Network Working Group
Internet-Draft
Intended status: BCP
Expires: January 14, 2010

C. Perkins
University of Glasgow

July 13, 2009

Guidelines for the use of Variable Bit Rate Audio with Secure RTP
draft-perkins-avt-srtp-vbr-audio-01.txt

Status of this Memo

Copyright Notice

Abstract

This memo discusses potential security issues that arise when using variable bit rate audio with the secure RTP profile. Guidelines to mitigate these issues are suggested.

Table of Contents

1. Introduction

The secure RTP framework (SRTP) [RFC3711] is a widely used framework for securing RTP sessions. SRTP provides the ability to encrypt the payload of an RTP packet, and optionally add an authentication tag, while leaving the RTP header and any header extension in the clear. A range of encryption transforms can be used with SRTP, but none of the pre-defined encryption transforms use any padding; the RTP and SRTP payload sizes match exactly.

When using SRTP with voice streams compressed using variable bit rate (VBR) codecs, the length of the compressed packets will therefore depend on the characteristics of the speech signal. This variation in packet size will leak significant amounts of information about the contents of the speech signal. For example [spot-me] shows that known phrases in an encrypted call can be recognised with high accuracy in certain circumstances, without breaking the encryption. Other work, referenced from [spot-me], has shown that the language spoken in encrypted conversations can also be recognised. This is potentially a significant security risk for some applications. This memo discusses ways in which this traffic analysis risk may be mitigated.

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 RFC 2119 [RFC2119].

2. Guidelines for use of VBR Audio with SRTP

To avoid the potential information leaks that might enable traffic analysis, VBR audio codecs that alter the size or spacing of their output according to the characteristics of the input speech signal SHOULD NOT be used with encrypted SRTP sessions.

It is safe to use variable rate coding to adapt the output of a voice codec the match characteristics of a network channel, for example for congestion control purposes, provided this adaptation done in a way that does not expose any information on the speech signal. That is, if the variation is driven by the available network bandwidth, not by the input speech (i.e. if the packet sizes and spacing are constant unless the network conditions change). VBR speech codecs can safely be used in this fashion with SRTP while avoiding leaking information on the contents of the speech signal that might be useful for traffic analysis.

3. Guidelines for use of Voice Activity Detection with SRTP

Many speech codecs employ some form of voice activity detection (VAD) to either suppress output frames, or generate some form of lower-rate comfort noise frames, during periods when the speaker is not active. If VAD is used on an encrypted speech signal, then some information about the characteristics of that speech signal can be determined by watching the patterns of voice activity. This information leakage is less than with VBR coding since only the lengths of continuous bursts of voice activity can be determined, and not the length of individual words or phonemes, but is still potentially a concern.

The information leakage due to VAD in SRTP audio sessions can be much reduced if the sender adds an unpredictable "overhang" period to the end of active speech intervals, so obscuring their actual length. an RTP sender using VAD with encrypted SRTP audio SHOULD insert such an overhang period at the end of each talkspurt, delaying the start of the silence/comfort noise by a random interval. The length of the overhang applied to each talkspurt must be randomly chosen in such a way that it is computationally infeasible for an attacker to predict the length of that talkspurt. The audio data comprising the overhang period must be packetised and transmitted in RTP packets in a manner that is indistinguishable from the other data in the talkspurt.

The application of such a random overhang period to each talkspurt will reduce the effectiveness of VAD in SRTP sessions when compared to non-SRTP sessions. It is, however, still expected that the use of VAD will provide a significant bandwidth saving for many encrypted sessions.

4. Security Considerations

The security considerations of [RFC3711] apply.

It might be thought that it is sufficient to pad the output of a VBR codec using RTP padding to generate constant size RTP data packets as a means of mitigating the traffic analysis attacks considered here (indeed, [spot-me] suggests such a mitigation). Section 3.1 of [RFC3711] discusses potential problems with this approach, which mean that it is NOT RECOMMENDED in general.

5. IANA Considerations

No IANA actions are required.

6. Acknowledgements

This memo is based on the discussion in [spot-me]. Recent versions of ZRTP [I-D.zimmermann-avt-zrtp] contain a similar recommendation; the purpose of this memo is to highlight these issues to a wider audience, since they are not specific to ZRTP. Thanks are due to Phil Zimmermann, Stefan Dohla, Mats Naslund, Jean-Marc Valin, Gregory Maxwell, David McGrew, and Mark Baugher for their comments and feedback on this memo.

7. References

7.1. Normative References

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

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

7.2. Informative References

[I-D.zimmermann-avt-zrtp]
Zimmermann, P., Johnston, A., and J. Callas, "ZRTP: Media Path Key Agreement for Secure RTP", draft-zimmermann-avt-zrtp-15 (work in progress), March 2009.

[spot-me] Wright, C., Ballard, L., Coull, S., Monrose, F., and G. Masson, "Spot me if you can: Uncovering spoken phrases in encrypted VoIP conversation", Proceedings of the IEEE Symposium on Security and Privacy 2008, May 2008.

Author’s Address

Colin Perkins
University of Glasgow
Department of Computing Science
Glasgow G12 8QQ
UK

Email: csp@csperkins.org