UbX'AUJ''=bZfUghf i Wh i fY

9jub@ua...  
CEææ { Á T É Á Šæ } \* ^ É Á Ræi ; [ á Á U q T æ ] | ^ ^ É Á Cææ { Á Ú @ [ • cææ \ ^ É Á Cææ : æ • c [ [ Á Væ • ] á { Á É Á Ú ~ á @ ^ ^ • @ Á Ú á } \* æ } æ { æ | æ } ^ w \*



É 8%ÉÉA1 [



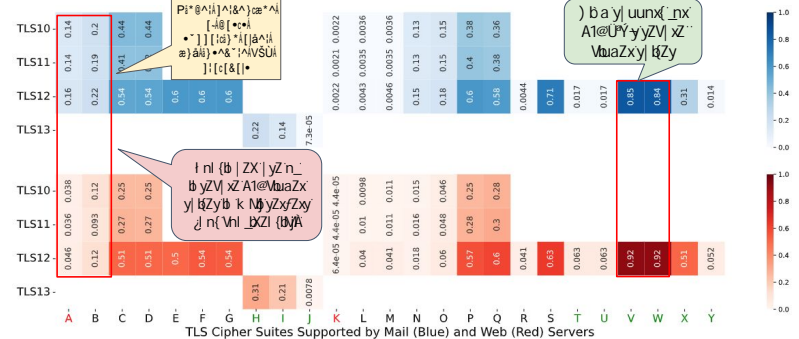
Invalid at ccTLD, implemented eTLD; Invalid at ccTLD, unimplemented eTLD; Valid at ccTLD, implemented eTLD; Valid at ccTLD and eTLD; Unimplemented (at ccTLD and eTLD); Valid at ccTLD, unimplemented eTLD

G97IF9...-BG97IF9

FÉ Úiá [ íá! ^ ^ æi&@A~ [ íc~Á- [ & ~ Á ] Áæá ~ à ^ c h [- Á c @ ^ Á ^ & [ • ^ • c ^ { Á Ç & [ [ í Á áá æ! ^ æ • D É Á C c h á { ] [ í cæ } c h c [ Á & [ ] • á á ! Á c @ ^ Á ^ } c á ! ^ Á ^ & [ • ^ • c ^ { Á , @ á ! ^ Á c ~ á á } \* Á á } - í æ • c í • & c ! ^ Á ^ Á & " í á c ~ É

Y [ í á á { æ } Á [- Á Ö P Ü Ú Ö Á á } [ ^ { ^ } cææ { Á æ c á V S Ö Á Ç ^ É \* É á æ ~ [ É á ] D á æ } á á & V S Ö Á Ç ^ É É á | D É Á V @ ^ Á { [ • c h ] [ ] æ í Á V S Ö Á É á & V S Ö Á & [ { á } æ c á ] • Á - [ í Á ^ æ & @ Á & [ í ! ^ Á ] [ ] á } \* Á & [ ~ ] c í ~ É

1 Ö É ö †

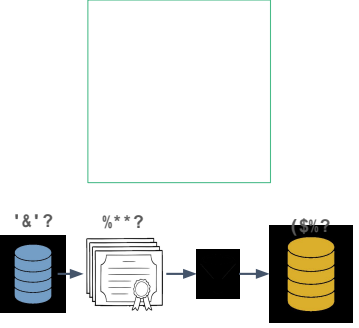


8 É A

Ö ~ Á c [ Á c @ ^ Á ] æ c ! ^ Á [- Á \* [ Ç ^ ! ] { ^ } c h , ^ á • í c Á } æ { á } \* Á • & @ ^ { ^ • É á á ^ } c á ~ á } \* Á \* [ Ç ^ ! ] { ^ } c h @ [ • c ] æ { ^ • Á • Á } [ ] c í á ç æ É Á V @ ~ • Á , ^ Á ^ [ ] [ ~ Á æ Á • c í á } \* ^ } c h - á [ ç í á } \* Á c & @ } á ~ ^ É á Ç í q x 0 e e v n f z í o i q w x 0 í o i w d 0 í o i q 0 í o i q d 0 Á c & É É É D

W : á } \* Á V S Ú Á Ü C E P Á - á ] á Á , ^ Á ^ c } á á c @ ^ Á [ • c ] æ { ^ Á á æ æ • ^ c á ^ Á r \* % ~ \$ \_

É 8 É



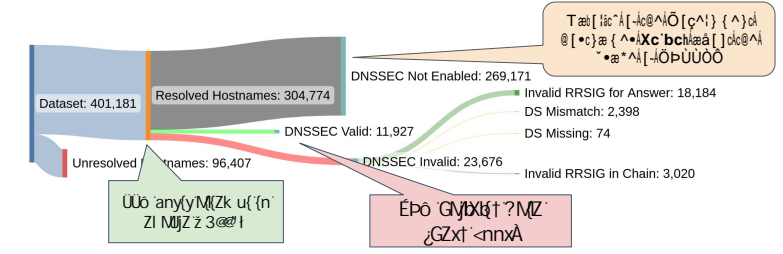
P æ { æ } á } á í & æ c á } \* Á c @ ^ Á ^ } c á ! ^ Á ^ & [ • ^ • c ^ { Á Ç ^ ! ] { ^ } c h @ [ • c ] æ { ^ • Á • ~ } [ í c á } \* Á ç æ í á ~ • Á V S Ú Á á } @ ^ Á • ~ í c ~ ] ^ á á á ^ Á V S Ú Á } [ c [ & [ í Á Ç ^ ! ] • í } É á V @ ^ Á á } ^ Á @ ^ æ c h { æ } Á [ ] Á c [ ] Á á } á á & æ c Á • c @ ^ Á } [ c [ & [ í Á ~ • æ \* ^ Á - [ í Á F í ] H É Á { æ á ] ^ í Ç ^ ! • É á } á á ! á á @ ^ æ c h { æ } • á } á á & æ c Á • c @ ^ Á } [ c [ & [ í Á æ } á á & á } @ ^ Á • ~ í c Á • ~ ] [ í c - [ í Á G í F í F í Á , ^ á • ^ í Ç ^ ! • É á V @ ^ Á á } @ ^ Á • ~ í c Á , í c @ Á f Y X á } æ á • Á C C É Á S D á } æ í Á & [ æ • í - á á æ c Á } • Á & ! ^ É á V U W \_ ] æ á • ^ í c Á , ^ æ É á æ } á á c @ ^ Á [ f Y Y b ] á } æ á • ^ í c @ ^ Á • Á & ! ^ Á & [ ] @ ^ Á • ~ í c Á • É É

Reason for Invalidity	Web TLS (N=40590, 16.11%)	Mail TLS (N=4956, 29.44%)
Hostname Mismatch	19085 (47.01%)	3449 (69.59%)
Self Signed Certificates	10473 (25.80%)	225 (4.53%)
Expired Certificates	8176 (20.14%)	667 (13.45%)
Missing SAN Names	2790 (6.87%)	544 (10.97%)

'yZxn] yXbyUvMjt' b'At@Zx[b]bVz' /y/bktf UZ (. ZZI' H ZUMIX 2 Nj

A

¾ 8%ÉÉA1 Ü [ b [



1

FÉ Y Á [ • í c h c @ æ c í á á • Á ç æ } æ á ! ^ Á ^ & [ • á á ! Á c @ ^ Á ^ } c á ! ^ Á ^ & [ • ^ • c ^ { Á É É Á • } ^ Á & æ | ^ Á - [ í Á c @ ^ Á ] [ ] \* Á ç æ á [ - Á c @ ^ Á ] c ^ ! ^ c É Á

GÉ U ~ í Á { ^ æ • í ^ { ^ } • Á • ^ Ç ^ ! } • í } æ \* c @ ^ æ á [ ~ c á c @ ^ Á ] [ æ á ] ^ c c ^ } c h [- Á á } • Á & ~ í á c ~ Á æ } á á } & [ ] • í • c ^ } & Á á } • Á ^ [- Á ^ • & ! ^ Á ] [ ] c [ & [ í Á á ^ } á ^ } c h [ ] Á c @ ^ Á æ } ] í & æ c á } É Á

U ~ V @ ^ Á á [ ] c á ] Á [- Á Ö P Ü Ú Ö Á á } \* Á Ç ^ ! ] { ^ } c h á [ { æ } • @ [ • c ] æ { ^ • Á • Á ] [ ] É

V ~ Ö [ Ç ^ ! ] { ^ } c h { ç } æ \* ^ á á { æ á ] á á , ^ á á ^ í Ç ^ ! • Á æ á ! Á & [ ] - á ~ ! ^ á á á • ^ í } c í Á , í c @ Á á • ^ í } c Á V S Ú Á } [ c [ & [ í Á æ } á á & á } @ ^ Á • ~ í c Á • Á • ~ ] [ ] c á É

W ~ T [ í Á \* [ Ç ^ ! ] { ^ } c h { æ á ] • ^ í Ç ^ ! • Á • ^ í Ç ^ ! } ç æ í á á [ í Á • ^ í - É • } • á á ^ í c á - í & æ c Á • Á & [ ] æ í Á á c [ Á , ^ á á & [ ] c ^ } c h ^ í Ç ^ ! • É

HÉ Y Á & [ ] c í á á ~ c Á æ á • í • @ Á ^ í - [ í { ç } & Á [ ] ^ Á • [ ~ ! ^ Á • & æ } ^ í á æ } á á c @ ^ Á } • ^ • c h & ~ í æ c á á á æ c Á • c @ [- Á • á á ] Á \* [ Ç ^ ! ] { ^ } c h @ [ • c ] æ { ^ • É