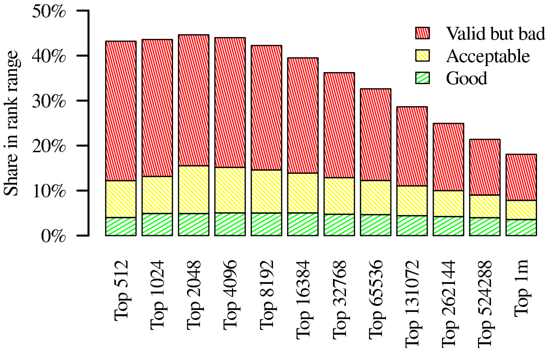

Figure 16: Certificate quality in relation to Alexa rank.

gated all issuers that had issued more than 2 certificates and found that all of them seemed to employ a PKI of their own, but without Root Certificates in Firefox. The reasons for this use of Version 1 are unknown to us, but plausible causes are software default settings or an arbitrary choice to use the simpler format of Version 1.

Certificate Quality: A Summarizing View.

We conclude our analysis with a summarizing view of certificate quality. Figure 16 reveals that the share of valid certificates (=total height of the bars) is negatively correlated with the Alexa rank. This does not come as a surprise, since operators of high-profile sites with a higher Alexa rank can be expected to invest more resources into a working HTTPs infrastructure. What is surprising, however, is that even in the top 512 or 1,024 – i.e., truly high-ranking sites – only just above 40% of certificates are absolutely valid. The figure also shows a classification of the certificates in three categories, which we have termed 'good', 'acceptable' and 'valid but bad'. *Good* certificates have correct chains, correct host names, a chain length of 2 at most, do not use MD5 as a signature algorithm, use non-weak keys that are either DSA keys or RSA keys of more than 1024 bit, and have a validity of at most 13 months (a year plus a grace period). For *acceptable* keys, we require the same but allow a chain length of 3 and validity is allowed to be up to 25 months. *Valid but bad* keys represent the remainder of keys (with correct chains and host names, but that otherwise fail our criteria). Interestingly, although high-profile sites are more likely to deliver valid certificates, the share of bad certificates among their valid certificates is much higher (about two thirds) when compared to that of all sites (about half). The conclusion we have to draw here is that even among the absolutely valid certificates the number of certificates with strong properties is disappointingly low.

6. DISCUSSION AND CONCLUSION

By combining and evaluating several actively and passively obtained data sets, which were in part also obtained over 1.5 years, we were able to derive a very comprehensive picture of the X.509 infrastructure as used with TLS/SSL. Our analysis supports with hard facts what has long been believed: that the X.509 certification infrastructure is, in great part, in a sorry state.

The most sobering result is probably the percentage of certificates that a client using the Mozilla Root Store would accept without a warning: just 18%. This can be traced to both incorrect certification chains (40% of all certificates exhibit this problem), but even more so to incorrect or missing host names in the subject or subject alternative name. With self-signed certificates, where conscientious operators would have an incentive to use the correct host name, the situation was much worse. The only positive point is that the percentages of absolutely valid certificates have increased since 2009, but then again only very slightly. Recall that these numbers refer to the top 1 million hosts – the percentage of certificates where the chains are correct is lower for the full IPv4 space than for the top sites, as we found by examining the EFF data set.

Moreover, many certification chains showed more than one error. Expired certificates were common, and so were certificates for which no Root Certificate could be found. A further problematic finding is that all our data sets reveal a high number of certificates that are shared between a large number of hosts. This is even the case for high-profile Web hosters – and often, the host names do not match the certificates there, either. Although offered by several CAs, Extended Validation certificates do no seem to be in wide use.

This truly seems a sorry state, and it does not come as a surprise that users just click away warnings, thus adding to the general insecurity of the WWW. As few CAs are responsible for more than half of the distinct certificates, one should think the situation should be better or at least easier to clean up.

There are some more positive tendencies that should be mentioned, however. Our evaluation shows that the more popular a given site is, the more likely it supports TLS/SSL at all, and the more likely it shows an absolutely valid certificate. On the whole, key lengths seem not to constitute a major problem. The same is true for signature algorithms. Keys with short bit lengths are becoming fewer, and the weak MD5 algorithm is clearly being phased out, too. Over the past 1.5 years, we also found an increase in the use of intermediate certificates while chain lengths remained remarkably short. This is a good development, as end-host certificates should not be issued by a Root Certificate that is used in online operations. However, use of intermediate certificates can be overdone and sometimes is. The latter will add to the already complex certification infrastructure that we have encountered, with a high number of distinct certification chains.

Concerning our passive monitoring, the data we obtained allowed us to evaluate negotiated properties of TLS/SSL associations, which cannot be obtained by active scans. We were able to determine the negotiated ciphers and digest mechanisms. Our observations lead us to conclude that most connections use secure ciphers with acceptable key lengths, and that key lengths are even increasing over time, with a good security margin. This is in line with the corresponding trends in certificate keys. However, MD5 is still commonly used to compute Message Authentication Codes (MACs). This is not really dramatic at the moment – even now, collision attacks take at least some seconds, and preimage attacks much longer. An attacker's time window is likely too short for a useful on-the-fly attack at this time (note that TLS/SSL uses a replay protection for the payload, too). However, we cannot see a compelling reason to continue using MD5 for MACs (except possibly for some very old legacy clients). We feel it should be discouraged and phased out.

With the above mentioned problems in certificates, however, we have to conclude that the positive movements do not address the most pressing problem, which is the certification structure itself. Unfortunately, the general state seems to persist as made evident by the fact that several critical factors like validity of certification chains, correct host names in certificates, or plainly the number of hosts that offer TLS/SSL do not seem to change. As this work has focused on the top 1 million hosts (scans) and the PKI as accessed by users (monitoring), this seems to indicate a pressing need.

7. ACKNOWLEDGEMENTS

We thank Nathan Evans and the Leibniz-Rechenzentrum (LRZ) in Munich for assisting us with obtaining the data. We especially thank Helmut Reiser of LRZ for his valuable feedback. This work has been partially funded by the European Union within the EU FP7 project ResumeNet (FP7-224619).

8. REFERENCES

[1] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, and W. Polk, "Internet X.509 public key infrastructure certificate and certificate revocation list (CRL) profile," RFC 5280 (Proposed Standard), May 2008.

[2] T. Dierks and E. Rescorla, "The transport layer security (TLS) protocol version 1.2," RFC 5246 (Proposed Standard), Aug. 2008, updated by RFCs 5746, 5878, 6176.

[3] J. Appelbaum, "Detecting certificate authority compromises and Web browser collusion," Blog entry: https://blog.torproject.org/blog/detecting-certificate-authority-compromises-and-web-browser-collusion, 2011, [online; last retrieved in May 2011].

[4] Mozilla Security Blog, "DigiNotar removal follow up," https://blog.mozilla.com/security/2011/09/02/diginotar-removal-follow-up/ [online; last retrieved in September 2011], 2011.

[5] C. Herley, "So long, and no thanks for the externalities: the rational rejection of security advice by users," in *Proc. 2009 Workshop on New Security Paradigms*. New York, NY, USA: ACM, 2009.

[6] C. Ellison and B. Schneier, "Ten risks of PKI: What you're not being told about public key infrastructure," *Computer Security Journal*, vol. 16, no. 1, 2000.

[7] P. Gutmann, "PKI: It's not dead, just resting," *IEEE Computer*, vol. 35, no. 8, August 2002.

[8] P. Eckersley and J. Burns, "An observatory for the SSLiverse," Talk at Defcon 18., July 2010, [last retrieved in May 2011]. [Online]. Available: https://www.eff.org/files/DefconSSLiverse.pdf

[9] P. Eckersley and J. Burns, "Is the SSLiverse a safe place?" Talk at 27C3. Slides from https://www.eff.org/files/ccc2010.pdf [online; last retrieved in May 2011], 2010.

[10] I. Ristic, "Internet SSL Survey 2010," Talk at BlackHat 2010. Slides from https://media.blackhat.com/bh-us-10/presentations/Ristic/BlackHat-USA-2010-Ristic-Qualys-SSL-Survey-HTTP-Rating-Guide-slides.pdf, 2010, [online; last retrieved in May 2011].

[11] I. Ristic, "State of SSL," Talk at InfoSec World 2011. Slides from http://blog.ivanristic.com/Qualys_SSL_Labs-State_of_SSL_InfoSec_World_April_2011.pdf, 2011, [online; last retrieved in May 2011].

[12] Alexa Internet Inc., "Top 1,000,000 sites (updated daily)," http://s3.amazonaws.com/alexa-static/top-1m.csv.zip, 2009–2011, [online; last retrieved in May 2011].

[13] H. K. Lee, T. Malkin, and E. Nahum, "Cryptographic strength of SSL/TLS servers: Current and recent practices," in *Proc. 7th ACM SIGCOMM Conference on Internet Measurement (IMC)*, San Diego, CA, USA, October 2007.

[14] S. Yilek, E. Rescorla, H. Shacham, B. Enright, and S. Savage, "When private keys are public – results from the 2008 Debian OpenSSL vulnerability," in *Proc. 9th ACM SIGCOMM Conference on Internet Measurement (IMC)*, Chicago, Illinois, USA, Nov. 2009.

[15] Data sets of active scans, http://pki.net.in.tum.de, 2011.

[16] A. Croll and S. Power, *Complete Web Monitoring*. O'Reilly Media, 2009.

[17] L. Braun, G. Münz, and G. Carle, "Packet sampling for worm and botnet detection in TCP connections," in *Proc. IEEE/IFIP Network Operations and Management Symposium (NOMS)*, Apr. 2010.

[18] S. Kornexl, V. Paxson, H. Dreger, A. Feldmann, and R. Sommer, "Building a time machine for efficient recording and retrieval of high-volume network traffic," in *Proc. 5th ACM SIGCOMM Conference on Internet Measurement (IMC)*, Berkeley, CA, USA, Oct. 2005.

[19] L. Braun, A. Didebulidze, N. Kammenhuber, and G. Carle, "Comparing and improving current packet capturing solutions based on commodity hardware," in *Proc. 10th ACM SIGCOMM Conference on Internet Measurement (IMC)*, Nov 2010.

[20] F. Fusco and L. Deri, "High speed network traffic analysis with commodity multi-core systems," in *Proc. 10th ACM SIGCOMM Conference on Internet Measurement (IMC)*, Nov 2010.

[21] V. Paxson, "Bro: a system for detecting network intruders in real-time," *Computer networks*, vol. 31, no. 23-24, 1999.

[22] H. Dreger, A. Feldmann, M. Mai, V. Paxson, and R. Sommer, "Dynamic Application-Layer Protocol Analysis for Network Intrusion Detection," in *Proc. USENIX Security Symposium*, Apr. 2006.

[23] Planet Lab, "Planet Lab Web site," https://www.planet-lab.org [online; last retrieved in May 2011].

[24] The International Grid Trust Federation, "IGTF Web site," http://www.igtf.net/ [online; last retrieved in May 2011].

[25] A. Klein, "Attacks on the RC4 stream cipher," *Designs, Codes and Cryptography*, vol. 48, 2008.

[26] E. Rescorla, "HTTP over TLS," RFC 2818 (Informational), 2000.

[27] CA/Browser Forum, "EV SSL certificate guidelines version 1.3," http://www.cabforum.org/Guidelines_v1_3.pdf, 2010, [online; last retrieved in May 2011].

[28] M. Stevens, A. Lenstra, and B. de Weger, "Chosen-prefix collisions for MD5 and colliding X.509 certificates for different identities," in *Advances in Cryptology – EUROCRYPT 2007*, ser. LNCS. Springer Berlin / Heidelberg, 2007, vol. 4515.

[29] NIST, "Approved Algorithms," http://csrc.nist.gov/groups/ST/toolkit/secure_hashing.html, 2006, [online; last retrieved in May 2011].

[30] A. Sotirov, M. Stevens, J. Appelbaum, A. Lenstra, D. Molnar, D. A. Osvik, and B. de Weger, "MD5 considered harmful today," http://dl.packetstormsecurity.net/papers/attack/md5-considered-harmful.pdf, 2008, [online; last retrieved in May 2011].

[31] T. Kleinjung, K. Aoki, J. Franke, A. Lenstra, E. Thomé, J. Bos, P. Gaudry, A. Kruppa, P. Montgomery, D. Osvik, H. te Riele, A. Timofeev, and P. Zimmermann, "Factorization of a 768-bit RSA modulus," in *Advances in Cryptology – CRYPTO 2010*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, vol. 6223.

[32] NIST, "Special publications (800 Series)," http://csrc.nist.gov/publications/PubsSPs.html, 2011, [online; last retrieved in May 2011].

Summary Review Documentation for

“The SSL Landscape – A Thorough Analysis of the X.509 PKI Using Active and Passive Measurements”

Authors: R. Holz, L. Braun, N. Kammenhuber, G. Carle

Reviewer #1

Strengths: The paper examines some extensive data sets. It compares and contrasts the information gleaned via active and passive measurements. The paper is (generally speaking) presented in a manner that makes the work appealing to a broad audience. The paper is on an important topic, and the results may help improve the security of the Internet.

Weaknesses: In places, the presentation needs improvement (i.e., some sections seem like they were written hastily). The paper needs to more clearly distinguish between Internet and Grid certificates.

Comments to Authors: Abstract: The location of the clients used in the active measurements did not.seem to have much effect on the results (nor would I have expected any), so I would not highlight that in the abstract, I would only mention it briefly in the methodology section. The abstract should summarize the key results, not make vague statements about being able to “perform additional analyses” (this is one example where the text seems to have been hastily written).

Introduction: Negligence is not the only possible reason for problems; malice, blackmail, etc. are some of the others. The fact that “several measurement points” were used is stated/implied multiple times, but since it really did not have an effect on the results, I would de-emphasize this. Immediately prior to describing the “remainder of the paper”, summarize the key findings of your study.

Section 2: Provide a diagram of a certificate (i.e., that shows the common information fields contained in a certificate). I would include a “conceptual” diagram of PKI; figure 1 is a bit detailed to start off with. Unless you can cite a credible source regarding the source of the CA attack, delete the statement “supposedly from Iran” (it is not relevant to the issue any way); also, please clarify that the attacker gained the signing authority of the CA. Rather than just providing the number of root CAs in the root store for August 2009 to December 2010, it would be interesting to look at much older versions of Netscape/Mozilla, and plot the growth.

Section 3: One paper to add is H. Lee, T. Malkin and E. Nahum, “Cryptographic strength of SSL/TLS servers: current and recent practices”, IMC 2007. (if not in this section, than in Section 5.2). Clarify that Ristic gave a talk at InfoSec 2011, rather than had a paper there

Section 4: Where will the data be released? Clarify which nmap scan was done, and why that type of scan(s) was chosen. Please clarify why a new scanning tool was developed; Lee et al. developed one that likely would have provided all of the required features (perhaps it is not publicly available), and workload generation tools like httperf also have the ability to do what is described in this section. Please clarify how much encrypted “Internet” was on non-standard ports (rather than encrypted “Grid” traffic); was it really worth trying to analyze all ports? Have the Bro patches been shared with Bro developers? Again, multiple client sites are mentioned; but it doesn’t seem like this mattered, so why keep mentioning it?

Section 5: Create a separate subsection for the discussion of filtering Grid-related certificates (e.g., “5.1 Data Pre-processing”). Although the few few paragraphs discuss filtering the Grid-related certificates, there is still discussion later in the paper about Grid traffic; this is confusing to the reader; please either completely filter it from the data set, or skip the discussion of filtering, and then do a more exhaustive job of describing the results for “Internet” certificates versus “Grid” certificates.

Reviewer #2

Strengths: The Web PKI is a critical component of the Internet, timely given the uptick in HTTPS, and as far as I am aware there is no comprehensive empirical characterization of it that matches this paper. If true that is reason enough to accept it. The results themselves are perhaps not very surprising in retrospect (lots of certificates are expired or have other problems!) but there are some gems (like 1000 different certificates that have duplicate private keys!) and it is valuable to see how much minor dysfunction is below the surface.

Weaknesses: Well, it’s a bit of a dry characterization that is not surprising for the most part. However, it is still valuable unless it has been reported previously.

Comments to Authors: Thanks for an interesting read - I’ve always wondered about certificates.

Please include your key results / contributions in the abstract and introduction. What are they? For instance, near the end of the paper it says “the more popular a site, the more likely it supports SSL, and the fewer the problems”. I did not get this from the results I read, perhaps because there are so many small results that it’s hard to keep the overall thrust together.

I was confused about self-certs. You describe these as for a host, but I believe it is common enough for an organization like a university to self-cert themselves as a root. Is this true, and can you separate this from other “missing root” errors? Basically, it is likely not an error for the users.

Can you tell us for each error what is common browser behavior? This would help to put the errors in perspective; I wasn’t sure if they would all result in a user alert, or not.

The error that surprised me most was that mismatches between CN and domain are very common. Surely this is a significant security problem? Please say more about what it means in practice.

Can you sketch a little more of the TLS mechanics wrt certificates? You imply that a server sends the certificate chain on SSL connect. I thought it was a multi-step process.

Reviewer #3

Strengths: It's good to have quantitative numbers on what everyone already suspects: that the x509 system is badly deployed and vulnerable.

Weaknesses: There are no surprises or take-away points from this paper that are not already published.

Comments to Authors: The main benefit of this paper in my mind is to more completely quantify what people already know and have previously studied [4,7,9]. That said, it's not clear to me that a more in-depth study of something that's known to be broken will be of interest to the IMC audience.

The text in S4.1 makes it sound like you scanned 1M _clients_ on port 443; you really scanned 1M popularly advertized web servers, right? The "acceptable use" clause in the CFP would make the former a Bad Thing, so you may want to consider making this more clear.

For this paper to be accepted, I would really like to see a more concrete example of the impact on users: are the sites with broken certificates still being visited (and thus people are just ignoring the "security") or are these just accidental clicking or probes. One thing that was not clear was the frequency distribution of certs: how common are the most common certs and what does that tail look like? Are there any very common certs that are broken and people keep using them?

I don't understand why multiple geographic locations at the same time is a benefit for your active probing or why different locations have different results: where there really connectivity problems or where you exploiting CDN phenomena? Fig 6 seems to show that the results were mostly comparable, no?

Figures 5 (a), 5(b) and 6 should probably be collapsed down to a single "stacked" bar graph for ease of comparison and space savings.

I thought that the explanation of the x509 protocol in in S2 was something that most people knew and could probably have been removed.

In general, I think a lot could have been done to shorten descriptions in the paper and make better use of space: maybe this points to this data being presented in a short paper format?

Reviewer #4

Strengths: The authors processed very large sets of collected data, a impressive amount of work. The results present an overall picture of the x509 certificate deployment, with various pieces of interesting results.

Weaknesses: The goal of the paper seems a bit unclear to me. After reading the paper one ends with lots data but a vague picture of the current stage of affairs. The paper started by claiming that "it always has been felt that the certification processes of this (x509) PKI may lack in stringency, resulting in a deployment where many certificates do not meet the requirements of a secure PKI.", but did not offer a concrete conclusion at the end with regard to that statement, i.e. does/not the deployment lack stringency? If so, what may be the causes of the problems? What need to change?

Comments to Authors: In this paper the authors processed impressive amount of collected data, and collected a variety of interesting stats regarding the deployment of x509 certificates (e.g. fig-3, the decreasing use of MD5; fig-4, more than 100K hosts sharing one single certificate?! Fig-5 & fig-6, percentages of different errors--but some inconsistency between the figures and text, see below).

High level comment: the overall results seem just like a collection, a bit lacking of stringency and focus.

The paper stated that "we evaluated our datasets with respect to the properties of certificates and certification chains, also correlating this with specific host properties (e. g. popularity)" Maybe the paper could be made stronger by first nailing down exactly what kinds of "properties" of the certificates that the paper would want to *focus* on, and for what purpose (e.g. is the goal to examine the effectiveness of the certificates in protecting the data).

Without a clearly defined focus, the results look rather diverse.

Reviewer #5

Strengths: The first study of X.509 Certificates in the wild for the purpose of determining the strength of PKI certification that focus on highly ranked domains and performed live monitoring of SSL/TLS to focus on used certificates, aiming at filtering out invalid certificates. The study reveals several pressing problems of the PKI infrastructure.
Extensive implementation and setup to capture SSL flows at a high speed and to comprehensively study deployed X.509 certificates. The authors successfully combined multiple tools (TNAPI, Bro, SNI, PlanetLab etc), which requires a substantial amount of effort.
Well written paper. Clearly explains the motivation of their study and their methodology.
Interesting statistics on the types of ciphers and their frequency used in SSL/TLS. Interesting statistics on how many times the same certificate occurs on multiple hosts.
Interesting statistics on the validity of the certificate chains. Some unexpected errors were found. 60% of trust chains were valid. It's good to have quantitative numbers on what everyone already suspects: that the x509 system is badly deployed and vulnerable.

Weaknesses: The methodology is quite straight forward and used a technique introduced in [13,14,15,16]. No innovative techniques were introduced.

It would be great if the authors were able to motivate all the statistics presented in the paper. For example, knowing the frequency with which ciphers appear in SSL would help us do ...

Related work: Mishari Al Mishari, Emiliano De Cristofaro, Karim El Defrawy and Gene Tsudik, "Harvesting SSL Certificate Data to Mitigate Web-Fraud", another large scale study of certificates.

Comments to Authors: Overall, I enjoyed reading the paper, I learned aspects of the real PKI deployment and obtained several statistics regarding the PKI. The authors could do a better job motivating the use of their statistics. It is a good paper. Its main weakness is that it is not introducing any non-obvious measurement techniques. But it is asking good questions and the answers the authors obtain are satisfactory.

Response from the Authors

In the following, we describe the major changes we have made to our original submission, as requested by the reviewers. The focus of the paper was improved by stating our goals and contributions in the introduction, as well as listing properties that we concentrate on. We restructured the paper to improve the storyline: basically, we mimic the steps that a client executes while validating a certificate (errors in certificate chains, host names etc.). This is followed by an investigation of certificate properties that impact the security of a SSL/TLS association (e.g., hash algorithms and public keys). Finally, we turn to deployment issues and discuss the security relevance of our findings. For every certificate property that we discuss, we added an introductory explanation that describes the importance of the property in question, and how a correct certificate should look like in this respect. At several points, we highlight typical client behavior when encountering a faulty certificate, in particular bad certification chains and self-signed certificates. As requested by the reviewers, we clarified the distinction between Grid and non-Grid certificates. Grid certificates are now only discussed when presenting the properties of SSL/TLS traffic, whereas they are filtered out from the discussion of certificate properties. As was proposed, we extended the discussion of data in some sections and reduced it in others. For example, we shortened the discussion of certificates from different locations where the results are not surprising. However, we disagree with the reviewers in the point that globally distributed scans should not be expected to show differences in the result; therefore, we added a section that outlines interesting differences encountered between our scanning locations. Further discussions of issuers of the most frequent certificates or keys affected by the Debian OpenSSL vulnerability have been added. The validity of certificates and host names contained therein is presented in greater detail, with reference to the effects in clients. We highlight that only about 18% of all certificates are actually counted as valid in a client. To make the paper more accessible, we furthermore updated the background section and added more information about X.509. In light of recent attacks on CAs (StartCom, DigiNotar), we moreover included descriptions of these attacks. Last but not least, we would like to thank our anonymous reviewers for their valuable feedback, and wish to particularly highlight the work of reviewer A.