An internet-wide analysis of TLS configurations of public LDAP servers

Steven Monteiro
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
s.c.monteiro@student.utwente.nl

ABSTRACT

LDAP is a protocol designed for querying and updating directory structures over IP. A common use case for the protocol is storing sensitive information such as passwords, creating a potential target for attackers. Despite this, we find no prior research quantifying the presence of public LDAP servers on the internet or investigating the security of these servers. This research investigates both of these points by performing an internet-wide scan for LDAP servers on well-known ports 389 and 636 and analyzing the TLS configurations of a sample of found instances. We discover over 6.6 million open ports, and observe over 29 thousand valid LDAP banners in our sample. We find major differences between port 389 and 636 in terms of preferred cipher suites and the validity of presented certificates. Some of our findings are encouraging from a security standpoint, while others leave to be desired.

Keywords

Internet-wide scan, LDAP, Reachability, Security, TLS

1. INTRODUCTION

The Lightweight Directory Access Protocol (LDAP) [31] is a protocol designed to query and update directory structures over the Internet Protocol (IP). Although multiple implementations exist (e.g. [22]), it is most frequently deployed in the form of the popular Active Directory software [18]. There are no limitations on the type of data that can be stored in an LDAP directory; however, common use cases are storing passwords as part of a single sign-on (SSO) setup and storing general contact information.

This combination of popularity and potentially sensitive data makes LDAP servers which are reachable over the internet a potential target for attackers. Despite this, we are not aware of any previous work which attempts to quantify LDAP servers in the wild or investigate their security. Preliminary evidence of a substantial presence can be found in the datasets of Rapid7's Project Sonar, which identifies 3.2 million hosts which have LDAP port 389 open alone [25]. To aid in security, LDAP provides support for TLS connections in the form of a STARTTLS command. We expect support for TLS in LDAP to be widespread since the most recent standard mandates its support if the implementation supports any form of non-anonymous authentication [10].

We first investigate what the reachability of publicly accessible LDAP servers on the internet is. Having established a set of reachable LDAP servers, we then ask what the TLS configurations of these public accessible servers are. We split this into three questions: Which protocol versions and cipher suites do they prefer if given free choice? What are the properties of certificates presented by these servers? And no less importantly, are these preferences and properties considered secure by current standards?

Looking ahead at our results, we discover over 6.6 million open ports across the two well-known LDAP ports 389 and 636. Using a stratified sample, we observe over 29 thousand LDAP banners. The corresponding servers are then used to determine TLS connection and certificate details. We find significant differences between these details on port 389 and 636. Additionally, we find that the most secure option is not always preferred by the server.

The paper will first discuss several key concepts used within the research, as well as related work in this area. We then give the methodology by which we answer our research questions. We follow with the detailed results obtained from our measurements and discuss the outcomes. Finally, we conclude the paper and give suggestions for future work to build upon our results.

2. BACKGROUND

In this section, we identify several key concepts used within the research, and acknowledge related work.

2.1 Banners

We use a port scan to determine which IP addresses may contain LDAP servers on port 389 and 636. In an ideal world, the number of responses would exactly equal the number of servers. Unfortunately, it is not uncommon for services to be run on ports other than the ones officially assigned to them. This creates the possibility of finding non-LDAP servers at LDAP ports, and these should not be included in our analysis for obvious reasons. Additionally, in an effort to ward off attacks, a middlebox intercepting the traffic between the scanning server and the scanned IP may cause the scanned IP to fail to respond to application-layer data despite responding to a TCP handshake. [14]

This problem can be addressed through a technique known as banner grabbing. Rather than only using raw TCP, as is the case with common scan types, banner grabbing establishes an application-layer connection and attempts to get the server to return its banner, which often contains

35th Twente Student Conference on IT July 2nd, 2021, Enschede, The Netherlands.

the server's vendor and version as well as the features it supports. A valid banner strongly suggests that there is a real server present. In the case of LDAP, the role of the banner is fulfilled by the root DSE[1], a directory entry located at the root of the directory structure. It is queried using the same syntax as a regular entry.

2.2 TLS

Transport Layer Security (TLS) is the *de facto* standard for encryption on the internet. It has been incorporated into many different application-layer protocols, most notably HTTPS, but also including LDAP. The protocol has had multiple iterations, starting with its predecessor Secure Sockets Layer (SSL). Although this paper uses the term "TLS" to refer to the encryption used in LDAP, we do not exclude SSL from our analysis.

TLS rests upon a complex architecture of cryptographic algorithms, which we do not dive into here for the sake of brevity. We instead briefly highlight the aspects of TLS that are most relevant for our research.

STARTTLS Depending on manner in which the application-layer protocol integrates TLS, connections over TLS can either be established immediately upon connecting, or initiated in plaintext and converted to TLS at any moment using a special STARTTLS command. LDAP has chosen for an implementation using STARTTLS, often in combination with a restriction that no queries can be made until a secure connection has been set up [9].

Protocol versions Not all versions of SSL and TLS are currently considered secure. All SSL and TLS versions with the exception of TLS 1.2 and TLS 1.3 have been deprecated due to various security vulnerabilities [30, 3, 19]. We highlight the POODLE attack, which allows an encrypted message to be extracted one byte at a time with no knowledge of the private key [21].

Cipher suites The cipher suite defines the set of cryptographic algorithms used in the connection. The strength of a cipher suite therefore depends on the strengths and weaknesses of each individual included algorithm.

Certificates TLS allows clients and servers alike to use X.509 certificates [4] to prove their identity, although it is common that only the server does so. The validity of the server certificate often strongly contributes to the trustworthiness of the server. A valid certificate meets criteria such as chaining up to a root certificate which has been trusted previously. A list of trusted root certificates, known as a root store, is often built into clients.

Key types and lengths The key type and key length are properties of a certificate. The key type is generally RSA or ECDSA in practice, but other algorithms can technically be supported [23]. The key length then determines the strength of the key, following the old adage that higher is better, although performance trade-offs may occur.

2.3 Related work

We are not aware of any previous work which addresses the questions posed in this research. Specifically, there appears to be no academic research into the reachability of LDAP servers in the wild or their security in any form. Rapid7's Project Sonar has been publishing raw datasets of networks scans that could be used in such research efforts [24], but no (human-written) reports have been published based on the LDAP datasets to our knowledge.

Durumeric et al. [7] introduced ZMap, which is used within this research to perform an internet-wide scan on known LDAP ports. Scanning speed and reliability is discussed, as well as ethical considerations arising from scanning large amounts of IP addresses, and the authors give recommendations related to these ethical concerns. We adopt their recommendations as described in Appendix A.

Holz et al. [12] performed active and passive measurement of TLS traffic and found that several issues can be identified related to X.509 certificates in the wild. Durumeric et al. [6] expanded upon this by performing a periodic internet-wide scan for HTTPS servers and raised comparable concerns.

Holz et al. [11] later addressed email and chat applications, which may also make use of TLS, by means of active and passive scanning. These applications resemble LDAP in their use of STARTTLS. The authors found "a worryingly high number of poorly secured servers".

3. METHODOLOGY

Our research methodology consists of multiple scans, which are each described here, and the analysis thereof. We first perform an internet-wide scan for IPv4 addresses that respond on the known LDAP ports 389 and 636. This is followed by a banner grab on all responsive hosts to establish that an actual LDAP server is present. If this is the case, we follow up with a STARTTLS command to obtain the properties of the TLS connection as well as the certificate presented by the server.

3.1 Internet-wide scan

The internet-wide scan is a SYN scan performed on all IP addresses within the IPv4 address space, with the exception of known bogon IPs such as those reserved for private networks [26]. Additionally, a blocklist is used consisting of IP ranges which have previously been requested to be excluded from internet-wide research scans, obtained through abuse complaints received by previous researchers. Two unique ports are scanned: 389 and 636. These two ports are the IANA-assigned ports for LDAP and LDAP over TLS respectively [13].

The scans are performed using the network scanning tool ZMap [7], which was chosen for its proven high scanning speed compared to similar tools such as Nmap. Since ZMap does not support scanning multiple ports in one invocation, the scan is split into two consecutive runs. The scanning is monitored using the tcpdump [28] tool.

A SYN scan artificially performs the first step of a TCP handshake and is therefore necessarily performed over TCP. This implies that only TCP services will be detected. This imposes a limitation on the completeness of our scan, since Active Directory is an LDAP server with rudimentary support for LDAP over UDP [17]. We make this trade-off because of the additional complexity of detecting UDP services, and note here that Rapid7's Project Sonar identified responses over UDP port 389 from approximately 15 thousand IP addresses during its scans [25].

As discussed in Section 2.1, this scan by itself does not establish the reachability of LDAP servers on port 389 and 636, but rather the reachability of a superset thereof. To establish that a server speaks LDAP, banner grabbing needs to be performed, which we describe in Section 3.3.

3.2 Stratification

Our further scans make use of the network scanner Nmap

[1] DSA-Specific Entry, where DSA stands for Directory System Agent.

192.0.2.0/24	198.51.100.0/24	203.0.113.0/24
192.0.2.8	**198.51.100.81**	203.0.113.95
192.0.2.234		**203.0.113.199**
		203.0.113.207

Table 1. An example of the stratification process. In this example, 3 IPs are sampled from a set of 6.

[16] due to its powerful scripting engine that includes a collection of pre-written scripts. Although Nmap is by no means less capable than ZMap, it is significantly slower when scanning an equal number of hosts [7]. A back-of-the-envelope calculation suggests that performing the application-level scans on the full set of responding hosts using Nmap would take approximately 14 days. In the interest of time, we use a stratified sample of the responding hosts on each port for further scans. The samples are created by grouping IP addresses by the prefix they originate from using the pyasn library [2]. A single IP is then selected uniformly at random from each prefix. An example of this process can be found in Table 1.

3.3 Banner and certificate grabbing

If and only if a host-port combination has responded to the SYN scan, it is a candidate for a banner grab to be performed to determine if an LDAP server is truly present. As explained in Section 2.1, this need arises from the fact that it is not uncommon for services to be run on ports other than the ones assigned to it, or to fail to respond to application-layer data despite responding to an initial SYN scan.

The banner grab consists of using the ldap-rootdse script [15] included in Nmap to query for the root DSE of the LDAP server. The LDAP protocol specification allows for this request to be explicitly refused by the server, but a well-formed refusal would still confirm the presence of an LDAP server. If no response is received at all, the connection is retried. This scan is similarly split into two consecutive runs, one for each port, and monitored using tcpdump.

The output of this scan is two sets of IP addresses that can confidently be stated to host LDAP servers. With these two sets, it is now possible to attempt to establish a TLS connection in order to retrieve the certificate and connection properties. This is done by the ssl-cert script in Nmap [8]. This script supports SSL 3.0, TLS 1.0, TLS 1.1, TLS 1.2, and TLS 1.3, as well as all cipher suites. In accordance with the protocol, this causes the server to select its most preferred protocol version and cipher suite. We again retry the connection if no response is received and monitor the scan using tcpdump.

3.4 Analysis

In order to answer our research questions about connection and certificate properties, we turn to analyzing the connections made by ssl-cert. Monitoring the scans using tcpdump allows us to analyze the scan traffic using Zeek [29]. The stated purpose of Zeek is as a security monitoring tool, observing network traffic and providing condensed logs of actions. We make use of its ability to analyze pre-existing pcap files produced by tcpdump to extract connection and certificate details without manually inspecting packets. In this paper, we focus on a number of connection-related and certificate-related properties: the protocol version of the connection, the cipher suite used in the connection, the validity of the presented certificate, and the key type and length of the certificate. We limit the analysis of the key type and length to end host certificates in the interest of time.

Hosts responding on both ports	2,893,630
Hosts responding only on port 389	317,856
Hosts responding only on port 636	562,866

Table 2. Responses to SYN scan.

	389		636	
Banner	23,401	(41.76%)	5,926	(12.18%)
No banner	28,956	(51.67%)	38,417	(78.93%)
Undecodable	3,684	(6.57%)	4,330	(8.90%)

Table 3. Responses to ldap-rootdse.

To validate certificates, we make use of Zeek's built-in functionality, which passes the actual validity checking on to the OpenSSL library. The Mozilla root store [20] is used as the list of trusted root certificates. We note that the main purpose of the Mozilla root store is for web browsing, and it is therefore conceivable that LDAP-specific roots have not been included. Furthermore, in addition to the certificates already in the root store, certificates can manually be trusted by users and system administrators, such as root certificates managed locally by organizations. As such, certificates that appear as untrusted may in fact be trusted by the intended clients.

4. RESULTS

All scans were performed between June 7 and June 18, 2021 inclusive from a single IP address located in Australia. At most one scan was running at any given time, and scans were not paused while in progress. However, the ZMap scan of port 636 was restarted from the beginning once due to a misconfiguration of the scanning server unrelated to the scanning activity. All scans made use of a blocklist containing 507 prefixes, representing 330,274,957 excluded IP addresses.[2]

4.1 Internet-wide scan

Table 2 shows that the majority of hosts responding to our SYN scan respond on both scanned ports. 8.42% of hosts only respond on port 389, and approximately 1.75 times as many only respond on port 636. It is tempting to conclude from these numbers that port 636 is more popular than port 389, but as discussed previously (Section 2.1), we have not established that the hosts speak LDAP or respond to any application-layer data at all.

4.2 Stratification

Employing the described stratification method (Section 3.2), 56,041 hosts are sampled from the responses on port 389, and 48,673 are sampled from port 636. 31,409 hosts appear in both samples. The two samples are treated separately in all scans and analyses described further on in this paper. We include percentages in all tables to facilitate the comparison between port 389 and port 636.

4.3 Banner and certificate grabbing

Table 3 visualizes the responses to the ldap-rootdse script. It immediately stands out that hosts are significantly more likely to return a banner on port 389 than on port 636. Several explanations are possible: port 636 could be more popular with services that stay silent upon receiving an LDAP handshake, or could have more middleboxes that

[2] 64 more IP addresses were excluded while scanning was in progress due to an abuse complaint received by email.

	389		636	
TLS 1.3	0	(0.00%)	1,568	(18.02%)
TLS 1.2	686	(99.56%)	7,071	(81.26%)
TLS 1.1	0	(0.00%)	2	(0.02%)
TLS 1.0	2	(0.29%)	61	(0.70%)
SSL 3.0	1	(0.15%)	0	(0.00%)

Table 4. Protocol versions selected during ssl-cert.

filter application-layer data. Research by Izhikevich et al. [14] suggests that this explanation is plausible, as also discussed in Section 2.1. A possibility within LDAP itself is that some LDAP servers do not accept any command other than STARTTLS while the connection is in plaintext [9].

"Undecodable" refers to responses that could not be decoded as LDAP messages, and were found in less than 9% of hosts on both ports. It is likely that the vast majority, if not all of these responses were sent by servers speaking a different protocol than LDAP. In particular, an HTTP server should return "400 Bad Request" upon receiving the query, a response which we recorded 1,059 times across both ports.

4.4 Analysis

During the analysis, it has become clear that it is not uncommon for a server to respond to a TLS handshake without finishing it. This behavior has been observed from 614 hosts. Since only partial data is available for the associated connections, they have been excluded from the analysis.

It is clear from Table 4 that the vast majority of scanned LDAP servers prefers the secure cryptographic protocols TLS 1.2 and TLS 1.3. The high adoption rate for TLS 1.2 is not surprising, since the specification was released in 2008 [5] and is as such already over a decade old.

18.02% of hosts selected TLS 1.3 as the protocol version on port 636. Unlike previous versions of SSL and TLS, TLS 1.3 encrypts the sending of the server certificate [27], which makes it impossible for passive monitors such as Zeek to obtain it. Although the information is available to us, this makes it more difficult to analyze the properties of TLS 1.3 connections. Due to time constraints, we decide to exclude TLS 1.3 connections and certificates from further analysis. We encourage future researchers investigating LDAP to include TLS 1.3 in their work.

A surprising result can be seen in Table 5, namely that the set of cipher suites preferred by LDAP servers on port 389 is completely different to the set on port 636. The number of unique cipher suites observed also differs majorly, with 18 unique cipher suites being observed on port 636, compared to just 3 on port 389. Arguably even more surprising is that in all but 5 cases, hosts responding on both ports select a different cipher suite when contacted on port 389 than on port 636. This strongly suggests that LDAP implementations tailor their responses based on port numbers.

We theorize that this allows the LDAP server to maintain compatibility with older clients that do not support a wide range of cipher suites. Specifically, port 389 would be used where a secure connection with port 636 is not possible. We however have insufficient data to support this, and we therefore recommend further research to determine the exact cause of this phenomenon.

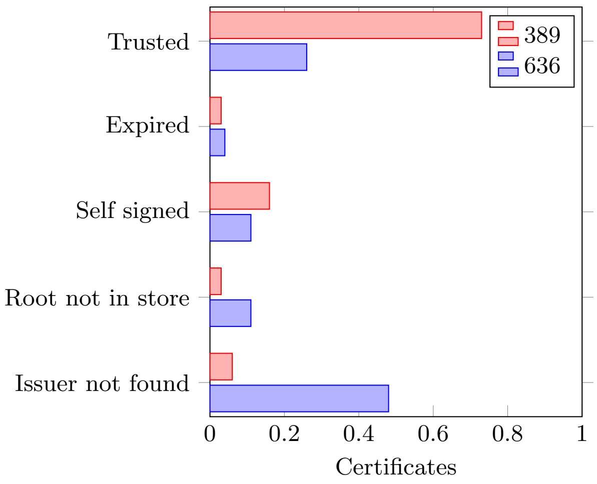

Figure 1. Validity status of certificates seen during ssl-cert.

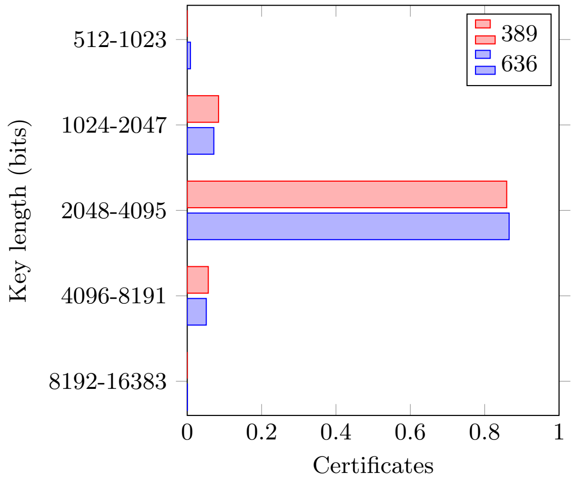

Figure 2. Key lengths of RSA end host certificates seen during ssl-cert.

Also noteworthy is the apparently common preference for suites that do not provide perfect forward secrecy (PFS). This is a potential security concern. In our sample, these are all suites that use RSA rather than Diffie-Hellman key exchange, that is, suites that do not have "DHE" or "ECDHE" in their names in Table 5.

Figure 1 again shows large differences between port 389 and 636, this time in terms of certificate validity. Whereas the majority of certificates from port 389 are trusted, certificates on port 636 are more likely to be signed by an unknown issuer, indicating that the issuing certificate was not sent by the server and not contained in the Mozilla root store. This suggests a potential significant presence of root certificates managed locally by organizations. We recommend that future work investigate these certificate chains in more detail.

Opposite to the cipher suites, we find that all but 1 of the hosts responding on both ports responded with the same certificate chain in both cases. In the single deviant case, the certificate was issued by an unknown issuer on port 389 and self signed on port 636.

We plot the observed key lengths of end host certificates making use of RSA in Figure 2. The figure reveals that key lengths of 2048-4095 bits are the most popular on both ports. This is not a surprising result, since 2048 bits is often the default setting of certificate generation tools, as

	389		636	
TLS_DHE_RSA_WITH_AES_128_CBC_SHA	0	(0.00%)	682	(9.56%)
TLS_DHE_RSA_WITH_AES_256_CBC_SHA	111	(16.11%)	5	(0.07%)
TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256	0	(0.00%)	247	(3.46%)
TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA384	0	(0.00%)	904	(12.67%)
TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384	0	(0.00%)	2,740	(38.41%)
TLS_RSA_WITH_AES_128_CBC_SHA	576	(83.60%)	18	(0.25%)
TLS_RSA_WITH_AES_256_CBC_SHA	0	(0.00%)	2,213	(31.02%)
TLS_RSA_WITH_AES_256_GCM_SHA384	0	(0.00%)	187	(2.62%)
Other	2	(0.29%)	138	(1.93%)

Table 5. Cipher suites selected during ssl-cert.

well as being the most common key length issued by certificate authorities (CAs). Across both ports, 626 (8.00%) hosts presented a certificate with a key length less than or equal to 1024 bits. These key lengths are considered weak by current cryptographic standards. In addition to the RSA certificates shown in Figure 2, we observe 20 (0.28%) certificates making use of elliptic curve cryptography (ECC) in the form of ECDSA. Of these, 1 certificate is trusted, issued by Let's Encrypt. We believe that the cause of this low adoption rate is the relative novelty of ECDSA combined with a desire to support legacy clients.

5. DISCUSSION

It is encouraging to see that the vast majority of LDAP servers in our sample select a secure protocol version, that is to say TLS 1.2 or 1.3. As mentioned previously (Section 3.4), all older protocol versions have been deprecated because of their use of weak cryptography. Similarly encouraging are the RSA key lengths observed in end host certificates. Over 90% of hosts on both ports use a key of at least 2048 bits. This is in line with the deprecation of keys 1024 bits and under.

We however must conclude that this positive news cannot be extended to all connection properties. Most notably, the preferred suites observed in our scans leave to be desired, with an overwhelming more than 80% of hosts on port 389 selecting a cipher that uses RSA key exchange, in addition to over 30% of hosts on port 636. RSA key exchange does not have the property of perfect forward secrecy (PFS), implying that a private key compromise can result in the decryption of past connections. It is our recommendation that LDAP implementations deprecate the use of cipher suites with RSA key exchange, or at least change the order of preference so that suites using Diffie-Hellman key exchange are preferred when supported.

We additionally find that a significant percentage of hosts presents certificates that are not trusted by the OpenSSL-based validation procedure in Zeek. In Section 3.4, we explain that this does not automatically imply that they are untrusted by the intended clients, since extra certificates can be added manually. If an organization making use of a local root certificate has taken proper security precautions to protect its private keys, this is not a security vulnerability. Their local nature unfortunately makes this impossible to assess.

Despite this, we think it is not unreasonable to expect the servers found in our scans to present a publicly trusted certificate, since all of the queried servers were after all publicly accessible via the internet. Such certificates can be obtained for free via Let's Encrypt [1]. Given this fact, we propose that servers currently serving an expired or self-signed certificate switch to a certificate from Let's Encrypt or any other provider of choice, and that organizations unwilling to do so update their firewall and VPN policies so as to make the server invisible to external clients which do not carry the necessary local root certificate.

6. CONCLUSION AND FUTURE WORK

In this paper, we described and executed a methodology to find public LDAP servers on the internet, a subject which has thus far been untouched in literature. We found over 6.6 million open ports distributed across two well-known LDAP ports. This was followed up by a banner grab making use of a stratified sampling technique, where we found over 29 thousand valid LDAP banners being returned. We subsequently attempted to establish TLS connections with these hosts and analyzed and discussed the properties of the connections and certificates presented by the server. In this process, we found major differences between the properties of port 389 and 636. Our study shows that while some statistics from our sample are encouraging from a security standpoint, such as the selected protocol version, others leave to be desired by introducing unnecessary weaknesses, such as the preferred cipher suite.

We acknowledge the limited scope of our sample and the analysis of our results. It is our hope that our results give rise to future work that is more comprehensive in scope. We highlight our decision to make use of a stratified sample, which may be replaced with a scan of the full set of found IP addresses given sufficient resources. We suggest that future work include the certificates and connection properties of TLS 1.3. Additionally, it remains unclear what causes the discrepancy between the results on port 389 and 636. Finally, we believe it may be of interest to investigate the certificate ecosystem of LDAP in more detail, so as to compare it to other services using STARTTLS or to HTTPS.

7. REFERENCES

[1] J. Aas, R. Barnes, B. Case, Z. Durumeric, P. Eckersley, A. Flores-López, J. A. Halderman, J. Hoffman-Andrews, J. Kasten, E. Rescorla, S. Schoen, and B. Warren. Let's Encrypt: An Automated Certificate Authority to Encrypt the Entire Web. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. ACM, Nov. 2019.

[2] H. Asghari and A. Noroozian. pyasn. https://pypi.org/project/pyasn/.

[3] R. Barnes, M. Thomson, A. Pironti, and A. Langley. Deprecating Secure Sockets Layer Version 3.0. RFC 7568, RFC Editor, June 2015.

[4] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, and T. Polk. Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation

List (CRL) Profile. RFC 5280, RFC Editor, May 2008.

[5] T. Dierks and E. Rescorla. The Transport Layer Security (TLS) Protocol Version 1.2. RFC 5246, RFC Editor, August 2008.

[6] Z. Durumeric, J. Kasten, M. Bailey, and J. A. Halderman. Analysis of the HTTPS certificate ecosystem. In *Proceedings of the 2013 conference on Internet measurement*. ACM, Oct. 2013.

[7] Z. Durumeric, E. Wustrow, and J. A. Halderman. ZMap: Fast Internet-wide Scanning and Its Security Applications. In *22nd USENIX Security Symposium (USENIX Security 13)*, pages 605–620, Washington, D.C., Aug. 2013. USENIX Association.

[8] D. Fifield. ssl-cert. https://nmap.org/nsedoc/scripts/ssl-cert.html.

[9] A. Findlay. Best Practices in LDAP Security. In *LDAPCon 2011*, Oct. 2011.

[10] R. Harrison. Lightweight Directory Access Protocol (LDAP): Authentication Methods and Security Mechanisms. RFC 4513, RFC Editor, June 2006.

[11] R. Holz, J. Amann, O. Mehani, M. Wachs, and M. A. Kaafar. TLS in the Wild: An Internet-wide Analysis of TLS-based Protocols for Electronic Communication. In *Proceedings 2016 Network and Distributed System Security Symposium*. Internet Society, 2016.

[12] R. Holz, L. Braun, N. Kammenhuber, and G. Carle. The SSL Landscape: A Thorough Analysis of the X.509 PKI Using Active and Passive Measurements. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement - IMC '11*. ACM Press, 2011.

[13] Internet Assigned Numbers Authority. Service Name and Transport Protocol Port Number Registry. https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml.

[14] L. Izhikevich, R. Teixeira, and Z. Durumeric. LZR: Identifying Unexpected Internet Services. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021.

[15] P. Karlsson. ldap-rootdse. https://nmap.org/nsedoc/scripts/ldap-rootdse.html.

[16] G. Lyon. Nmap. https://nmap.org/.

[17] Microsoft. 3.1.1.3.1.11 LDAP Search Over UDP. https://docs.microsoft.com/en-us/openspecs/windows_protocols/ms-adts/3fad0ec9-414c-432a-ba0b-837c74091dd6.

[18] Microsoft. Active Directory Lightweight Directory Services Overview. https://docs.microsoft.com/en-us/previous-versions/windows/it-pro/windows-server-2012-r2-and-2012/hh831593(v=ws.11).

[19] K. Moriarty and S. Farrell. Deprecating TLS 1.0 and TLS 1.1. BCP 195, RFC Editor, March 2021.

[20] Mozilla Corporation. Mozilla's CA Certificate Program. https://wiki.mozilla.org/CA.

[21] B. Möller, T. Duong, and K. Kotowicz. This POODLE Bites: Exploiting The SSL 3.0 Fallback. Technical report, Google, September 2014.

[22] OpenLDAP Foundation. OpenLDAP. https://openldap.org/.

[23] T. Polk, R. Housley, and L. Bassham. Algorithms and Identifiers for the Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. RFC 3279, RFC Editor, April 2002.

[24] Rapid7. Open Data. https://opendata.rapid7.com/.

[25] Rapid7. TCP Scans. https://opendata.rapid7.com/sonar.tcp/.

[26] Y. Rekhter, R. G. Moskowitz, D. Karrenberg, G. J. de Groot, and E. Lear. Address Allocation for Private Internets. BCP 5, RFC Editor, February 1996.

[27] E. Rescorla. The Transport Layer Security (TLS) Protocol Version 1.3. RFC 8446, RFC Editor, August 2018.

[28] The Tcpdump Group. tcpdump. https://www.tcpdump.org/.

[29] The Zeek Project. Zeek. https://zeek.org/.

[30] S. Turner and T. Polk. Prohibiting Secure Sockets Layer (SSL) Version 2.0. RFC 6176, RFC Editor, March 2011.

[31] K. D. Zeilenga. Lightweight Directory Access Protocol (LDAP): Technical Specification Road Map. RFC 4510, RFC Editor, June 2006.

APPENDIX

A. ETHICAL CONSIDERATIONS

As described in our methodology (Section 3), we scan a large amount of IP addresses and attempt to establish connections with responsive hosts. While we do not perform these actions with malicious intent or exploit any vulnerabilities, network administrators often view this kind of activity as an attack on their network. Durumeric et al. [7] cover this problem and list seven recommendations to alleviate it. We take their suggestions into account as follows: prior to scanning, we inform the local network administrators and verify that the network can sustain our scans at the configured speed; and the source IP address of the scan hosts a web page stating the goal of the research and an email address that can be used to opt out of the scans.

Furthermore, we make use of a preexisting blocklist that has been built up over time by other researchers following similar guidelines. During the scanning, we received one abuse complaint via email. The prefix was immediately blocked on the scanning server to prevent further connections to the IPs, and the IPs were later discarded from the dataset.