Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: April 16, 2015

T. Fossati
Alcatel-Lucent
H. Tschofenig
ARM Ltd.

October 13, 2014

Datagram Transport Layer Security (DTLS) over Global System for Mobile Communications (GSM) Short Message Service (SMS)
draft-fossati-dtls-over-gsm-sms-01

Abstract

This document specifies the use of Datagram Transport Layer Security (DTLS) over the Global System for Mobile Communications (GSM) Short Message Service (SMS).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 16, 2015.

Table of Contents

1. Introduction

This document specifies the use of DTLS [RFC6347] over GSM SMS [GSM-SMS] for securing end-to-end communication between Mobile Stations (i.e. devices implementing the GSM SMS communication standard).

DTLS provides communications privacy for applications that use datagram transport protocols and allows client/server applications to communicate in a way that is designed to prevent eavesdropping and detect tampering or message forgery.

SMS is a generic transport protocol for narrow-band end-to-end communication between devices, and is an integral part of the GSM network technology.

1.1. Usage of DTLS and SMS in CoAP M2M Environments

One of the main reasons for defining a DTLS/SMS binding is its envisaged usage in machine-to-machine (M2M) communication.

Specifically, M2M environments based on the CoAP protocol mandate DTLS for securing transactions between endpoints -- as detailed in Section 9 of [RFC7252], and further articulated in [I-D.ietf-dice-profile], while the [OMA-LWM2M] architecture identifies SMS as an alternative transport for CoAP messages.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

This specification requires readers to be familiar with all the terms and concepts that are described in [GSM-SMS], [WAP-WDP], [RFC5246], and [RFC6347].

3. DTLS over SMS

3.1. Data Coding Scheme

The remainder of this specification assumes Mobile Stations are capable of producing and consuming 8-bit binary data encoded Transport Protocol Data Units (TPDU).

3.2. Handshake Overview

DTLS adds an additional roundtrip to the TLS [RFC5246] handshake to serve as a return-routability test for protection against certain types of DoS attacks. Thus a full blown DTLS handshake comprises up to 6 "flights" (i.e. logical message exchanges), each of which is then mapped on to one or more lower layer PDUs using the segmentation and reassembly (SaR) scheme described in Section 4.2.3 of [RFC6347]. The overhead for said scheme is 6 bytes per Handshake message which, given a realistic 10+ messages handshake, would amount around 60 bytes across the whole handshake sequence.

(Note that the DTLS SaR scheme is defined for handshake messages only. In fact, Record Layer messages are never fragmented and MUST fit within a single transport layer datagram, whatever be the limit imposed by the underlying transport.)

SMS provides an optional segmentation and reassembly scheme as well, known as Concatenated short messages (see Section 9.2.3.24.1 of [GSM-SMS]). However, since the SaR scheme in DTLS can't be circumvented, the Concatenated short messages mechanism SHOULD NOT be used during handshake to avoid redundant overhead. Before starting the handshake phase (either actively or passively), the DTLS implementation MUST be explicitly configured with the PMTU of the SMS transport in order to correctly instrument its SaR function. The PMTU SHALL be 133 bytes if WDP-based multiplexing is used (see Section 3.5), 140 bytes otherwise.

It is RECOMMENDED to use the established security context over the longest possible period (possibly until a Closure Alert message is

received, or after a very long inactivity timeout) to avoid the expensive re-establishment of the security association.

3.2.1. X.509 Certificate-based Authentication Caveats

X.509 certificate-based authentication (used in Certificate mode CoAP) exacerbates the number of TPDUs -- especially those involved in flight 4 and 5 -- needed to complete the handshake phase.

In such case, given the typical latency of the SMS transport, the time to finalise the handshake could be in the order of 10s of seconds (maybe even minutes).

More importantly, the large number of TPDUs involved increases the likelihood to incur packet loss which DTLS does not handle efficiently. In fact, the DTLS timeout and retransmission logics apply to whole flights, but not to message fragments individually. So, loss or delay of a single fragment may disrupt the current flight, which needs to be entirely retransmitted.

Depending on the delay and packet loss characteristics of the network link, completing a DTLS handshake which involves exchanging X.509 data may prove to be a daunting task [[CREF1: TODO: substantiate with figures]].

For these reasons, it is advisable to carefully consider whether the use of X.509 certificate-based authentication is compatible with the characteristics of the network link between the involved parties.

3.3. Message Segmentation and Re-Assembly

[RFC6347] requires that each DTLS message fits within a single transport layer datagram

The content of an SMS message is carried in the TP-UserData field, and its size may be up to 140 bytes. As already mentioned in Section 3.2, longer (i.e. up to 34170 bytes) messages can be sent using a segmentation and reassembly scheme known as Concatenated SMS (see Section 9.2.3.24.1 of [GSM-SMS]).

This scheme consumes 6-7 bytes (depending on whether the short or long segmentation format is used) of the TP-UserData field, thus reducing the space available for the actual content of the SMS message to 133-134 bytes per TPDU.

Though in principle a PMTU value higher than 140 bytes could be used (which may look like an appealing option given its more efficient use of the transport) there is a significant number of disadvantages to

consider (apart from the fixed tax of 7 bytes per TPDU to be paid to the SaR function):

1. high sensitivity to packet loss -- since there is no automatic recovery mechanism in case one TPDU in the chain is lost, and since the SaR function is transparent to the application layer, then a PMTU worth of data may be discarded even if just 1/255th of the data were lost;

2. some networks may support the Concatenated SMS function partially, if at all;

3. TPDU reordering may delay data delivery to the application;

4. high buffering requirement on both ends of the communication path.

For these reasons, the Concatenated short messages mechanism SHOULD NOT be used, and it is RECOMMENDED to leave the same PMTU settings used during the handshake phase (Section 3.2), i.e. 133 bytes if WDP-based multiplexing is enabled (Section 3.5), 140 bytes otherwise.

Note that, after DTLS handshake has completed, any fragmentation and reassembly logics that pertains the application layer - e.g. segmenting CoAP messages into DTLS records and reassembling them after the crypto operations have been successfully performed - needs to be handled by the application that uses the established DTLS tunnel.

3.4. DTLS State Machine Timers Adjustments

[RFC6347] recommends an initial timer value of 1 second with exponential back off up to no less then 60 seconds. Given the latency characteristics of typical SMS delivery, the 1 second value can easily lead to spurious retransmissions, which in turn may lead to link congestion.

Choosing an appropriate timer value is a difficult problem due to the wide variance in latency in SMS delivery. This specification RECOMMENDS an initial timer value of 10 seconds with exponential back off up to no less then 60 seconds.

If SMS-STATUS-REPORT messages are enabled, their receipt is not to be interpreted as the signal that the specific handshake message has been acted upon by the receiving party. Therefore, it MUST NOT be taken into account by the DTLS timeout and retransmission function.

Handshake messages MUST carry a validity period (TP-VP parameter in a SMS-SUBMIT TPDU) that is not less than the current value of the retransmission timeout. In order to avoid persisting messages in the network that will be discarded by the receiving party, handshake messages SHOULD carry a validity period that is the same as, or just slightly higher than, the current value of the retransmission timeout.

If an RTT estimator (e.g. [I-D.bormann-core-cocoa]) is already available in the protocol stack of the device, it could be used to dynamically update the setting of the retransmit timeout.

3.5. Multiplexing Security Associations

Unlike IPsec, DTLS records do not contain any association identifiers. Applications must arrange to multiplex between associations on the same endpoint which, when using UDP/IP, is usually done with the host/port number.

If the DTLS server allows more than one client to be active at any given time, then the WAP User Datagram Protocol [WAP-WDP] can be used to achieve multiplexing of the different security associations. (The use of WDP provides the additional benefit that upper layer protocols can operate independently of the underlying wireless network, hence achieving application-agnostic transport handover.)

The total overhead cost for encoding the WDP source and destination ports is 7 bytes out of the total available for the SMS content.

The receiving side of the communication gets the source address from the originator address (TP-OA) field of the SMS-DELIVER TPDU. This way an unique 4-tuple identifying the security association can be reconstructed at both ends. (When replying to its DTLS peer, the sender will swaps the TP-OA and TP-DA parameters and the source and destination ports in the WDP.)

4. New Versions of DTLS

As DTLS matures, revisions to and updates for [RFC6347] can be expected. DTLS includes mechanisms for identifying the version in use, and presumably future versions will either include backward compatibility modes or at least not allow connections between dissimilar versions. Since DTLS over SMS simply encapsulates the DTLS records transparently, these changes should not affect this document and the methods of this document should apply to future versions of DTLS.

Therefore, in the absence of a revision to this document, this document is assumed to apply to all future versions of DTLS. This document will only be revised if a revision to DTLS or SMS makes a revision to the encapsulation necessary.

5. Security Considerations

Security considerations for DTLS as specified in [RFC6347] apply.

In most networks, sending SMS messages is not a free service, therefore DoS attacks tend to be a lot less common than in IP networks. However, it is RECOMMENDED not to disable the cookie exchange protection, unless the involved risks are fully understood, and the chance of a DoS attack is deemed low enough to drop the protection mechanism in order to save one round-trip per handshake.

DTLS lays on top of SMS, and therefore it doesn't provide any security service to it. The SMS implementation must be able to protect itself from any special SMS message that can be used to alter the device state or configuration in an undesired way (e.g. fiddling with the private key store). Any SMS client must make sure that malicious use of such messages is not possible, for example, by filtering out certain SMS User Data header fields.

The layering of DTLS on top of the SMS transport does not introduce any new security issues. We believe that the recommendations contained in this specification (i.e. initial RTO increase, use of WDP for multiplexing security associations, avoidance of SMS SaR) have no impact on the security of DTLS.

6. Acknowledgements

Thanks to Tim Carey, Thierry Garnier, Zhiyuan Hu, Kathleen Moriarty, Eric Rescorla, Padmakumar Subramani, for helpful comments and discussions that have shaped this document.

7. IANA Considerations

This specification contains no request to IANA.

8. References

8.1. Normative References

[GSM-SMS] ETSI, "3GPP TS 23.040 V7.0.1 (2007-03): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service (SMS) (Release 7)", March 2007.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012.

[WAP-WDP] Wireless Application Protocol Forum, "Wireless Datagram Protocol", June 2001.

8.2. Informative References

[I-D.bormann-core-cocoa]
Bormann, C., Betzler, A., Gomez, C., and I. Demirkol, "CoAP Simple Congestion Control/Advanced", draft-bormann-core-cocoa-02 (work in progress), July 2014.

[I-D.ietf-dice-profile]
Tschofenig, H., "A Datagram Transport Layer Security (DTLS) 1.2 Profile for the Internet of Things", draft-ietf-dice-profile-04 (work in progress), August 2014.

[OMA-LWM2M]
OMA, "Lightweight Machine to Machine Technical Specification", 2013.

[RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, June 2014.

Authors' Addresses

Thomas Fossati
Alcatel-Lucent
3 Ely Road
Milton, Cambridge CB24 6DD
UK

Email: thomas.fossati@alcatel-lucent.com

Hannes Tschofenig
ARM Ltd.
110 Fulbourn Rd
Cambridge CB1 9NJ
UK

Email: hannes.tschofenig@gmx.net
URI: http://www.tschofenig.priv.at