Certificate Misissuance – A Review Paper

Martin Weise

Vienna University of Technology, Austria
martin.weise@student.tuwien.ac.at

***Abstract*—Many modern internet applications rely on secure connections through HTTPS. The underlying public key infrastructure requires certificates to verify a website's identity and issue a digital certificate to it. If a certificate deviates from standards or community best practices, there is a compatibility issue that can lead to a fraction of users that cannot make use of secure connections. With this background, Kumar et al have developed a certificate linter that verifies these standards to encourage compatibility. In this paper I am going to reproduce their findings by using the same methodology and input data.**

I. Introduction

Secure connections using the HTTPS protocol require a functional public key infrastructure with certificate authorities issuing correct certificates in terms of best practices and standards. The former can be found in the CA/Browser forum baseline requirements, the latter in RFC 5280. Both sources are used by the original authors of the paper I am going to review and are codified into the certificate linter `zlint` [1].

One of the authors quickly provided me with the snapshot [2, 3] used in the original paper and successfully verified the checksum. In this paper I focus on reproducing the results of tables I, III and Fig. 2 of the original paper. My results are listed in detail in the appendix of this paper.

II. Experimental Evaluation

Since the tool `zlint` is open source and can be used as a Go library, I wrote a small application [4] to verify the results using the same version of `zlint` [1] as the authors in the original paper. It reads the text file from the snapshot on per-line base, collecting all certificates that were valid on 23. July 2017. A total of $3,790$ certificates were unable to be parsed, $3,788$ (99.95%) of them because of unparseable asn1 structures. These parseable certificates were processed each by `zlint`, saving the relevant processing results data into a comma separated file which in a next step is imported into a PostgreSQL relational database. The data then is aggregated using a structured query language, trying to reproduce the results of the paper.

My reproduced results for the largest authorities (see TABLE I) partly differ. Especially the total number of certificates issued by GeoTrust Inc. and GlobalSign is approximately half of the expected number. These certificates also contain much more errors and warnings than in the original paper. This could be the result of different interpretations of myself since the query sums up only unique (by fingerprint) certificates that match the regex "ˆGeoTrust.*" for GeoTrust Inc. and "ˆGlobalSign.*" respective.

A similar reproduced results occurs for the most common `zlint` errors and warnings (see TABLE III). The lint "ExtKeyUsage not critical" hardly found certificates (107 instead of 26K in the original paper) that violated RFC 5280 §4.2. This could be a indicator that my lint definitions somehow strongly differ from the one used in the original paper or the snapshots somehow are not the same. For better understanding which lints I used, the assumed lint name is given in the table. It is assumed that `zlint`'s built-in attributes `ErrorsPresent` and `WarningsPresent` are true exactly if-and-only-if a error (or a warning respective) is present.

The reproduced results for the misissuance rates (see Fig. 2) highly differ from the ones in the original paper. Taking into consideration the possibility of different definitions of lints due to different versions of the software, there is still a major difference in total certificates per year. This could be due to different interpretations or methodology. I interpreted Fig. 2 of the original paper as "certificates that were valid at least 1 day in the given year determined by the fields `NotBefore` and `NotAfter` of a certificate".

III. Conclusion

Not all results from the original paper were reproducible. Since one of the authors provided me with the infos needed and verified the correctness of the snapshot, a possible reason for my different reproduced results may be caused by implementation errors and/or different interpretations of the table data in the original paper.

Acknowledgements

I thank Wilfried Mayer for the technical support and virtual machine help during the verification process. Also, I thank Edgar Weippl for the reliable supervision. This work was supported in part by SBA Research gGmbH.

References

[1] K. Deepak, J. Rudenberg, A. David, and P. Murley, "Zlint," 2018, [Online]. Available: https://github.com/zmap/zlint/, accessed: 2018-09-20, master branch hash: "7057a53b1d1daf9ee70547255cf2fa516ff50c3c".

[2] K. Deepak, "Private mail conversation with Mr. Deepak," 2018, note: The snapshot identifier is "ioqhe2qn4mc02brp-certificates.20170723T082310.json.lz4" and has a sha1-hash of "f163dfdc0307e8547e23727cf950bb080e00dc15".

[3] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman, "A search engine backed by internet-wide scanning," in *Proceedings of the*

22Nd ACM SIGSAC Conference on Computer and Communications Security, ser. CCS '15. New York, NY, USA: ACM, 2015, pp. 542–553. [Online]. Available: http://doi.acm.org/10.1145/2810103.2813703

[4] M. Weise, "Certificate missisuance experiment," 2018, [Online]. Available: https://github.com/CheckResearch/TrackingCertificateMisissuanceintheWild_Experiment_01, accessed: 2019-02-04.

APPENDIX

The numbering of the table corresponds with the numbering in the original paper. I denote the relative deviation in % to the values in the original paper (reference values) in parentheses.

TABLE I: Largest authorities

Issuer	Certificates	w/ Errors	w/ Warnings
Lets Encrypt	37M (0.0%)	21 (61.5%)	8
Comodo	6.3M (−6.3%)	79,900 (2382%)	82,248 (941%)
cPanel	4.7M (0.0%)	131 (0.0%)	1 (0.0%)
Symantec	2.8M (0.0%)	29,241 (26.8%)	2.7M (0.0%)
GeoTrust Inc.	672K (−64.6%)	4,019 (−29.4%)	661K (−65.2%)
GoDaddy	1.6M (0.0%)	40,047 (4.8%)	6,992 (34.8%)
GlobalSign	613K (−48.9%)	11,274 (1247%)	10,644 (4391%)

TABLE III: Most common `zlint` errors and warnings.

Error	Certificates	Lint
Subject CN not from SAN	30K (−56.8%)	`e_subject_common_name_not_from_san`
SAN extension missing	3,217 (−91.8%)	`e_ext_san_missing`
Invalid character in DNSName	25K (−15.8%)	`e_dnsname_bad_character_in_label`
AKID missing	53 (−99.8%)	`e_ext_authority_key_identifier_missing`
SAN email field present	3,168 (−73.6%)	`e_ext_san_rfc822_name_present`
Invalid TLD in DNSName	4,253 (−34.6%)	`e_dnsname_not_valid_tld`
Warning	**Certificates**	**Lint**
SKID missing	2.8M (−50.6%)	`w_ext_subject_key_identifier_missing_sub_cert`
ExtKeyUsage not critical	107 (−100.0%)	`w_ext_key_usage_not_critical`
Explicit Text not UTF-8	68K (−63.0%)	`w_ext_cert_policy_explicit_text_not_utf8`
Policy contains NoticeRef	602 (−99.1%)	`w_ext_cert_policy_contains_noticeref`
AIA missing CA URL	3,237 (−92.1%)	`w_sub_cert_aia_does_not_contain_issuing_ca_url`
ExtKeyUsage Extra Values	3,298 (−78.0%)	`w_sub_cert_eku_extra_values`

Fig. 2: Misissuance rates

Year	Certificates	w/ Errors	w/ Warnings
2012	9.6M (320.1%)	7M (1,055%)	6.8M (2,280%)
2013	11M (283.8%)	8.1M (841.8%)	8.1M (3,269%)
2014	15M (312.3%)	11M (1,095%)	11M (11,157%)
2015	28M (304.2%)	22M (1,102%)	23M (65,637%)
2016	53M (5.6%)	41M (1,109%)	43M (177,247%)
2017	116M (13.4%)	51M (1,412%)	50M (216,753%)