Vertrauliche Video-Konferenzen im Internet

Luigi Lo Iacono, Christoph Ruland

Institut für Digitale Kommunikationssysteme
Universität Siegen
Hölderlinstrasse 3
57068 Siegen
{lo_iacono,ruland}@nue.et-inf.uni-siegen.de

Zusammenfassung: Echtzeit-orientierte Multimedia-Kommunikation im Internet eröffnet eine Vielzahl neuer Anwendungen. Diese innovative Kommunikationsplattform ist gerade für weltweit operierende Unternehmen von Interesse. So können z.B. durch die Verwendung von VoIP-Lösungen oder Groupware-Applikationen Kosten gesenkt und gleichzeitig die Zusammenarbeit der Mitarbeiter optimiert werden. Dies trifft auch für Video-Konferenzsysteme zu. Anstelle regelmäßiger Meetings, die meist mit Dienstreisen eines Großteils der Teilnehmer verbunden sind, können Konferenzen virtuell durch die Übertragung von Sprach- und Videodaten über das Internet abgehalten werden.

Die Akzeptanz der beschriebenen Kommunikationsanwendungen hängt stark von den Faktoren Dienstgüte und Sicherheit ab. Die Übertragung der echtzeit-orientierten Mediendaten muss möglichst kontinuierlich erfolgen, so dass sowohl eine ruckelfreie Wiedergabe der Sprache als auch der Bewegtbilder möglich ist. Da Konferenzen firmenintern und vertraulich sind, werden sie hinter verschlossener Tür abgehalten. Das Pendant in der elektronischen Welt muss eine Entsprechung anbieten. Security-Mechanismen haben allerdings einen Einfluss auf Dienstgüteparameter. Dies muss bei der Entwicklung von Techniken zum Schutz multimedialer Kommunikation berücksichtigt und abgestimmt werden. Dieser Beitrag zeigt anhand des Beispiels eines Video-Konferenzsystems für das Internet, wie Sicherheitsmechanismen in echtzeit-orientierte Multimedia-Kommunikationsanwendungen unter Berücksichtigung von Quality of Service (QoS) integriert werden können.[1]

1 Einleitung

Das Internet ist seit seiner Entstehung Mitte der 80er Jahre rapide gewachsen. Die enorme Entwicklung umfasst nicht nur die Anzahl der an das Konglomerat von Netzen angeschlossenen Endsysteme sondern auch die Anwendungen, die das Internet als Kommunikationsplattform nutzen. Höhere Übertragungsraten im Netz, steigende Verarbeitungsgeschwindigkeiten in den Endsystemen und sinkende Kosten ermöglichen neue innovative Dienste und Applikationen wie z.B. die Internet-Telefonie (VoIP), Media-on-Demand Dienste wie Audio/Video-Server, Video-Liveübertragungen, kollaboratives verteiltes Arbeiten und Softwaredistribution.

[1] Dieses Projekt wird teilweise von der Deutschen Forschungsgemeinschaft gefördert (Projekt-Nr.: DFG – RU 600/8-1, Projekt-Titel: Internet Security System für VoIP unter Berücksichtigung von Quality of Service).

Gerade für weltweit operierende Unternehmen ist diese innovative Kommunikationsplattform von Interesse. So können z.B. durch die Verwendung von VoIP-Lösungen oder Groupware-Applikationen Kosten gesenkt und gleichzeitig die Zusammenarbeit der Mitarbeiter optimiert werden. Dies trifft auch für Video-Konferenzsysteme zu. Anstelle regelmäßiger Meetings, die meist mit Dienstreisen eines Großteils der Teilnehmer verbunden sind, können Konferenzen virtuell durch die Übertragung von Sprach- und Videodaten über das Internet abgehalten werden.

Die Akzeptanz der beschriebenen Kommunikationsanwendungen hängt stark von den Faktoren Dienstgüte und Sicherheit ab. Die Übertragung der echtzeit-orientierten Mediendaten muss möglichst kontinuierlich erfolgen, so dass sowohl eine ruckelfreie Wiedergabe der Sprache als auch der Bewegtbilder möglich ist. Da Konferenzen firmenintern und vertraulich sind, werden sie hinter verschlossener Tür abgehalten. Das Pendant in der elektronischen Welt muss eine Entsprechung anbieten. Security-Mechanismen haben allerdings einen Einfluss auf Dienstgüteparameter. Dies muss bei der Entwicklung von Techniken zum Schutz multimedialer Kommunikation berücksichtigt und abgestimmt werden.

Im Rahmen eines von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projektes arbeiten die Autoren zusammen mit einem Projektpartner aus Süd Korea an der Entwicklung und Implementierung von Sicherheitssystemen für multimediale Kommunikationsanwendungen im Internet unter Berücksichtigung von Quality of Service (QoS) Aspekten. Ergebnisse dieser Arbeit werden hier vorgestellt, wobei dieser Beitrag anhand des Beispiels eines Video-Konferenzsystems für das Internet zeigt, wie Sicherheitsmechanismen in echtzeit-orientierte Multimedia-Kommunikationsanwendungen unter Berücksichtigung von QoS integriert werden können.

Dazu werden zunächst der Aufbau von Video-Konferenzsystemen schematisch dargestellt und die für die unterschiedlichen Kommunikationspfade verfügbaren Protokolle und Standards erläutert. Anschließend wird der Einsatz und Nutzen verschiedener Security-Standards zum Schutz von Video-Konferenzsystemen diskutiert. Hierbei liegt der Fokus auf dem Transportpfad und insbesondere auf der vertraulichen Übertragung der Mediendaten. Das zur Zeit bei der IETF in der Standardisierung befindliche Internet Draft SRTP stellt sich beim Vergleich als besonders geeignet heraus, da es einerseits innovative Techniken einsetzt und andererseits zum Schutz aller bis dato verfügbaren multimedialen Kommunikationsanwendungen im Internet geeignet ist. Multimedia Kommunikationsanwendungen nutzen RTP zum Medientransport. Da SRTP die Sicherheitsdienste Vertraulichkeit und Integrität sowohl für RTP als auch für das im RTP-Umfeld eingesetzte Kontrollprotokoll RTCP bietet und über einen Schutz vor Replay-Attacken verfügt, können alle auf RTP-basierten Multimedia-Anwendungen von SRTP profitieren. Der Beitrag schließt mit der Beschreibung der entwickelten SRTP-Implementierung und der Integration dieser in das Open Source Projekt openH323. Das im openH323-Projekt enthaltene OpenPhone ist eine Video-Konferenzanwendung, die erweitert wurde und die SRTP-Implementierung zum Aufbau vertraulicher Kommunikationen nutzt. Mit der präsentierten Lösung ist es auf handelsüblichen PCs ohne spezielle kryptographische Hardware oder Rechenbeschleuniger möglich, eine vertrauliche Video-Konferenz im Internet durchzuführen. Des Weiteren ist die Implementierung so effizient, dass die Sicherheitsdienste Vertraulichkeit und Daten-

integrität erbracht werden können, ohne dafür zusätzliche Dienstgüte im Netz anfordern bzw. bereitstellen zu müssen.

2 Video-Konferenzsysteme

Video-Konferenzsysteme gehören zur Gruppe der Multimedia-Kommunikationsanwendungen. Diese sind durch ihren isochronen Übertragungsmodus gekennzeichnet. Weiterhin ist die Kommunikation durch zwei separate Kommunikationspfade charakterisiert. Ein Kommunikationspfad wird zur Signalisierung und der zweite zum Medientransport verwendet. Abbildung 1 zeigt die hierfür im Internet verfügbaren Standards.

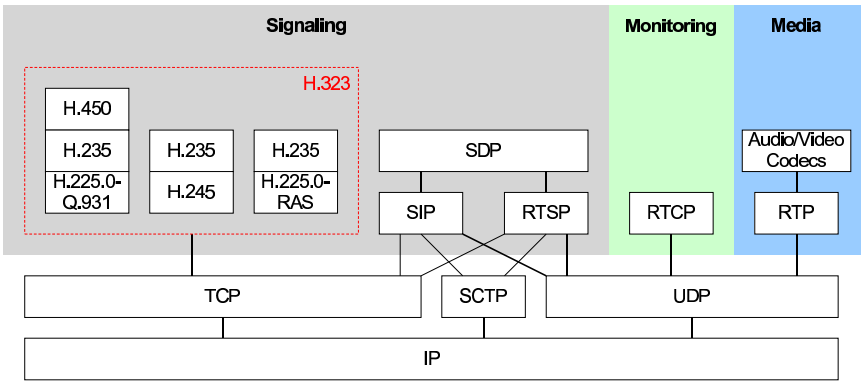

Abbildung 1: Multimedia Protokollstack

Der Signalisierungspfad wird zur Auffindung der Kommunikationspartner, zur Aushandlung der Transportadressen, der Codecs, etc. und zur Steuerung der Multimedia-Sitzung benötigt. Für diesen Zweck stehen verschiedene Standards und Protokolle zur Verfügung. Darunter sind die von der ITU-T standardisierte Protokollfamilie H.323 [I32300] und die von der IETF erarbeiteten und verabschiedeten Standards SIP [Ha99] und RTSP [SRL98].

Über den Transportpfad werden die Mediendaten ausgetauscht. Zur Zeit findet hierfür im Internet ausschließlich das RTP [SCFJ96] Protokoll Verwendung.

2.1 Signalisierung

Die erste und damit älteste Spezifikation für paketbasierte Multimedia-Kommunikation ist die von der ITU-T verabschiedete Norm H.323 [I32300]. Die erste Version wurde bereits 1990 veröffentlicht. H.323 umfasst eine ganze Sammlung von Teilstandards, die alle Aspekte der Sprach- und Video-Übertragung in paketbasierten Netzen enthält. Zum Verbindungsaufbau referenziert H.323 die Norm H.225 [I22500] (Call Signaling). H.245 [I24500] (Control Signaling) beschreibt die Möglichkeiten zur Aushandlung von Parametern wie z.B. Transportadressen und geeigneter Audio/Video-Codecs und die Steuerung

der Sitzung. Mögliche Sprach-Codecs werden der G.700-Reihe (G.711, G.722, G.723, G.728, G.729) und verfügbare Video-Codecs der H.260-Reihe (H.261, H.263) entnommen. H.323 ist der Defacto-Standard für VoIP und Video-Konferenzen im Internet.

Neben der H.323 Protokollfamilie hat die IETF eine Reihe von Standards entwickelt, die die Steuerung und Signalisierung von Multimedia-Sitzungen im Internet ermöglichen. Hierzu zählt das Session Initiation Protocol (SIP) [Ha99], welches u.a. in Zusammenhang mit dem Aufbau von VoIP- oder Video-Konferenz-Sitzungen zum Tragen kommt. Im Bereich der Media-on-Demand-Dienste ist das Real-time Streaming Protocol (RTSP) [SRL98] angesiedelt. Es bietet Funktionen zum Abruf von Medienströmen und Kommandos zur Steuerung des Abspielverhaltens (Start, Stop, Pause, Vorwärts, Rückwarts). Die Beschreibung von Multimedia-Sitzungen erfolgt mit Hilfe des Session Description Protocol (SDP) [HJ98].

2.2 Transport

Der Transport der echtzeit-orientierten Daten wird mit Hilfe des Real-time Transport Protocol (RTP) [SCFJ96] realisiert. RFC 1889 definiert neben RTP ein weiteres Protokoll, das zur Kontrolle der Übertragungsqualität dient und Parameter wie Delay, Jitter und Paketverlustrate berechnet. Es handelt sich hierbei um das Real-time Transport Control Protocol (RTCP).

Zusätzlich zu den genannten Monitoring-Funktionalitäten wird RTCP zum Transport von Teilnehmerinformationen (Name, E-Mail), von anwendungsspezifischen Informationen (APP-Paket) und zur Beendigung von RTP-Sitzungen (BYE-Paket) genutzt. Folglich ist auch dieser Kommunikationskanal vor passiven und aktiven Angriffen zu schützen.

3 Sicherheitsmechanismen

Bei der Auswahl und Integration von kryptographischen Verfahren in Multimedia-Kommunikatiosanwendungen müssen beide Kommunikationspfade geschützt werden. Sowohl für die Signalisierung als auch für den Transport gilt es, geeignete Sicherheitsmechanismen zu implementieren.

Der Signalisierungskanal stellt prinzipiell keine neuen Anforderungen an die Security-Protokolle, die nicht auch schon die konventionellen Internet-Anwendungen wie Email, FTP oder http stellen. Daher ist es empfehlenswert, zur Sicherung des Signalisierungspfades, bekannte und gut untersuchte Security-Protokolle zu verwenden. H.323 und RTSP referenzieren in diesem Zusammenhang zum Schutz der Signalisierungsnachrichten IPSec und SSL/TLS.

Dieser Beitrag beschränkt sich auf die Integration von Sicherheitsdiensten in den Medientransport und speziell auf die Vertraulichkeit der Mediendaten, denn hier findet die realzeit-orientierte Kommunikation statt, die neue Anforderungen an die Sicherheitsprotokolle stellt. Dabei werden verschiedene Internet Sicherheitsprotokolle auf ihre Tauglichkeit im Umfeld isochroner Übertragung untersucht und insbesondere deren Auswirkungen auf QoS-Parameter herausgearbeitet, die durch Datenexpansion, Fehlerfortpflanzung und zusätzliche Verzögerung bedingt sind.

Abgesehen von anwendungsabhängigen Sicherheitsmechanismen gibt es im Internet zwei Standard-Sicherheitsprotokolle: SSL und IPSec. Während Secure Socket Layer (SSL) bzw. Transport Layer Security[2] (TLS) [DA99] TCP-Verbindungen insbesondere zwischen Clients und Servern sichert, wurde IPSec [KA98] für die Sicherheit der IP-Protokollelemente zwischen beliebigen Internetadressen entworfen. Für die Sicherheit von Multimedia Kommunikationsanwendungen kommt nur IPSec in Frage. SSL/TLS kann nicht verwendet werden, da Multimedia Anwendungen aufgrund ihres Echtzeit-Charakters UDP als Protokoll der Transportschicht nutzen.

3.1 IPSec

IPSec realisiert die Sicherheitsdienste Vertraulichkeit, Datenunversehrtheit und Datenauthentikation für IP-Pakete. Ein eingeschränkter Schutz vor Verkehrsflussanalysen ist ebenso enthalten wie ein Schutz vor wiederholter Sendung von IP-Paketen. IPSec wurde sowohl für das Internet-Protokoll Version 4 als auch Version 6 konzipiert. Es arbeitet mit zwei verschiedenen Protokollen: *dem Encapsulating Security Payload (ESP) und dem Authentication Header (AH)*.

Der ESP-Mechanismus von IPSec verschlüsselt und authentifiziert die Daten eines IP-Pakets. Somit bietet ESP Vertraulichkeit und Datenauthentikation. Des Weiteren beinhaltet der ESP-Header eine Sequenznummer, wodurch beim Empfänger eines ESP/IP-Pakets in Verbindung mit einem lokalen Replay-Fenster, in dem die Sequenznummern bereits empfangener Pakete gespeichert sind, wiederholt gesendete IP-Pakete erkannt werden können.

Im Gegensatz zum ESP bietet der AH keine Verschlüsselung der Daten. Er gewährleistet nur die Datenintegrität. Ein optionaler Schutz gegen Replay-Angriffe ist, wie beim ESP, aufgrund der im AH-Header ebenfalls enthaltenen Sequenznummer äquivalent möglich. Der große Unterschied zur Authentikation des ESP ist, dass AH Teile des IP-Headers bei der Generierung der Authentifizierungsdaten (kryptographische Prüfsumme) mit einschließt.

Die Sicherheitsprotokolle können in zwei Modi realisiert werden: *Transport Mode und Tunneling Mode*. Im Transport Mode werden die IP-Pakete als Einheiten mit Sicherheitsinformationen versehen und (optional) verschlüsselt und übertragen. Im Tunneling Mode werden die gesicherten IP Pakete selber in IP-Pakete verpackt, die dann (getunnelt) als Standard-IP-Pakete übertragen werden. Typischerweise verwenden Endgeräte den Transport Mode, während der Tunneling Mode z.B. von Firewalls und Security Gateways genutzt wird. Beide Modi lassen sich aber auch kombinieren oder sogar kaskadieren.

ESP und AH fügen jedem IP-Paket einen eigenen Header hinzu. Dies bedeutet eine nicht zu vernachlässigende Datenexpansion. Gerade beim Sprachkanal fällt diese stark ins Gewicht, da dort typischerweises jedes Sprachpaket nur wenige Bytes an Nutzdaten enthält. Wird der in VoIP- oder Video-Konferenzumfeld gängige Vocoder G.723.1 verwendet, enthält jeder Frame 30 ms Sprache, die in 20 Bytes codiert werden. Daraus folgt, dass ein RTP-Paket, das 60 ms Sprache überträgt, 56 Bytes (16 Bytes RTP-Header) groß ist. Das

[2] Mit TLS standardisier die IETF Netscape's SSL.

IP-Paket hat dann eine Größe von 84 Bytes. Durch die Anwendung der Security-Protokolle ESP und AH wächst die Paketgröße dramatisch an, da sowohl ESP als auch AH ihre Header jedem Paket voranstellen. Extreme Ausmaße nimmt die Datenexpansion im Tunnel-Modus an, da hierbei neben dem ESP-Header ein vollständiger IP-Header hinzugefügt wird. Dabei wächst die Paketgröße um ungefähr 52,4 % auf 128 Bytes an.

Eine häufig bei IPSec-gesicherter Kommunikation verwendete Betriebsart für Blockchiffren ist der Cipherblock Chaining (CBC) Modus [ISO97]. Dieser verschlüsselt jeden Eingabeblock P_i mit der Verschlüsselungsfunktion E und dem geheimen Schlüssel k in dem der Chiffretext des vorherigen Blocks C_{i-1} zuvor mit dem Eingabeblock XOR-verknüpft wird: $C_i = E_k\left(C_{i-1} \oplus P_i\right)$. Folglich wirkt sich ein Fehler in der Übertragung über zwei Ausgabeblöcke aus.

Die zur Erbringung der Sicherheitsdienste nötigen kryptographischen Systemparameter und insbesondere die Schlüssel können manuell und automatisch in sogenannten Security Associations (SA) zwischen den Kommunikationspartnern festgelegt werden. Da das manuelle Einstellen durch einen Security-Administrator nur für eine kleine Anzahl an IT-Systemen handhabbar ist, wird in der Regel die automatische Variante zur Aushandlung der SAs verwendet. Hierfür referenziert IPSec das Internet Key Exchange (IKE) [HC98] Protokoll. Eine SA ist eine unidirektionale Verbindung zwischen zwei Endsystemen. Für Video-Konferenzen, die zwei RTP-Kanäle mit den dazugehörigen RTCP-Verbindungen öffnen, sind folglich auch jedem Endsystem 6 SAs zu etablieren und vorzuhalten.

Konferenzen beziehen oft mehr als nur zwei Kommunikationspartner ein. Die Verwendung von IPSec ist in diesen Mehrteilnehmer-Szenarien nicht möglich, da IKE die Etablierung von SAs nur zwischen zwei Kommunikationsteilnehmern unterstützt.

Die betrachteten Sicherheitsdienste können Ende-zu-Ende realisiert werden (meist im Transport Modus) oder eine Teilstrecke sichern, die über offene Netze geht (Tunnel Modus). Die Ende-zu-Ende Sicherheit ist erstrebenswert, verursacht allerdings im IPSec-Umfeld Probleme im Zusammenspiel mit Paketfiltern. Diese sind nicht in der Lage, im verschlüsselten Payload des IP-Pakets die in diesem Fall für den Medienstrom verwendeten UDP-Ports auszulesen.

3.2 H.235

Im Anhang J der H.323-Norm werden Sicherheitsmerkmale der paketbasierten Multimedia Kommunikation beschrieben (Baseline Security Profile, Sophisticated Security Profile, Media Encryption). Diese Profile enthalten aber nur Anforderungen an Sicherheitsdienste und keine Sicherheitsmechanismen oder technologieabhängigen Vorgaben. Diese werden im Standard H.235 beschrieben.

H.235 kann als anwendungsspezifische Sicherheitsnorm gesehen werden. Sie definiert für die H.323 Protokollfamilie Sicherheitsmechanismen, die sich sowohl auf die Signalisierungsnachrichten als auch auf den Medienstrom beziehen. Dabei legt die H.235 Norm erst im Anhang Sicherheitsprofile fest, die spezifizieren, was eine H.235-Implementierung mindestens unterstützen muss (siehe Abbildung 2). Das Baseline Security Profile sieht einen Integritätsschutz der Signalisierungsnachrichten (RAS, H.225.0 und H.245) mittels

Baseline security profile (Annex D)
Voice encryption profile (Annex D, optional)
Signature security profile (Annex E)

Security Services	Call Functions					
	RAS	H.225.0	H.245	RTP		
Authentication	Password HMAC-SHA1-96	Password HMAC-SHA1-96	Password HMAC-SHA1-96			
	Digital signature MD5/SHA1-RSA	Digital signature MD5/SHA1-RSA	Digital signature MD5/SHA1-RSA			
Non-Repudiation	Digital signature MD5/SHA1-RSA	Digital signature MD5/SHA1-RSA	Digital signature MD5/SHA1-RSA			
Integrity	Password HMAC-SHA1-96	Password HMAC-SHA1-96	Password HMAC-SHA1-96			
	Digital signature MD5/SHA1-RSA	Digital signature MD5/SHA1-RSA	Digital signature MD5/SHA1-RSA			
Confidentiality				RC2/CBC 56 bit	DES/CBC 56 bit	3DES/CBC 168 bit
Access Control						
Key Management	Subscription-based password assignment	Subscription-based password assignment	Integrated H.235 session key management (key distribution, key update)			
		Authenticated DH key exchange				
	Certificates	Certificates				

Abbildung 2: Sicherheitsdienste in H.235

MACs vor und eine Authentikation mit Verwendung von Passwörtern. Die Vertraulichkeit des Medienstroms ist optional und beschränkt sich auf RTP (entspricht RFC 1889). Mechanismen zum Schutz der Integrität der RTP-Pakete sind nicht enthalten. RTCP bleibt gänzlich ungeschützt. Das Schlüsselmanagement beschränkt sich in dem genannten Profil auf das a priori Vereinbaren der Schlüssel.

Die Datenexpansion verursacht durch die in H.235 definierten Sicherheitsmechanismen ist relativ beschränkt. Da der Sicherheitsdienst Nachrichtenauthentizität für RTP-Pakete nicht geboten wird, wird folglich auch keine kryptographische Prüfsumme dem Paket beigefügt. Das Paket vergrößert sich nur anhand der evtl. nötigen Auffüllung des letzten Blocks mit Padding-Bits, um ein Vielfaches der Blockgröße des Verschlüsselungsalgorithmus' zu erhalten.

Die einzige Betriebsart für Blockchiffren, die im H.235 Standard spezifiziert ist, ist der CBC Modus. Dieser hat neben der Eigenschaft, dass der Klartext ein Vielfaches der Blocklänge der zugrundeliegenden Blockchiffre sein muss, durch die Verkettung von Blöcken eine Fehlerfortpflanzung, die sich über den nachfolgenden Block erstreckt (vgl. Abschnitt 3.1). Im Falle von H.235 werden so gestörte RTP-Pakete dennoch decodiert und abgespielt, da der Sicherheitsdienst Nachrichtenauthentizität für RTP-Pakete nicht im Standard enthalten ist.

Multi-Point Control Units (MCUs) ermöglichen Multimedia-Konferenzen zwischen drei oder mehr Terminals und werden häufig als Konferenz-Server bezeichnen. Die Signalisierung wird über eine zentrale Komponente, dem sog. Multi-Point Controller (MC), abgewickelt. Jeder Endpunkt der Konferenz (Slave) baut eine TCP-Verbindung zum MC

(Master) auf. Über diesen Kanal kontrolliert der MC durch das Versenden entsprechender Nachrichten die Endpunkte der Konferenz und kann insbesondere dafür Sorge tragen, dass alle Slaves mit den gleichen kryptographischen Systemparametern konfiguriert werden. Die Medienverteilung dagegen kann auf einer zentralisierten Master/Slave-Konfiguration oder einer dezentralisierten Multicast-Konfiguration basieren. Im ersten Ansatz ist der sog. Multi-Point Processor (MP) für die Verarbeitung der eingehenden Medienströme verantwortlich, die jeder Endpunkt an den MP sendet. In der zweiten Konfiguration wird der MP nicht benötig und die Terminals senden und empfangen alle Medien über die spezifizierte Multicast-Adresse.

Die Ende-zu-Ende Vertraulichkeit in H.235 schränkt die Funktionsweise von Paketfiltern nicht ein. Anders als bei IPSec wird nur der Payload von RTP-Paketen verschlüsselt, wodurch die Felder im Transport-Header für Firewalls auslesbar bleiben.

3.3 SRTP

Wie bereits erläutert und in Abbildung 1 grafisch veranschaulicht wurde, ist RTP/UDP derzeit die einzige Kombination, die von allen Multimedia-Kommunikationssystemen zum Transport der Medien-Daten benutzt wird. Durch den Ansatz Sicherheitsmechanismen in RTP zu integrieren, können diese dem Transportpfad jeder multimedialen Kommunikationsanwendung im Internet zu Gute kommen. Sicherheitsdienste auf RTP-Ebene können sowohl in H.323- als auch in SIP- und RTSP-basierten Multimedia Applikationen integriert werden. Genau dieser Ansatz wird z.Zt. von einer Arbeitsgruppe der IETF verfolgt, die sich allgemein mit dem Thema Audio-/Videotransport beschäftigt (AVT Working Group). Zu den Veröffentlichungen dieser Arbeitsgruppe gehört unter anderem ein Internet Draft mit dem Titel "The Secure Real-Time Transport Protocol" (SRTP) [Ba02], in dem Prozeduren und Mechanismen zur sicheren Übertragung von RTP- und RTCP-Paketen beschrieben werden.

SRTP ist als Profil für RTP definiert und bietet Sicherheitsdienste sowohl für RTP als auch RTCP an. Darunter sind die Vertraulichkeit und die Datenauthentikation der RTP- und RTCP-Pakete und ein Schutz vor Replay-Attacken. Im Folgenden gehen wir nur kurz auf die wesentlichen Aspekte und Eigenschaften von SRTP ein. Für ausführlichere Beschreibungen verweisen wir auf [Ba02], [LR02].

Die Verschlüsselung wird in SRTP mit einer Stromchiffre realisiert. Ein Schlüsselstromgenerator erzeugt den Schlüsselstrom, der anschließend mit den Klartext XOR-verknüpft wird (siehe Abbildung 3).

Verschlüsselt werden die Nutzdaten des RTP-Pakets. Diese Art der Chiffrierung hat kein zusätzliches Padding zur Folge, da jedes Klartextbit einzelnd verarbeitet wird (nicht im Block). Durch die Verschlüsselung wird folglich keine zusätzliche Datenexpansion verursacht und Fehler wirken sich auch nur auf das betroffene Bit aus. Als vorgegebener Schlüsselstromgenerator *KG* wird der AES [N01] im Segmented Integer Counter (SIC) Modus [DH79], [Mc00] verwendet. Alternativ dazu spezifiziert [Ba02] den AES im F8 Modus [3G01]. Beide Kryptosysteme haben den Vorteil, dass die Generierung des Schlüsselstrom unabhängig von dem Klartext erfolgen kann. Folglich kann die rechenintensive

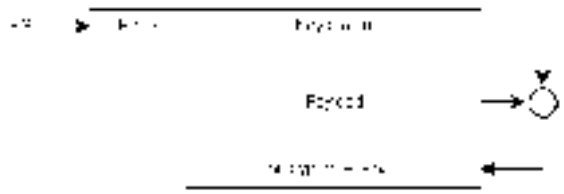

Abbildung 3: Verschlüsselung in SRTP

Erzeugung des Schlüsselstroms im Vorfeld erfolgen – unter besonderer Nutzung von Freizyklen des Prozessors (z.B. Sprachpausen) –, wobei die eigentliche Verschlüsselung nur noch aus der effizienten XOR-Verknüpfung von Klartext und Schlüsselstrom besteht. Die Verzögerung durch die Verschlüsselungsoperation kann somit minimiert werden.

Die Datenauthentikation wird mit Hilfe von Message Authentication Codes (MACs) realisiert. Hierfür ist der HMAC/SHA1 Algorithmus spezifiziert [KBC97]. Da die zur MAC-Berechnung verwendete kryptographische Hashfunktion einen Hashwert der Länge 160 Bit erzeugt, ist auch die Länge der MACs 160 Bit. Um die Datenexpansion zu reduzieren, sieht SRTP die Verkürzung auf die 32 höchstwertigen Bits vor. Dies ist allerdings nicht ganz unproblematisch, da diese Maßnahme Kollisionsattacken ermöglicht (Geburtstagsparadoxon) [KBC97].

SRTP definiert entgegen der beiden bisher diskutierten Sicherheitsprotokolle keine Schlüsselmanagement-Funktionalitäten bzw. -Protokolle. Es ist vielmehr möglich abhängig von der zugrundeliegenden Anwendung ein geeignetes Key-Management zu wählen bzw. zu entwickeln.

Die Unterstützung von Gruppenkommunikation ist z.B. möglich, da SRTP für Multicast-Konfigurationen geeignet ist. Hierfür ist – wie gesagt – ein passendes Schlüsselmanagement Voraussetzung.

Die Koexistenz der Ende-zu-Ende Verschlüsselung mit SRTP und Packetfiltern gilt das gleiche wie für H.235. In SRTP werden die RTP-Nutzdaten verschlüsselt. Die UDP-Header bleiben für Firewalls auslesbar und auswertbar.

3.4 Zusammenfassender Vergleich der Sicherheitsmechanismen

IPSec kann prinzipiell zur Durchführung vertraulicher Video-Konferenzen im Internet genutzt werden. Die starke Verbreitung gerade als VPN-Technik bietet dies sogar an, da die Abwicklung der multimedialen Kommunikation über bereits vorhandene Infrastrukturen erfolgen kann. Allerdings wurde ersichtlich, dass IPSec in vielerlei Hinsicht keine optimale Lösung darstellt. Einerseits expandieren die ESP- und AH-Header den Datenstrom in nicht zu vernachlässigender Art und Weise, so dass es zu steigenden Übertragungskosten kommt und die bereitgestellte Übertragungsgeschwindigkeit entsprechend zu erhöhen ist. Andererseits verursacht das zur Aushandlung der SAs verwendete IKE Protokoll einen enormen Aufwand und ist der Grund für die fehlende Multiteilnehmer-Unterstützung. Nicht zu Letzt ist das komplexe IKE häufig der Grund für mangelnde Interoperabilität.

Schließlich erschwert der Einsatz von IPSec zur Ende-zu-Ende Verschlüsselung zudem das Zusammenspiel mit Paketfiltern.

Sicherheitsprotokolle auf Anwendungsschicht beeinträchtigen die Funktionsweise von Paketfiltern nicht. H.235 und SRTP gehören zu denen, die im Multimedia-Umfeld zur Verfügung stehen. H.235 hat allerdings zu viele Lücken. Gerade der außer Acht gelassene RTCP-Kontrollkanal bietet einige Angriffsmöglichkeiten (z.B. Abhören von Teilnehmerinformationen und Denial of Service).

SRTP schließt diese Lücken und bietet effiziente und moderne kryptographische Verfahren zur Erbringung der Sicherheitsdienste Vertraulichkeit und Datenauthentizität. Ein spezifisches Protokoll zum Aushandeln, Austausch und Verwalten kryptographischer Parameter und Schlüssel ist nicht explizit in SRTP enthalten. Vielmehr können geeignete Schlüsselmanagement-Protokolle in Abhängigkeit gegebener Applikationen bzw. gegebener Anforderungen dediziert ausgewählt werden.

4 Implementierung

Zur Implementierung des SRTP-Protokolls muss zunächst eine Software-Bibliothek mit kryptographischen Verfahren bereitstehen, die die in Tabelle 1 aufgelisteten Algorithmen enthält.

	Mandatory	**Optional**	**Default**
Encryption	AES/SIC, NULL	AES/F8	AES/SIC
Integrity	HMAC/SHA1	-	HMAC/SHA1
Key Derivation	AES/SIC	-	AES/SIC

Tabelle 1: Kryptographische Verfahren in SRTP

Stehen die kryptographischen Primitiven zur Verfügung, kann SRTP durch die Erweiterung eines RTP-Frameworks realisiert werden. Es empfiehlt sich eine sog. Bump-in-the-Stack Implementierung, die zwischen der RTP-Schicht und der Transportschicht liegt. SRTP bekommt in dieser Anordnung die RTP-Pakete von der übergeordneten Schicht, verarbeitet diese entsprechend des vereinbarten Krypto-Kontextes und leitet diese an die Transportschicht weiter. Auf der Empfängerseite wird analog verfahren. Hier bekommt die SRTP-Schicht das SRTP-Paket von der Transportschicht übergeben. Diese kann es entsprechend Verarbeiten und als RTP-Paket den Protokollstack nach oben reichen.

Die meisten RTP-Implementierungen sind in Verbindung mit einer Multimedia Kommunikationsanwendung anzutreffen. Dies gilt auch für die im openH323-Projekt[3] entwickelte Open Source Implementierung des H.323 Standards. Hier ist neben den Signalisierungsstandards, Vocodern und Video-Codecs ein RTP-Stack enthalten.

[3] http://www.openh323.org/

Das objektorientierte in C++ entwickelte Framework und speziell die darin enthaltenen Klassen `RTP_Session` und `RTP_UDP` erlauben durch entsprechende Ableitung und Erweiterungen die Integration von SRTP (siehe Abbildung 4).

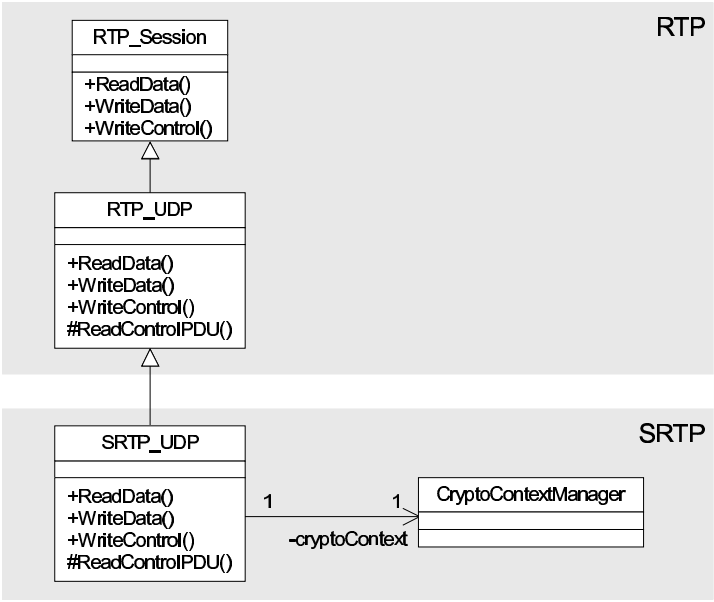

Abbildung 4: Integration von SRTP in ein RTP Framework

RTP-Pakete werden über den Aufruf der Funktionen `ReadData()` und `WriteData()` empfangen bzw. gesendet. Das Senden und Empfangen von RTCP-Paketen ist über die Funktionen `WriteControl()` und `ReadControlPDU()` möglich. Die Klasse `SRTP_UDP` definiert diese Funktionen neu und überschreibt damit die ursprüngliche Implementierung, so dass die Pakete abhängig von den gewählten Sicherheitsdiensten (festgehalten im CryptoContext) verarbeitet werden.

OpenPhone ist ein grafisches H.323-Terminal für Windows, das im Rahmen des OpenH323-Projekts entwickelt wurde und den OpenH323-Stack als Basis für die Kommunikation nutzt. Mit dieser Applikation ist es möglich, eine Videokonferenz zwischen zwei Kommunikationspartnern aufzubauen. Durch einige wenige Eingriffe kann OpenPhone so modifiziert werden (hin zum Secure OpenPhone), dass ein sicherer Medientransport über das in OpenH323 integrierte SRTP-Framework möglich wird. Für den Benutzer werden diese Änderungen durch ein erweitertes Optionen-Menü sichtbar (siehe Abbildung 5).

Durch die Optionen erhält der Benutzer die Möglichkeit, diverse für die Initialisierung des Krypto-Kontexts und somit der SRTP-Session benötigten Parameter einzustellen. Anzugeben sind – gruppiert nach Sicherheitsdiensten – die Algorithmen für jedes Protokoll und schließlich die kryptographischen Schlüssel und zugehörige Systemparameter. Bei akti-

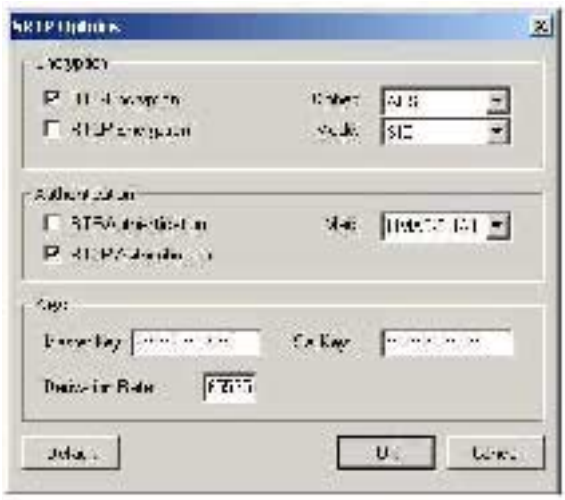

Abbildung 5: SRTP Optionen des Secure OpenPhone

vierter Authentikation ist der Schutz vor Wiederholungsangriffen implizit eingeschlossen, da diese nur in Kombination Sinn machen. Bei wiederholten Paketen oder fehlgeschlagener MAC-Verifikation öffnet sich ein kleines Pop-Up-Fenster, welches dem Benutzer eine Statistik über die Anzahl der verworfenen Pakete anzeigt.

Analog kann die Erweiterung anderer RTP-Implementierung erfolgen, wie z.B. die des Java Media Frameworks (JMF)[4]. Applikationen auf Basis der Signalisierungsstandards RTSP (im JMF enthalten) und SIP[5] können hiermit implementiert werden.

5 Tests und Messergebnisse

Bei der Verwendung von Secure OpenPhone lässt sich subjektiv kein Unterschied zwischen den Betrieb ohne Security und dem mit Security feststellen. Dies unterstreichen ebenfalls die auf Basis der Implementierung durchgeführten Messungen. Abbildung 6 zeigt exemplarisch die Zeiten, die die verschiedenen Verschlüsselungsalgorithmen zur Verarbeitung verschiedener RTP-Pakete benötigen. Jeder Balken zeigt dabei die Summe aus der Initialisierungszeit und der Verschlüsselungszeit. Die Messungen wurden auf einem PC mit 350 MHz Intel Pentium III Prozessor und 128 MB Hauptspeicher ausgeführt.

Die Default-Einstellung zur Verschlüsselung gemäß [Ba02] ist der AES im SIC Modus. Die Messergebnisse sind ganz links im Diagramm aus Abbildung 6 eingezeichnet. Wie daraus abgelesen werden kann, ist die Verzögerung durch die Verschlüsselungsoperation für ein RTP-Paket, das vier GSM-Frames enthält, deutlich unter 0,2 ms. Hieraus folgt, dass die Implementierung von SRTP so effizient erfolgen kann, dass zur Erbringung der

[4] http://java.sun.com/jmf/

[5] http://dns.antd.nist.gov/proj/iptel/

in Abschnitt 3.3 genannten Sicherheitsdienste keine zusätzliche QoS bereitgestellt werden muss.

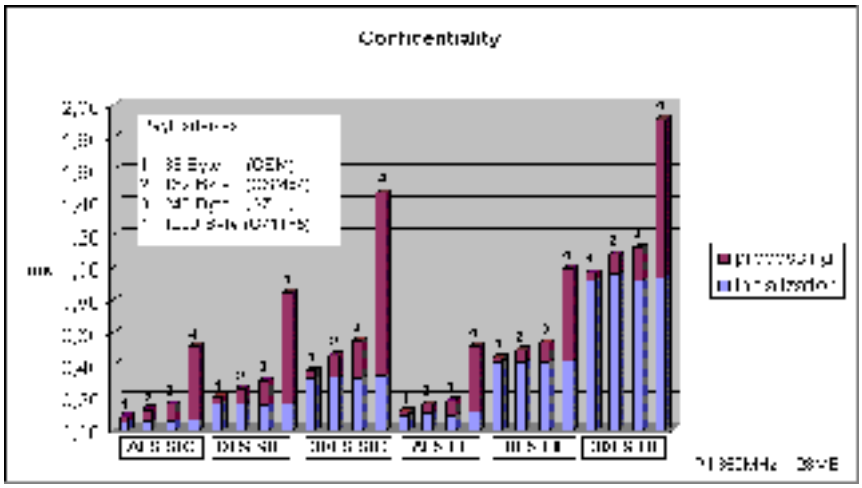

Abbildung 6: Messergebnisse

6 Fazit

Video-Konferenzen sind vor passiven und aktiven Angriffen zu schützen. Dies gilt bereits für Konferenzsysteme, die über geschlossene Übertragungsnetze wie das ISDN kommunizieren [Sch03]. Ein höheres Gefährdungspotenzial ist im Internet gegeben. Passive Angriffe sind hier mit weitaus geringerem Aufwand realisierbar [Lo02]. Dazu kommen eine Reihe aktiver Angriffe, die durch Manipulationen der Signalisierungs- oder Medien-Daten erfolgen. Bleiben Veränderungen am Medienstrom unerkannt, ist z.B. ein Diensteanbieter in der Lage, bei bestehenden Verbindungen den verwendeten Codec zu ändern, ohne dass die Verbindungsteilnehmer darüber Kenntnis bekommen. Liegt in diesem Szenario eine volumenabhängige Tarifierung vor, ist die Motivation, einen Codec mit höherem Übertragungsvolumen einzustellen, offensichtlich.

Neben der Security spielt die Netzwerk-QoS für echtzeit-orientierte Kommunikationsanwendungen eine wesentliche Rolle. Security-Mechanismen beeinflussen die QoS-Anforderungen. Wie sich die Vertraulichkeit bei Video-Konferenzsystemen auf QoS-Parameter auswirkt wurde in diesem Beitrag untersucht. Verfügbare Standards zum Schutz multimedialer Kommunikation wurden analysiert und diskutiert. Einige sind im Umfeld echtzeit-orientierter Kommunikation einsetzbar, andere nicht. SRTP stellt sich als geeignetes Protokoll zum Schutz der Medien-Daten heraus. Die vorgestellte SRTP-Implementierung ermöglicht die Durchführung von Video-Konferenzen im Internet mittels handelsüblicher PCs ohne zusätzlicher kryptographischer Hardware oder Rechenbeschleuniger. Die Verarbeitungszeiten der Verfahren zur Erbringung der Sicherheitsdienste sind dabei dermaßen

effizient, dass keine zusätzlichen Anforderungen an die zur Verfügung stehende Dienstgüte gestellt werden müssen.

Das Interesse an SRTP nimmt auch auf Seiten der Hersteller zu. Cisco arbeitet z.B. an SRTP-Implementierungen als zukünftige Erweiterung der eigenen VoIP-Produktpalette.

Durch den Ansatz, Sicherheitsdienste in RTP zu integrieren, kann jede Internet Multimedia Kommunikationsanwendung davon profitieren, da RTP zur Zeit das einzige Protokoll ist, das Dienste für realzeit-orientierte Anwendungen anbietet. Anwendungen, die auf den Signalisierungsnormen SIP und RTSP basieren, können die Sicherheitsdienste von SRTP nutzen und die benötigten Parameter aushandeln. Hierfür wurde das in beiden Signalisierungsstandards verwendete SDP entsprechend erweitert. Für H.323 ist dies allerdings noch offen. Theoretische Ansätze hierzu finden sich in [LR02]. Ob und wie sich SRTP in H.323-Umgebungen wie z.B. dem DFN Dienst *DFNVC* integrieren und nutzen lässt, wirft herausfordernde Forschungs- und Entwicklungsfragen auf.

Literatur

[3G01] 3rd Generation Partnership Project (3GPP), Technical Specification Group Services and System Aspects: Specification of the 3GPP Confidentiality and Integrity Algorithms – Technical Specification – Document 1: f8 and f9 Specification, Dezember 2001.

[Ba02] Baugher, M., McGrew, D., Oran, D., Blom, R., Carrara, E., Naslund, M., Norrman, K.: The Secure Real-time Transport Protocol. IETF Internet Draft, Juni 2002.

[DA99] Dierks, T., Allen, C.: The TLS Protocol Version 1.0. IETF RFC 2246, Januar 1999.

[DH79] Diffie, W., Hellman, M.: Privacy and Authentication: An Introduction to Cryptography. In Proceedings of the IEEE, 67(3), 1979; S. 397-4n27.

[Ha99] Handley, M., Schulzrinne, H., Schooler, E., Rosenberg, J.: SIP: Session Initiation Protocol. IETF RFC 2543, März 1999.

[HC98] Harkins, D., Carrel, D.: The Internet Key Exchange (IKE). IETF RFC 2409, November 1998.

[HJ98] Handley, M., Jacobson, V.: SDP: Session Description Protocol. IETF RFC 2327, April 1998

[I22500] ITU-T Recommendation H.225.0 Version 4: Call Signalling Protocols and Media Stream Packetization for Packet Based Multimedia Communications Systems. ITU-T, November 2000.

[I23500] ITU-T Recommendation H.235 Version 2: Security and encryption for H-Series (H.323 and other H.245-based) mulitmedia terminals. ITU-T, November 2000.

[I24500] ITU-T Recommendation H.245 Version 7: Control Protocol for Multimedia Communication. ITU-T, Februar 2000.

[I32300] ITU-T Recommendation H.323 Version 4: Packet Based Multimedia Communication Systems. ITU-T, November 2000.

[ISO97] ISO/IEC 10116: Information technology – Security techniques – Modes of operation for an n-bit blockcipher. International Organization for Standardization, 1997.

[KA98] Kent, S., Atkinson, R.: Security Architecture for the Internet Protocol. IETF RFC 2401, November 1998.

[KBC97] Krawczyk, H., Bellare, M., Canetti, R.: HMAC: Keyed-Hashing for Message Authentication. IETF RFC 2104, Februar 1997.

[Lo02] Lo Iacono, L.: Rote Telefone: Abhören von IP-Telefonaten. iX, Magazin für Professionelle Informationstechnik, Heise Verlag, Mai 2002.

[LR02] Lo Iacono, L., Ruland, C.: Confidential Multimedia Communication in IP Networks. In Proceedings of 8th IEEE International Conference on Communication Systems, Singapur, 2002.

[Mc00] McGrew, D.: Segmented Integer Counter Mode: Specification and Rationale. Cisco Systems, Inc., Oktober 2000.

[N01] NIST: Advanced Encryption Standard (AES). FIPS PUB 197, November 2001.

[SCFJ96] Schulzrinne, H., Casner, S., Frederik, R., Jacobson, V.: RTP: A Transport Protocol for Real-Time Applications. IETF RFC 1889, Januar 1996.

[Sch03] Scheidemann, V.: Audio-Angriffe auf Videokonferenzen. In Tagungsband 8. Deutscher IT-Sicherheitskongress des BSI, SecuMedia Verlags-GmbH, 2003.

[Sch96] Schulzrinne, H.: RTP Profile for Audio and Video Conferences with Minimal Control. IETF RFC 1890, January 1996.

[SRL98] Schulzrinne, H., Rao, A., Lanphier, R.: Real Time Streaming Protocol (RTSP). IETF RFC 2326, April 1998.