Network Working Group
Internet-Draft
Intended status: BCP
Expires: September 5, 2009

C. Perkins
University of Glasgow
March 4, 2009

Guidelines for the use of Variable Bit Rate Audio with Secure RTP
draft-perkins-avt-srtp-vbr-audio-00.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) [1] 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 [3] 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 [3], 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 [2].

2. Guidelines for the use of VBR Audio with SRTP

To avoid the potential information leaks that might enable traffic analysis, VBR audio codecs SHOULD NOT be used with encrypted SRTP sessions.

Similarly, the use of voice activity detection with silence suppression or comfort noise can be considered an extreme form of VBR coding, which changes both the size and spacing of packets, and so leaks some information on the characteristics of the speech signal. Accordingly, it SHOULD NOT be used with encrypted SRTP sessions.

It is safe to use variable rate coding to adapt a speech signal to the characteristics of a network channel, for example for congestion control purposes, provided this is done in a way which 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 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. Security Considerations

The security considerations of [1] apply.

It might be thought that it is sufficient to pad the output of a VBR codec to a constant size using the RTP padding feature as a means of mitigating the traffic analysis attacks considered here (indeed, [3] suggests such a mitigation). Section 3.1 of [1] discusses potential problems with this approach, which mean that it is NOT RECOMMENDED in general.

4. IANA Considerations

No IANA actions are required.

5. Acknowledgements

This memo is based on the discussion in [3]. Recent versions of ZRTP [4] contain a similar recommendation; the purpose of this memo is to highlight the issue to a wider audience, since it is not specific to ZRTP.

6. References

6.1. Normative References

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

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

6.2. Informative References

[3] 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.

[4] 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.

Author's Address

Colin Perkins
University of Glasgow
Department of Computing Science
Sir Alwyn Williams Building
Lilybank Gardens
Glasgow G12 8QQ
UK

Email: csp@csperkins.org