Internet Engineering Task Force (IETF)
Request for Comments: 8358
Updates: 5485
Category: Informational
ISSN: 2070-1721

R. Housley
Vigil Security

March 2018

Update to Digital Signatures on Internet-Draft Documents

Abstract

RFC 5485 specifies the conventions for digital signatures on Internet-Drafts. The Cryptographic Message Syntax (CMS) is used to create a detached signature, which is stored in a separate companion file so that no existing utilities are impacted by the addition of the digital signature.

The RFC Editor recently published the first RFC that includes non-ASCII characters in a text file. The conventions specified in RFC 7997 were followed. We assume that non-ASCII characters will soon start appearing in Internet-Drafts as well. This document updates the handling of digital signatures on Internet-Drafts that contain non-ASCII characters in a text file.

This document updates RFC 5485.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are candidates for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc8358.

Table of Contents

1. Introduction

RFC 5485 [IDSIG] specifies the conventions for digital signatures on Internet-Drafts. The Cryptographic Message Syntax (CMS) [CMS] is used to create a detached signature, which is stored in a separate companion file so that no existing utilities are impacted by the addition of the digital signature.

The RFC Editor recently published the first RFC that includes non-ASCII characters in a text file. The conventions specified in RFC 7997 [RFCED] were followed. We assume that non-ASCII characters will soon start appearing in Internet-Drafts as well. This document updates the handling of digital signatures on Internet-Drafts that contain non-ASCII characters in a text file.

This document updates RFC 5485 [IDSIG], which contains the conventions that have been used by the IETF Secretariat to digitally sign Internet-Drafts for the past few years. The IETF Secretariat generates the digital signature shortly after the Internet-Draft is posted in the repository.

The digital signature allows anyone to confirm that the contents of the Internet-Draft have not been altered since the time that the document was signed.

The digital signature is intended to provide a straightforward way for anyone to determine whether a particular file contains the Internet-Draft that was made available by the IETF Secretariat. The signing-time associated with the signature provides the wall clock time at which the signature was generated; it is not intended to provide a trusted timestamp.

1.1. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [STDWORDS] [STDWORDS2] when, and only when, they appear in all capitals, as shown here.

1.2. ASN.1

The CMS uses Abstract Syntax Notation One (ASN.1) [X.680]. ASN.1 is a formal notation used for describing data protocols, regardless of the programming language used by the implementation. Encoding rules describe how the values defined in ASN.1 will be represented for transmission. The Basic Encoding Rules (BER) [X.690] are the most widely employed rule set, but they offer more than one way to

represent data structures. For example, definite length encoding and indefinite length encoding are supported. This flexibility is not desirable when digital signatures are used. As a result, the Distinguished Encoding Rules (DER) [X.690] were invented. DER is a subset of BER that ensures a single way to represent a given value. For example, DER always employs definite length encoding.

2. Detached Signature Files

All Internet-Draft file names begin with "draft-". The next portion of the file name depends on the source of the document. For example, documents from IETF working groups usually have "ietf-" followed by the working group abbreviation, and this is followed by a string that helps people figure out the subject of the document.

All Internet-Draft file names end with a hyphen followed by a two digit version number and a suffix. The suffix indicates the type of file. For example, a text file will have a suffix of ".txt". Today, plain text files are the most common, but the RFC Editor has announced plans to make use of other formats [RFCSERIES]. Each file format employs a different suffix.

Going forward, one cannot assume that a text file with a suffix of ".txt" will contain only ASCII characters.

The companion signature file has exactly the same file name as the RFC or Internet-Draft, except that ".p7s" is added to the end. This file name suffix conforms to the conventions in RFC 5751 [MSG]. Here are a few example names:

```
Internet-Draft: draft-ietf-example-widgets-03.txt
Signature File: draft-ietf-example-widgets-03.txt.p7s

Internet-Draft: draft-ietf-example-widgets-03.pdf
Signature File: draft-ietf-example-widgets-03.pdf.p7s

Internet-Draft: draft-housley-internet-draft-sig-file-00.txt
Signature File: draft-housley-internet-draft-sig-file-00.txt.p7s
```

3. Additional Content Types

The CMS is used to construct the detached signatures for Internet-Drafts. The CMS ContentInfo content type MUST always be present, and it MUST encapsulate the CMS SignedData content type. Since a detached signature is being created, the CMS SignedData content type MUST NOT encapsulate the Internet-Draft. The CMS detached signature is summarized in RFC 5485 [IDSIG].

The SignedData.SignerInfo.EncapsulatedContentInfo.eContentType value MUST identify the format of the Internet-Draft that is being signed. Section 5 of RFC 5485 [IDSIG] lists the file formats and the associated content type. This document expands that list as follows:

File Format	Content Type
ASCII text	id-ct-asciiTextWithCRLF
UTF-8 text (includes non-ASCII)	id-ct-utf8TextWithCRLF
HyperText Markup Language (HTML)	id-ct-htmlWithCRLF
EPUB	id-ct-epub
Extensible Markup Language (XML)	id-ct-xml
Portable Document Format (PDF)	id-ct-pdf
PostScript	id-ct-postscript

The object identifiers associated with the content types listed above table are:

```
id-ct OBJECT IDENTIFIER ::= { iso(1) member-body(2)
     us(840) rsadsi(113549) pkcs(1) pkcs9(9) smime(16) 1 }

id-ct-asciiTextWithCRLF OBJECT IDENTIFIER ::= { id-ct 27 }

id-ct-utf8TextWithCRLF OBJECT IDENTIFIER ::= { id-ct 37 }

id-ct-htmlWithCRLF OBJECT IDENTIFIER ::= { id-ct 38 }

id-ct-epub OBJECT IDENTIFIER ::= { id-ct 39 }

id-ct-xml OBJECT IDENTIFIER ::= { id-ct 28 }

id-ct-pdf OBJECT IDENTIFIER ::= { id-ct 29 }

id-ct-postscript OBJECT IDENTIFIER ::= { id-ct 30 }
```

4. Need for Canonicalization

In general, the content of an Internet-Draft is treated like a single octet string for the generation of the digital signature. Unfortunately, the text and HTML files require canonicalization to avoid signature validation problems. The primary concern is the manner in which different operating systems indicate the end of a line of text. Some systems use a single new-line character, other systems use the combination of the carriage-return character followed by a line-feed character, and other systems use fixed-length records padded with space characters. For the digital signature to validate properly, a single convention must be employed.

4.1. ASCII, UTF-8, and HTML File Canonicalization

The canonicalization procedure follows the conventions used for text files in the File Transfer Protocol (FTP) [FTP]. Such files must be supported by FTP implementations, so code reuse seems likely.

The canonicalization procedure converts the data from its internal character representation to the standard 8-bit NVT-ASCII representation (see TELNET [TELNET]). In accordance with the NVT standard, the <CRLF> sequence MUST be used to denote the end of a line of text. Using the standard NVT-ASCII representation means that data MUST be interpreted as 8-bit bytes.

Trailing space characters MUST NOT appear on a line of text. That is, the space character must not be followed by the <CRLF> sequence.

Thus, a blank line is represented solely by the <CRLF> sequence.

The form-feed nonprintable character (0x0C) is expected in Internet-Drafts. Other non-printable characters, such as tab and backspace, are not expected, but they do occur. Non-printable or non-ASCII characters (ones outside the range 0x20 to 0x7E) MUST NOT be changed in any way not covered by the rules for end-of-line handling in the previous paragraph.

Trailing blank lines MUST NOT appear at the end of the file. That is, the file must not end with multiple consecutive <CRLF> sequences.

In some environments, a Byte Order Mark (BOM) (U+FEFF) is used at the beginning of a file to indicate that it contains non-ASCII characters. In UTF-8 or HTML files, a BOM at the beginning of the file is not considered to be part of the file content. One or more consecutive leading BOMs, if present, MUST NOT be processed by the digital signature algorithm.

Any end-of-file marker used by an operating system is not considered to be part of the file content. When present, such end-of-file markers MUST NOT be processed by the digital signature algorithm.

Note: This text file canonicalization procedure is consistent with the NVT-ASCII definition offered in Appendix B of RFC 5198 [UFNI].

4.2. XML File Canonicalization

Utilities that produce XML files are expected to follow the guidance provided by the World Wide Web Consortium (W3C) in Section 2.11 of [R20081126]. If this guidance is followed, no canonicalization is needed.

A robust signature generation process MAY perform canonicalization to ensure that the W3C guidance has been followed. This guidance says that a <LF> character MUST be used to denote the end of a line of text within an XML file. Therefore, any two-character <CRLF> sequence and any <CR> that is not followed by <LF> are to be translated to a single <LF> character.

4.3. No Canonicalization of Other File Formats

No canonicalization is needed for file formats currently used or planned for Internet-Drafts other than ASCII, UTF-8, HTML, and XML files. Other file formats, including PDF [PDF], PostScript [PS], and EPUB [EPUB] are treated as a simple sequence of octets by the digital signature algorithm.

5. IANA Considerations

IANA has registered object identifiers for three content types in the "SMI Security for S/MIME CMS Content Type (1.2.840.113549.1.9.16.1)" registry as follows:

Description	OID	Specification
id-ct-utf8TextWithCRLF	1.2.840.113549.1.9.16.1.37	[RFC8358]
id-ct-htmlWithCRLF	1.2.840.113549.1.9.16.1.38	[RFC8358]
id-ct-epub	1.2.840.113549.1.9.16.1.39	[RFC8358]

6. Security Considerations

The security considerations in RFC 5485 [IDSIG] are unchanged.

7. Deployment and Operational Considerations

The deployment considerations in RFC 5485 [IDSIG] are unchanged.

8. References

8.1. Normative References

[CMS] Housley, R., "Cryptographic Message Syntax (CMS)", STD 70, RFC 5652, DOI 10.17487/RFC5652, September 2009, <https://www.rfc-editor.org/info/rfc5652>.

[EPUB] International Digital Publishing Forum, "EPUB Content Documents 3.1", January 2017, <http://www.idpf.org/epub/31/spec/epub-contentdocs.html>.

[IDSIG] Housley, R., "Digital Signatures on Internet-Draft Documents", RFC 5485, DOI 10.17487/RFC5485, March 2009, <https://www.rfc-editor.org/info/rfc5485>.

[PDF] International Organization for Standardization, "Document management -- Electronic document file format for long-term preservation -- Part 3: Use of ISO 32000-1 with support for embedded files (PDF/A-3)", ISO 19005-3:2012, 2012.

[PS] Adobe Systems Incorporated, "PostScript Language Reference Manual, third edition", Addison-Wesley Publishing Company, ISBN 0-201-37922-8, 1999.

[R20081126] Bray, T., Paoli, J., Sperberg-McQueen, M., Maler, E., and F. Yergeau, "Extensible Markup Language (XML) 1.0 (Fifth Edition)", World Wide Web Consortium Recommendation REC-xml-20081126, November 2008, <http://www.w3.org/TR/2008/REC-xml-20081126>.

[STDWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[STDWORDS2] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[X.680] ITU-T, "Information Technology - Abstract Syntax Notation One: Specification of Basic Notation", Recommendation X.680, ISO/IEC 8824-1:2002, 2002.

[X.690] ITU-T, "Information technology -- ASN.1 encoding rules: Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER) and Distinguished Encoding Rules (DER)", ITU-T Recommendation X.690, ISO/IEC International Standard 8825-1:2008, November 2008.

8.2. Informative References

[FTP] Postel, J. and J. Reynolds, "File Transfer Protocol", STD 9, RFC 959, DOI 10.17487/RFC0959, October 1985, <https://www.rfc-editor.org/info/rfc959>.

[MSG] Ramsdell, B. and S. Turner, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification", RFC 5751, DOI 10.17487/RFC5751, January 2010, <https://www.rfc-editor.org/info/rfc5751>.

[RFCED] Flanagan, H., Ed., "The Use of Non-ASCII Characters in RFCs", RFC 7997, DOI 10.17487/RFC7997, December 2016, <https://www.rfc-editor.org/info/rfc7997>.

[RFCSERIES] Flanagan, H. and N. Brownlee, "RFC Series Format Requirements and Future Development", RFC 6949, DOI 10.17487/RFC6949, May 2013, <https://www.rfc-editor.org/info/rfc6949>.

[TELNET] Postel, J. and J. Reynolds, "Telnet Protocol Specification", STD 8, RFC 854, DOI 10.17487/RFC0854, May 1983, <https://www.rfc-editor.org/info/rfc854>.

[UFNI] Klensin, J. and M. Padlipsky, "Unicode Format for Network Interchange", RFC 5198, DOI 10.17487/RFC5198, March 2008, <https://www.rfc-editor.org/info/rfc5198>.

Acknowledgements

The idea for the Internet-Draft signature file came from a discussion with Scott Bradner at IETF 69 in Chicago, IL, USA. Many helpful suggestions came from Jim Schaad, Pasi Eronen, Chris Newman, and Glen Barney. Glen Barney also played a key role in implementing Internet-Draft signatures as specified in RFC 5485 [IDSIG].

Author’s Address

Russell Housley
Vigil Security, LLC
918 Spring Knoll Drive
Herndon, VA 20170
United States of America

Email: housley@vigilsec.com