



As noted, SRTP encrypts only the payload (i.e., the audio or video) for confidentiality. The authentication algorithm protects the integrity of the entire original RTP packet. The optional Master Key Identifier (MKI) and the recommended authentication tag are the only SRTP additions to the original RTP packet.

The SRTP MKI identifies which master key was used to derive the session keys currently in use for the encryption and/or authentication of the current packet. While sometimes not used, a typical MKI is 4 bytes in length, and is used in a system that may employ multiple key exchanges. While key exchanges are beyond the scope of SRTP, they are discussed a bit later in this paper.

The authentication tag is also of a configurable length, but is usually either 4 or 10 bytes long. Authentication ensures that attackers can neither modify packets in the stream nor insert (forge) additional packets. The authentication operation is performed after the encryption operation and protects the entire RTP packet. Since the sequence number is part of this protection, the authentication tag provides protection against replay.

Both the Advanced Encryption Standard Counter Mode (AES-CM) encryption method and the NULL encryption method are mandatory to implement for SRTP. The NULL method is used when no encryption (only authentication) is desired. When NULL is used, the original RTP payload remains unchanged. AES-CM is the default encryption method used in SRTP. A major reason that AES-CM was chosen was because there is no payload expansion produced (the encrypted payload is of the same length as the original payload). Another feature of AES-CM allows the processing of out-of-order packets, which also implies being able to process packets in parallel. The standard also specifies a third, optional encryption method using AES in f8-mode (AES-f8). Universal Mobile Telecommunications System (UMTS) 3G mobile networks use AES-f8.

HMAC-SHA1 as defined in RFC 2104 is the mandatory authentication algorithm. The standard recommends that RTP streams be protected with a 10-byte (80 bit) authentication tag. Notably, common VoIP payloads can be as small as 20 bytes, making the authentication tag 50 percent of the payload's size. To reduce this overhead, a 4-byte (32-bit) authentication tag can be used if the security risk using a smaller-sized tag is acceptable to the specific application, which may be the case in many VoIP implementations.

SRTP can create all the authentication and encryption keys it requires from a single master key. To do this, it uses a key derivation algorithm based on AES-CM. It is important to note that SRTP does not define the exchange of the master key. SRTP does not define any key exchange algorithms. There are numerous key exchange proposals currently before the IETF. The most commonly implemented key exchange protocol for VoIP is the Security Descriptions (SDS) protocol as defined in RFC 4568. This is the key exchange method we are using in the Interop Labs SRTP demonstrations. However, there are some limitations to SDS, such as requiring an existing secure transport for the SDS messages, and limitations in handling audio before call set up is complete ("early media"). A recent IETF meeting in March 2007 discussed alternate key exchange protocols and decided to pursue DTLS-SRTP and ZRTP (both based on Internet drafts in progress) as the basis for future work.

SRTP is an efficient, concise Internet security protocol which works well and has achieved good interoperability. However, the lack of a well accepted key exchange protocol has held back widespread implementation to date. Hopefully, future IETF work will result in a popular key exchange protocol which, when coupled with SRTP, will achieve ubiquitous VoIP security.