A (brief) survey of S/MIME certificate issuance

Overview

- Now that there's a standard governing S/MIME certificate issuance, it would be informative to analyze known S/MIME certificates against the standard
- There's no Certificate Transparency for S/MIME, so the set of known S/MIME certificates is (probably) quite small compared to total issuance
- Despite a lack of Certificate Transparency, there are a few sources:
  - Censys.io
  - CA/B Forum mailing lists

Methodology for mailing list extraction

- Download message archives for the following CABF mailing lists:
  - Cabfpub
  - Servercert-wg
  - Smcwg-public
- Follow and download all links to .p7s attachments
- Extract all end-entity certificates found in the PKCS #7 files

- This yielded about 160 S/MIME certificates

Methodology for Censys.io

- Download all end-entity certificates with the id-kp-emailProtection EKU and are marked as “trusted” or “was-trusted” by Censys.io

- This yielded 1960 certificates

- Combined totals from the mailing lists and Censys was roughly 2090 certificates (some certificates were found in both locations)

Breakdown by validation type

Validation level	Count
Individual	1406
Sponsored	532
Mailbox	145
Organization	11

This breakdown is heuristic based, as no CA includes the reserved policy OIDs (yet). Determination of validation level was done by examining the subject attributes:

- If givenName and surname are present, but organizationName is absent, then Individual
- If givenName, surname, and organizationName are present, then Sponsored
- If organizationName and commonName are present, and they are not the same value and the commonName doesn't contain a "@", then Sponsored
- If organizationName is present and all of the above are not true, then Organization
- If none of the above are true, then Mailbox

S/MIME certificate profile compliance

- In the dataset, 73 unique “issuers” were identified
  - For Issuer DNs that contain an organizationName, that value is the “issuer” ID
  - For Issuer DNs that do not contain a commonName, that value is the “issuer” ID
- This means that there is likely not a 1:1 mapping of Issuer IDs to CA owners
- Mapping to a root certificate owner may also not be the most useful, as a single CA owner may have multiple issuance systems

Fatal

- These errors are the result of incorrectly encoded fields or extensions that prevent further analysis

- 11 issuers

Fatal

- These errors are the result of incorrectly encoded fields or extensions that prevent further analysis

- 11 issuers
- Most common problem is that subject email addresses were not encoded using IA5String

Errors

- All issuers have not included the required reserved policy OIDs as well the organizationIdentifier attributes, so those findings were ignored

Finding	Issuer count
Invalid algorithm encoding (old SHA-1)	18
Validity period too long	15
Invalid explicitText encoding type	13
Prohibited keyUsage bit	12
Prohibited EKU	11
Subject email address is not in SAN	4
Prohibited subject attribute	4
RSA modulus length is too small	3

Errors, cont’d

- And a long tail of various errors, such as:

- No key usage extension
- Authority key identifier extension was absent
- Prohibited SAN type
- Invalid keyUsage extension encoding

Future areas for exploration

- Expand set of data sources in the absence of CT
  - More mailing lists
  - Crawling Google, etc.
- Refine the definition of an “Issuer” to better convey scope / weighting
- Determine scope of “viewable” issuance as opposed to “total” issuance
  - Perhaps CRLs can be useful for this?
  - If determined, then metrics such as “average validity period” would be feasible