BAB II
TINJAUAN PUSTAKA

2.1. Steganografi dan Perkembangannya

Kata steganografi menjadi sangat populer setelah kasus pemboman gedung WTC, 11 September 2001 di Amerika Serikat, teroris menyembunyikan pesan-pesan kegiatan terornya dalam berbagai media yang dapat dijadikan penampung untuk menyembunyikan *file* seperti pada *image*, *audio* dan *video* [2]. Pada peristiwa tersebut disebutkan oleh "pejabat pemerintah dan para ahli dari pemerintahan AS" yang tidak disebut namanya bahwa "para teroris menyembunyikan peta-peta dan foto-foto target dan juga perintah untuk aktivitas teroris di ruang *chat sport*, *bulletin boards* porno dan *web site* lainnya".

Steganografi berasal dari bahasa Yunani yaitu *steganós* yang berarti tersembunyi/menyembunyikan dan *gráphy* yang artinya tulisan, sehingga secara keseluruhan artinya adalah tulisan yang disembunyikan [1]. Steganografi telah digunakan sejak sekitar 2500 tahun yang lalu untuk kepentingan politik, militer, diplomatik, serta untuk kepentingan pribadi sebagai alat. Catatan pertama tentang steganografi ditulis oleh seorang sejarawan Yunani [1], Herodotus, yaitu ketika Histiaeus dipenjarakan oleh Raja Darius di Susa pada abad 5 SM. Histiaeus harus mengirim pesan rahasia kepada Aristagoras, dengan cara mentato pesan pada kulit kepala seorang budak dan ketika rambut budak itu mulai tumbuh, Histiaeus mengutus budak itu ke Militus untuk mengirim pesan di kulit kepalanya tersebut kepada Aristagoras. Cara tersebut dilakukan untuk menghindari isi pesan diketahui oleh pihak yang tidak diinginkan di tengah perjalanannya.

Definisi informal pertama tentang skema steganografi diformulasikan oleh Simmons dalam ”The Prisoners’ Problem and The Subliminal Channel” [7]. Dua narapidana, Alice dan Bob berada dalam ruangan sel yang berbeda di bawah pengawasan sipir penjaga, Eve. Mereka saling berkomunikasi berencana untuk melarikan diri, namun sipir penjaga hanya akan mengijinkan komunikasi mereka jika semua komunikasi melalui sipir penjaga. Apabila pesan yang disampaikan Alice kepada Bob tidak mencurigakan, maka pesan akan diteruskan. Dengan kondisi demikian Alice berusaha menyamarkan berita yang disampaikan kepada Bob sebagaimana pesan

biasa yang tidak mencurigakan, namun sebenarnya di dalamnya mengandung pesan rahasia yang hanya diketahui oleh mereka berdua.

Dalam perkembangannya dengan menggunakan media elektronik dan jaringan komunikasi, pesan bisa dipertukarkan dengan cepat dan mudah tanpa dibatasi jarak dan waktu. Namun kemudahan pertukaran pesan melalui media elektronik mempunyai beberapa resiko, di antaranya resiko penyadapan, pengubahan, dan perusakan pesan, sehingga diperlukan suatu cara yang bisa mengurangi dampak negatif atas terjadinya resiko tersebut.

Karena alasan tersebut, muncullah penyandian terhadap pesan dengan enkripsi dan dekripsi. Enkripsi (*enciphering,* standar nama menurut *ISO 7498-2)* dilakukan pada pesan yang akan dikirim dengan cara mengubah pesan asli ke dalam bentuk lain yang sulit untuk dimengerti. Sedangkan dekripsi (*deciphering,* standar nama menurut *ISO 7498-2)* dilakukan pada pesan hasil enkripsi yang diterima dengan cara mengubahnya kembali ke bentuk aslinya dengan kunci yang memang telah diketahui sebelumnya.

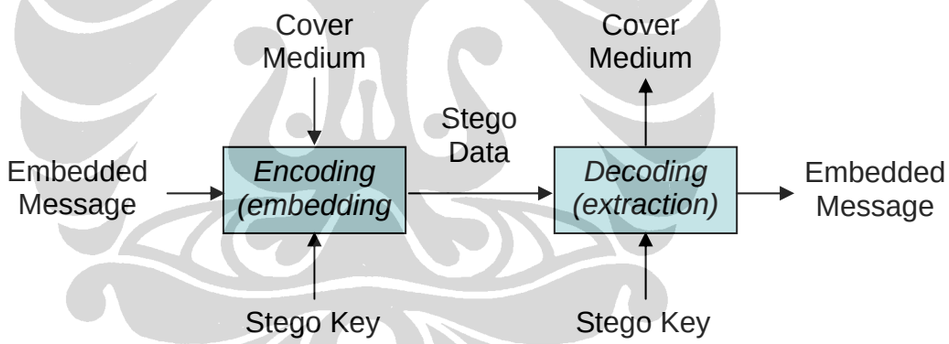
Dengan teknik tersebut, pesan yang dikirimkan selalu dalam bentuk yang sulit dimengerti sehingga yang mengetahui maksudnya hanyalah pengirimnya dan penerima yang dituju. Ilmu dan seni untuk menjaga kerahasian pesan dengan cara menyandikannya ke dalam bentuk yang tidak dapat dimengerti lagi maknanya disebut kriptografi [8]. Semula penyandian sederhana telah efektif untuk menjaga kerahasiaan pesan, tetapi lalu muncul kriptanalisis. Kriptanalisis [8] adalah ilmu dan seni untuk memecahkan pesan terenkripsi menjadi bentuk aslinya tanpa mengetahui kunci yang diberikan. Dampak penggunaan teknik kriptografi yang cukup mencolok adalah suatu informasi yang tidak terbaca/tidak dipahami, tidak lazim sebagaimana mestinya. Hal tersebut akan memberikan suatu simpulan bagi pihak yang tidak berwenang untuk melakukan proses kriptanalisis. Karena telah dikenali oleh penyadap, resiko diketahui, diubah, dan dirusaknya pesan asli menjadi lebih besar.

Menyadari hal tersebut, para ahli kriptografi mencoba mencari cara baru yang lebih efektif dengan melakukan penyembunyian pesan ke dalam pesan lain sehingga pesan tersebut tidak disadari keberadaannya. Cara ini disebut steganografi yaitu suatu seni atau ilmu yang mengkomunikasikan sebuah pesan dengan suatu cara, tanpa bisa dideteksi oleh pihak lawan [1].

Berbeda dengan kriptografi yang menjaga kerahasian pesan dengan cara mengubah bentuk pesan agar tidak dapat dipahami oleh orang lain, steganografi merupakan suatu teknik penyembunyian pesan pada suatu medium. Perlu diperhatikan dalam steganografi, suatu pesan tidak harus diubah, tetapi pesan tersebut disembunyikan pada suatu medium agar pesan tersebut tidak terlihat [9]. Salah satu keuntungan steganografi dibandingkan dengan kriptografi adalah bahwa pesan yang dikirim tidak menarik perhatian sehingga media penampung pesan tidak menimbulkan kecurigaan bagi pihak ketiga [10]. Steganografi banyak digunakan untuk memberikan tanda (seperti *copyright*) pada suatu karya cipta yang menandakan bahwa karya cipta tersebut bukan bajakan. Selain itu steganografi juga seringkali digunakan untuk menyembunyikan data/pesan rahasia yang ditujukan kepada orang tertentu.

Meskipun mempunyai maksud yang sama untuk mengamankan informasi namun terdapat perbedaan mendasar antara kriptografi dan steganografi. Kriptografi menyembunyikan/mengolah isi (*content*) pesan sehingga pesan tidak dapat terbaca. Sedangkan steganografi menyembunyikan keberadaan (*existence*) pesan sehingga tidak menimbulkan kecurigaan (*conspicuous*) akan adanya pesan [1].

Steganografi memerlukan media untuk menyembunyikan pesannya, bisa berupa teks, gambar, audio maupun video.



Gambar 2.1. Skema Metode Steganografi [9]

Pesan yang akan disembunyikan sering diistilahkan sebagai *embedded message* (*hiddentext*), datanya bisa berupa file, image, teks, audio, video, dan lain-lain. Data yang dijadikan media untuk menyembunyikan pesan disebut *cover medium/object (covertext). Cover medium* yang telah ditambahkan pesan rahasia dengan

steganografi disebut *stego-data/object (stegotext)*. Sedangkan kunci yang digunakan pada proses penyembunyian maupun rekonstruksinya disebut *stego-key* [9]. Pada keadaan yang ideal, siapapun yang melakukan *scan* terhadap data tersebut tidak akan mengetahui bahwa data tersebut mengandung data lain yang rahasia sehingga pengambilan data hanya dapat dilakukan oleh penerima yang berhak. Secara skematis, metode steganografi dapat dilihat pada Gambar 2.1.

Blok utama yang membangun sebuah algoritma steganografi [9] adalah pemilihan *cover medium,* algoritma *embedding* dan *extracting,* serta manajemen *stego key*.

Secara matematis skema steganografi sudah dapat ditentukan. Anggap K_s adalah *stego key* dari susunan $\mathcal{K}$, semua kunci rahasia stego. $\mathcal{M}$ adalah susunan semua pesan yang dapat ditempelkan, dan $\mathcal{C}$ adalah susunan semua cover medium. Skema steganografik dibentuk oleh dua pemetaan, pemetaan penempelan, *Emb*, dan pemetaan ekstraksi, *Ext* :

$$Emb : \mathcal{C} \times \mathcal{K} \times \mathcal{M} \rightarrow \mathcal{C}$$

$$Ext : \mathcal{C} \rightarrow \mathcal{M}_1 \qquad (2.1)$$

Sehingga $Ext(Emb(\mathbf{c}, K_s, \mathbf{m})) = \mathbf{m}$ untuk semua $\mathbf{c} \in \mathcal{C}$, $K_s \in \mathcal{K}$, and $\mathbf{m} \in \mathcal{M}$.

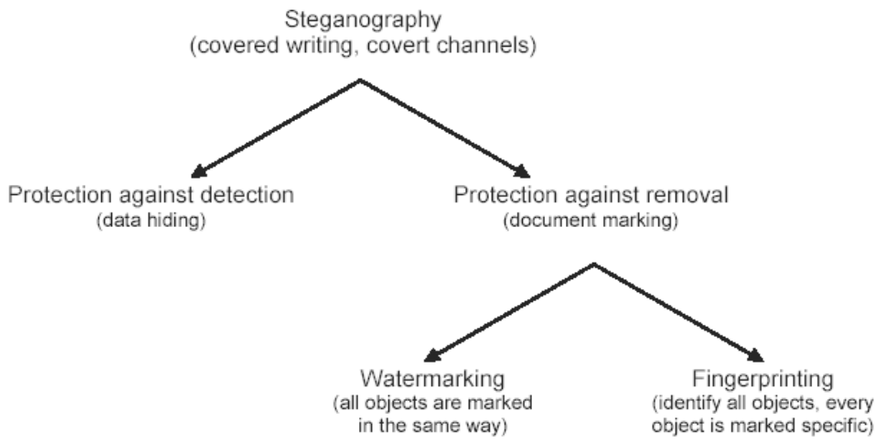
$\mathbf{s} = Emb(\mathbf{c}, K_s, \mathbf{m})$ disebut *stego work*.

Banyaknya teknik dalam steganografi menyebabkan diperlukan adanya pengelompokan atas jenis-jenisnya. Pengelompokan ini diharapkan akan memudahkan pengguna steganografi untuk memilih teknik yang sesuai dan pembuat steganografi untuk mengembangkan teknik-teknik baru dengan lebih terarah.

Steganografi sendiri sering dibedakan menjadi dua jenis menurut tujuannya [11], yaitu steganografi untuk menghindari deteksi (*data hiding*) dan steganografi untuk menghindari penghapusan data (*document marking*). Jenis pertama kemudian lebih sering disebut sebagai steganografi itu sendiri. Jenis kedua dibagi lagi menjadi dua yaitu *watermarking* dan *fingerprinting,* seperti terlihat pada Gambar 2.2. *Watermarking* lebih banyak dibahas bersama dengan steganografi.

Manfaat utama dari *watermarking* adalah untuk identifikasi dan menyertakan potongan informasi yang unik pada suatu media tanpa disadari orang lain. Kedua hal tersebut umumnya ditujukan untuk menandakan keaslian dari suatu karya yang pada akhirnya bisa meminimalkan tindak pembajakan atas karya tersebut.

Perbedaan antara *data hiding* dan *watermarking* adalah kehadiran pihak lawan yang aktif.



Gambar 2.2. Skema Representasi Prosedur Steganografi [11]

Dalam aplikasi *watermarking* seperti *copyright protection* dan *authentication*, pihak lawan aktif berusaha mencoba untuk mengeluarkan, membuat tidak *valid* atau memalsukan *watermark* [11]. Dalam *data hiding* tidak ada pihak lawan di dalamnya yang bertindak mengeluarkan pesan yang disembunyikan. Meskipun demikian, teknik *information hiding* membutuhkan ketahanan terhadap distorsi-distorsi yang mungkin dapat terjadi.

Steganografi dapat digunakan juga untuk melakukan perawatan atas kerahasiaan informasi yang berharga dari kemungkinan sabotase, pencuri, atau dari pihak yang tidak berwenang. Sayangnya, steganografi juga dapat digunakan untuk alasan yang ilegal. Sebagai contoh, steganografi dapat digunakan oleh para teroris untuk menyamarkan komunikasi mereka dari pihak luar.

Pada steganografi *digital* modern, penyembunyian pesan dilakukan pada media penampung *image*, audio dan video elektronik yang dikirim dengan menggunakan jaringan internet. Data dimasukkan ke dalam data redundan (data yang tersedia tetapi seringkali tidak diperlukan), seperti *field* pada protokol komunikasi, gambar grafik, dan sebagainya [9].

Dalam semua metode steganografi, secara alamiah teknik penyembunyiannya dapat dipisahkan dan dianalisa untuk mengetahui apa yang terjadi dalam keseluruhan proses. Hal tersebut dikategorikan dalam enam kategori steganografi [12] yaitu :

1. Sistem Substitusi

   Sistem steganografi substitusi menggantikan *bit-bit* yang tidak perlu dari suatu media dengan *bit-bit* dari pesan rahasia. Beberapa teknik steganografi yang ada menggunakan metode *Least Significant Bit* (LSB) untuk memproses pesan rahasianya.

2. Teknik *Domain Transform*

   Pada dasarnya teknik ini menyembunyikan data pesan dalam ”ruang transform” dari suatu sinyal. Sebagai ilustrasi adalah kompresi gambar dalam format JPEG yang akan berubah menjadi lebih kecil, pada saat perubahan itu terdapat ruang transform yang dapat digunakan untuk menyembunyikan informasi.

3. Teknik *Spread-Spectrum*

   Dalam *spread spectrum* yang langsung, aliran informasi yang ditransmisikan dibagi menjadi potongan-potongan kecil. Setiap potongan ditempatkan sebagai kanal frekuensi dari spektrum.

4. Metode Statistik

   Metode statistik ini menggunakan apa yang disebut skema steganografi ”1-*bit*”. Skema ini menempelkan satu *bit* informasi hanya dalam suatu pembawa *digital* sehingga mengakibatkan perubahan statistik meskipun sedikit.

5. Teknik Distorsi

   Metode steganografi ini membuat perubahan dalam obyek media untuk menyembunyikan informasi. Pesan rahasia didapatkan kembali jika algoritma yang dibandingkan berubah, berbeda dengan aslinya.

6. Metode Penurunan Media

   Secara khusus obyek media dipilih untuk menyembunyikan pesan, namun sebenarnya metode ini membuat/menurunkan sebuah media yang digunakan untuk menyembunyikan informasinya.

Pada sistem substitusi, terdapat beberapa metode yang bisa diterapkan untuk beberapa tipe media, di antaranya adalah metode *Least Significant Bit* (LSB).

2.2. Metode Steganografi LSB

Cara paling umum untuk menyembunyikan pesan adalah dengan memanfaatkan *Least Significant Bit* (LSB) [5]. Walaupun terdapat kekurangan pada metode ini, tetapi kemudahan implementasinya membuat metode ini tetap digunakan sampai sekarang. Contoh ilustrasinya sebagai berikut : jika digunakan *image* 24 *bit* warna sebagai media, sebuah *bit* dari masing-masing komponen *Red*, *Green*, dan *Blue*, dapat digunakan sehingga 3 *bit* dapat disimpan pada setiap *pixel*. Sebuah *image* 800 x 600 *pixel* dapat digunakan untuk menyembunyikan 1.440.000 *bit* (180.000 bytes) data rahasia. Misalnya, di bawah ini terdapat 3 *pixel* dari *image* 24 *bit* warna :

(00100111 11101001 11001000)
(00100111 11001000 11101001)
(11001000 00100111 11101001)

jika diinginkan untuk menyembunyikan karakter A (**100000011**) dihasilkan :

(00100111 11101000 11001000)
(00100110 11001000 11101000)
(11001000 00100111 11101001)

dapat dilihat bahwa hanya 3 *bit* saja yang perlu diubah untuk menyembunyikan karakter A ini.

Jika pesan = 10 *bit*, maka jumlah *byte* yang digunakan = 10 *byte*.

Contoh susunan *byte* yang lebih panjang :

00110011 10100010 11100010 10101011 00100110
10010110 11001001 11111001 10001000 10100011

Pesan : **1110010111**

Hasil penyisipan pada *bit* LSB :

00110011 10100011 11100011 10101010 00100110
10010111 11001000 11111001 10001001 10100011

Pada metode LSB [5], ukuran data yang akan disembunyikan bergantung pada ukuran *cover-object*.

Perubahan pada LSB ini akan terlalu kecil untuk terdeteksi oleh mata manusia sehingga pesan dapat disembunyikan secara efektif. Proses ekstraksi pesan dapat dengan mudah dilakukan dengan mengekstrak LSB dari masing-masing *pixel* pada stego secara berurutan dan menuliskannya ke *output file* yang akan berisi pesan

tersebut. Keuntungan metode LSB adalah mudah dalam pengimplementasian dan proses *encoding* yang cepat. Pada perkembangannya metode steganografi LSB selain diterapkan pada media *image*, juga bisa diterapkan pada media audio [10].

Kriteria steganografi yang baik meliputi tiga hal [11] yaitu :

1. *Imperceptible/Undetectability*

   Keberadaan pesan rahasia tidak dapat dipersepsi.

2. *Fidelity*

   Mutu *cover-object* tidak jauh berubah akibat *embedded*.

3. *Recovery*

   Data yang disembunyikan harus dapat diungkapkan kembali.

Kriteria *robustness* tidak terlalu penting karena tujuan yang utama dari steganografi adalah untuk menghindari kecurigaan (lawan tidak menyadari keberadaan pesan tersembunyi).

Sedangkan efektifitas teknik steganografi [1] bisa diukur menggunakan tiga prinsip berikut :

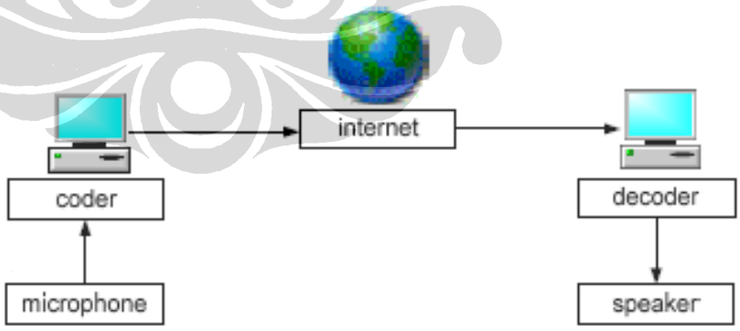
1. Besarnya data, semakin besar/banyak data yang bisa disembunyikan berarti semakin baik tekniknya.
2. Tingkat kesulitan deteksi, berhubungan dengan seberapa mudah seseorang dapat mendeteksi adanya pesan yang disembunyikan. Biasanya ada hubungan secara langsung antara besarnya data yang dapat disembunyikan dan seberapa mudah seseorang dapat mendeteksinya. Semakin besar jumlah informasi yang bisa disembunyikan pada suatu media, semakin meningkatkan kesempatan seseorang untuk dapat mendeteksi bahwa ada informasi yang tersembunyi dalam media tersebut.
3. Tingkat kesulitan penghapusan, melibatkan prinsip-prinsip bahwa seseorang yang sedang menyadap media semestinya tidak dapat dengan mudah menghapus data yang disembunyikan.

Satu hal esensial yang menjadi kelebihan steganografi adalah kemampuannya untuk menipu persepsi manusia, manusia tidak memiliki insting untuk mencurigai adanya arsip-arsip yang memiliki informasi yang tersembunyi di dalamnya, terutama bila arsip tersebut tampak seperti arsip normal lainnya. Namun steganografi juga bukanlah pengamanan yang sempurna karena metode pendeteksi pesan dalam

steganografi juga banyak dikembangkan yang disebut steganalisis [9], yaitu suatu teknik yang digunakan untuk mendeteksi penggunaan steganografi pada suatu data/media. Seorang steganalis tidak berusaha untuk melakukan dekripsi terhadap informasi yang tersembunyi dalam suatu data/media, yang dilakukan adalah berusaha untuk menemukannya. Terdapat beberapa cara yang dapat digunakan untuk mendeteksi steganografi seperti melakukan pengamatan terhadap suatu data/media dan membandingkannya dengan salinan data/media yang dianggap belum direkayasa, atau berusaha mendengarkan dan membandingkan perbedaannya dengan data/media lain bila data/media tersebut adalah dalam bentuk audio.

2.3. Pengertian *Voice over Internet Protocol* (VoIP)

Voice over Internet Protocol (juga disebut VoIP, *IP Telephony*, *Internet Telephony* atau *Digital Phone*) [3] adalah suatu teknologi yang memungkinkan para pihak melakukan komunikasi suara, video, dan data melalui suatu media telekomunikasi yang memanfaatkan protokol internet. Data suara yang bersifat analog dikonversi menjadi kode *digital*, kemudian dialirkan melalui suatu media jaringan dalam bentuk paket-paket data. Ilustrasi sederhana tentang komunikasi VoIP terlihat pada Gambar 2.3. VoIP termasuk sistem yang dapat menekan biaya pemeliharaan, karena dapat dengan mudah dipindah, ditambah dan diubah. Sistem VOIP juga mampu menangani jumlah panggilan yang banyak secara bersamaan. Terlepas dari regulasi di Indonesia [13] yang masih belum jelas terhadap VoIP sebagai suatu layanan komunikasi, secara umum layanan VoIP akan memberi masyarakat banyak pilihan dalam berkomunikasi.



Gambar 2.3. VoIP Secara Sederhana [3]

Dalam perancangan jaringan VoIP, yang perlu diperhatikan adalah masalah *delay* dan *bandwidth* [14].

Delay didefinisikan sebagai waktu yang dibutuhkan untuk mengirimkan data dari sumber (pengirim) ke tujuan (penerima), sedangkan *bandwidth* adalah kecepatan maksimum yang dapat digunakan untuk melakukan transmisi data antar komputer pada jaringan IP atau internet. Kualitas suara bagus tidaknya tergantung dari waktu *delay*. Besarnya *delay* maksimum yang direkomendasikan oleh *International Telecommunication Union Telecommunication Standardization Sector* (ITU-T) untuk aplikasi suara adalah 150 ms, sedangkan *delay* maksimum dengan kualitas suara yang masih dapat diterima pengguna adalah 250 ms.

Beberapa *delay* yang dapat mengganggu kualitas suara dalam perancangan jaringan VoIP [14] dapat dikelompokkan menjadi :

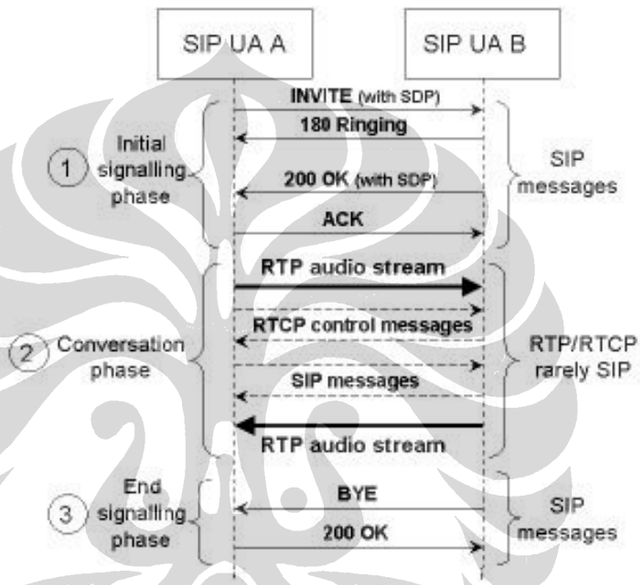
1. *Propagation delay* (*delay* yang terjadi akibat transmisi melalui jarak antar pengirim dan penerima).
2. *Serialization delay* (*delay* pada saat proses peletakan *bit* ke dalam *circuit*).
3. *Processing delay* (*delay* yang terjadi saat proses *coding, compression, decompression* dan *decoding*).
4. *Packetization delay* (*delay* yang terjadi saat proses paketisasi *digital voice sample).*
5. *Queuing delay* (*delay* akibat waktu tunggu paket sampai dilayani).
6. *Jitter buffer* (*delay* akibat adanya *buffer* untuk mengatasi *jitter*).

Dalam perancangan VoIP, *bandwidth* merupakan suatu yang harus diperhitungkan agar dapat memenuhi kebutuhan pelanggan yang dapat digunakan sebagai parameter untuk menghitung jumlah peralatan yang dibutuhkan dalam suatu jaringan. Perhitungan ini juga sangat diperlukan dalam efisiensi jaringan dan biaya serta sebagai acuan pemenuhan kebutuhan untuk pengembangan di masa mendatang.

Terdapat empat macam protokol utama yang memungkinkan terjadinya komunikasi VoIP [4], diantaranya :

1. *Signalling Protocols*, adalah protokol yang membuat, memodifikasi, dan memutuskan komunikasi antara pihak yang sedang berkomunikasi. Yang saat ini populer adalah *Session Initiation Protocol* (SIP), H.323, dan H.248/Megaco [3].

2. *Transport protocols* yang menyediakan *end-to-end connection*. *Real-time Transport Protocol* (RTP) [3] merupakan protokol yang sesuai untuk aplikasi pengiriman audio secara *real-time*. RTP menggunakan koneksi *User Datagram Protocol* (UDP) atau terkadang menggunakan *Transmission Control Protocol* (TCP) untuk mengirimkan data suara *digital*.
3. *Speech codecs*, yaitu protokol yang bertugas melakukan kompresi/dekompresi data suara yang akan dikirimkan atau diterima melalui jaringan IP. Contoh : G.711, G.729, G.723.1, Speex [6], dan sebagainya.
4. *Supplementary protocols*, yaitu protokol yang berisi fungsional-fungsional yang melengkapi fungsionalitas untuk bekerjanya layanan VoIP. Misalnya RTCP atau RSVP [3].



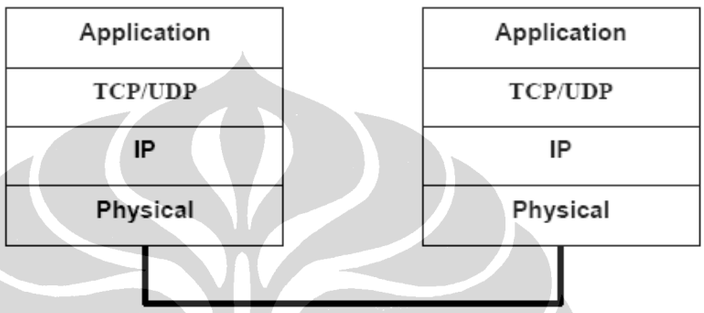
Gambar 2.4. *Set up* Komunikasi VoIP dengan Protokol SIP/RTP [4]

Secara umum, telepon IP terdiri dari dua fase, antara lain [4] *signalling phase* dan *conversation phase*. Pada kedua fase tersebut terjadi lalu lintas ataupun koneksi di antara pihak yang melakukan komunikasi dengan VoIP. Pada umumnya, fase *signalling* menggunakan SIP [3], sedangkan untuk transportasi audio dengan RTP. Pada saat terjadi proses *signalling*, terjadi pertukaran SIP *message* antara *end-point* pihak yang terlibat dalam komunikasi VoIP. Setiap SIP *message* melalui SIP *network server*. Setelah fase *signalling* selesai, fase *conversation* dimulai. Yaitu, aliran audio

(melalui RTP) dialirkan dari pihak pemanggil (*caller*) ke pihak yang dipanggil (*callee*). Ilustrasi lebih lengkap, seperti contoh *set up* komunikasi VoIP yang terlihat pada Gambar 2.4. Sebagai tambahan, RTCP merupakan *control protocol* bagi RTP dan didesain untuk melakukan kontrol kualitas atas servis dalam sebuah sesi komunikasi dengan VoIP.

2.4. Protokol TCP/IP

TCP/IP [14] merupakan sebuah protokol yang digunakan pada jaringan internet. Protokol ini terdiri dari dua bagian besar, yaitu TCP dan IP. Ilustrasi pemrosesan data untuk dikirimkan dengan menggunakan protokol TCP/IP diberikan pada Gambar 2.5 di bawah ini.



Gambar 2.5. Mekanisme Protokol TCP/IP [14]

Dalam mentransmisikan data pada *layer transport* ada dua protokol yang berperan yaitu TCP dan UDP. TCP merupakan protokol yang *connection-oriented* yang artinya menjaga reliabilitas hubungan komunikasi *end-to-end*. Konsep dasar cara kerja TCP adalah mengirim dan menerima segmen-segmen informasi dengan panjang data bervariasi pada suatu datagram internet. TCP menjamin reliabilitas hubungan komunikasi karena melakukan perbaikan terhadap data yang rusak, hilang atau kesalahan kirim. Dalam hubungan VoIP, TCP digunakan pada saat *signalling*, TCP digunakan untuk menjamin *set up* suatu *call* pada sesi *signalling.* TCP tidak digunakan dalam pengiriman data suara pada VoIP karena pada suatu komunikasi data VoIP penanganan data yang mengalami keterlambatan lebih penting daripada penanganan paket yang hilang.

UDP [14] merupakan *transport protocol* yang lebih sederhana digunakan untuk situasi yang tidak mementingkan mekanisme reliabilitas. UDP pada VoIP digunakan

untuk mengirimkan *audio stream* yang dikirimkan secara terus menerus. UDP digunakan pada VoIP karena pada pengiriman *audio streaming* yang berlangsung terus-menerus lebih mementingkan kecepatan pengiriman data agar tiba di tujuan tanpa memperhatikan adanya paket yang hilang walaupun mencapai 50% dari jumlah paket yang dikirimkan [3]. Karena UDP mampu mengirimkan data *streaming* dengan cepat, maka dalam teknologi VoIP, UDP merupakan salah satu protokol penting yang digunakan sebagai *header* pada pengiriman data selain RTP dan IP. Untuk mengurangi jumlah paket yang hilang saat pengiriman data (karena tidak terdapat mekanisme pengiriman ulang) maka pada teknologi VoIP pengiriman data banyak dilakukan pada *private network* [3].

IP [14] didesain untuk interkoneksi sistem komunikasi komputer pada jaringan *paket switched*. Pada jaringan TCP/IP, sebuah komputer diidentifikasi dengan alamat IP. Tiap-tiap komputer memiliki alamat IP yang unik, masing-masing berbeda satu sama lainnya. Hal ini dilakukan untuk mencegah kesalahan pada transfer data.

2.5. Protokol ITU-T H.323

H.323 [14] merupakan protokol komunikasi yang direkomendasikan oleh ITU-T *Study Group* 16 yang mendefinisikan sistem dan protokol untuk komunikasi multimedia termasuk audio, video dan data di jaringan internet. Perangkat-perangkat telepon internet yang berbeda-beda namun mengacu ke standar H.323 tersebut, akan dapat saling berinteroperasi satu sama lain. Pada H.323 terdapat beberapa protokol dalam pengiriman data yang mendukung agar data terkirim *real-time*. Beberapa protokol H.323 pada *layer network* dan *transport* adalah RTP, RTCP, dan RSVP.

RTP dapat digunakan untuk beberapa macam data *stream* yang *real-time* seperti data suara dan data video.

RTCP (*Real-Time Control Protocol*) [14] digunakan untuk mengirimkan paket *control* setiap terminal yang berpartisipasi pada percakapan yang digunakan sebagai informasi untuk kualitas transmisi pada jaringan.

RSVP (*Resource Reservation Protocol*) [14] bekerja pada *layer transport* dan digunakan untuk menyediakan *bandwidth* agar data suara yang dikirimkan tidak mengalami *delay* ataupun kerusakan saat mencapai alamat. RSVP merupakan

signalling protocol tambahan pada VoIP yang mempengaruhi QoS (*Quality of Service*) [3].

2.6. Protokol SIP (*Session Initiation Protocol*)

SIP [14] merupakan protokol persinyalan yang bertujuan untuk mengendalikan inisiasi, modifikasi, serta terminasi sesi-sesi multimedia, termasuk sesi komunikasi audio atau video. SIP merupakan protokol berbasis teks yang mirip dengan protokol *Hypertext Transfer Protocol* (HTTP) dan *Simple Mail Transfer Protocol* (SMTP).

SIP memiliki fungsi-fungsi yang didefinisikan sebagai berikut :

1. *User location* SIP menyediakan kemampuan untuk menemukan lokasi pengguna akhir yang bermaksud akan membangun sebuah sesi atau mengirimkan sebuah permintaan.
2. *User capabilities* SIP memungkinkan determinasi kemampuan media dari perangkat yang terlibat di dalam sesi.
3. *User availability* SIP memungkinkan determinasi keinginan pengguna untuk melakukan komunikasi.
4. *Session set up* SIP memungkinkan modifikasi, transfer, dan terminasi dari sebuah sesi aktif.

Sebagai bahan perbandingan, pada Tabel 2.1 terlihat perbedaan antara protokol H.323 dan SIP dari fitur-fitur yang digunakan.

Tabel 2.1. Perbedaan antara Protokol H.323 dan SIP [3]

Feature	H.323	SIP
Architecture	Stack Implementation	Element Implementation
Complexity	Complex	Simple
Standards body	ITU	IETF
Protocol	Mostly TCP	Mostly UDP
Protocol Encoding	Binary (ASN.1, Q.931)	Text (HTTP-ish)
Server processing	State-full	State-less, Transaction oriented
Addressing	Flat alias, E.164, email	SIP, E.164, URLs
Call Setup delay	V1: 6-7x RTT to V3: 1.5-2.5x RTT	1.5x RTT
Mid-call failure	Fail	Live
Loop Detection	V1:No, v3: Path Value	Yes – "via" field, time, hops
Manageability	Yes	No
Call control	Yes	Yes

2.7. Kompresi Data Suara

Teknik digitalisasi suara manusia dibedakan ke dalam dua kategori utama [6]. Pertama adalah proses konversi sinyal analog ke sinyal *digital* oleh *Analog-Digital Converter* (ADC) di sisi pengirim dan *Digital-Analog Converter* (DAC) di sisi penerima. Teknik yang umum dipakai adalah *Pulse Code Modulation* (PCM), *Differential* PCM (DPCM) dan *Delta Modulation* (DM).

Kategori kedua adalah proses kompresi data suara untuk menghasilkan data suara dengan laju *bit* yang serendah mungkin namun dengan tetap menjaga kualitas suara agar dapat dimengerti dengan jelas oleh pendengar. Proses ini sering disebut sebagai *vocoder* (*voice coder*) [6]. Teknik kompresi suara yang digunakan pada sistem komunikasi VoIP bertujuan untuk penghematan besarnya *bandwidth* atau data paket suara yang akan ditransmisikan. Sebuah kanal suara (audio) yang baik tanpa di kompresi akan mengambil *bandwidth* sekitar 64 kbps. Dengan adanya teknik kompresi, sebuah kanal suara dapat dihemat menjadi sekitar 6 kbps (*half-duplex*).

ITU-T membuat beberapa standar untuk *voice coding* [3] yang direkomendasikan untuk implementasi VoIP. Beberapa standar yang sering dikenal antara lain: G.711, G.723.1, G.726, dan sebagainya.

G.711 adalah suatu standar internasional untuk kompresi audio dengan menggunakan teknik *Pulse Code Modulation* (PCM) dalam pengiriman suara. PCM [14] mengkonversikan sinyal analog ke bentuk *digital* dengan melakukan *sampling* sinyal analog tersebut 8000 kali/detik dan dikodekan dalam kode angka. Jarak antar sampel adalah 125 μ detik. Sinyal analog pada suatu percakapan diasumsikan berfrekuensi 300 Hz – 3400 Hz. Sinyal tersampel lalu dikonversikan ke bentuk diskrit. Sinyal diskrit ini direpresentasikan dengan kode yang disesuaikan dengan amplitudo dari sinyal sampel. Format PCM menggunakan 8 *bit* untuk pengkodeannya. Laju transmisi diperoleh dengan mengalikan 8000 sampel/detik dengan 8 *bit*/sampel, menghasilkan 64.000 *bit*/detik. *Bit rate* 64 kbps ini merupakan standar transmisi untuk satu kanal telepon *digital*.

Standar G.711 merupakan teknik kompresi yang tidak efisien, karena akan memakan *bandwidth* 64 kbps untuk kanal pembicaraan. Agar *bandwidth* yang digunakan tidak besar dan tidak mengesampingkan kualitas suara, maka solusi yang digunakan untuk pengkompresi digunakan standar G.723.1.

Pengkode sinyal suara G.723.1 adalah jenis pengkode suara yang direkomendasikan untuk terminal multimedia dengan *bit* rate rendah. G.723.1 memiliki *dual rate speech coder* yang dapat di-*switch* pada batas 5.3 kbit/s dan 6.3 kbit/s. Dengan memiliki *dual rate speech coder* ini maka G.723.1 memiliki fleksibilitas dalam beradaptasi terhadap informasi yang dikandung oleh sinyal suara, sehingga hasil rekonstruksinya menjadi sangat mirip dengan aslinya. Sinyal eksitasi untuk *bit-rate* rendah dikodekan dengan *Algebraic Code Excited Linier Prediction* (ACELP) [3] sedangkan untuk *rate* tinggi dikodekan dengan menggunakan *Multipulse Maximum Likelihood Quantization* (MP-MLQ) [3]. Rate yang lebih tinggi menghasilkan kualitas yang lebih baik.

Tabel 2.2. Perbandingan Teknik Kompresi dan Kinerjanya [3]

Teknik Kompresi	kbps	MIPS	ms	MOS
G.711 PCM	64	0,34	0,125	4,1
G.726 ADPCM	32	14	0,125	3,85
G.728 LD-CELP	16	33	0,625	3,61
G.729 CS-ACELP	8	20	10	3,92
G.729 x2 Encoding	8	20	10	3,27
G.729 x3 Encoding	8	20	10	2,68
G.729a CS-ACELP	8	10,5	10	3,7
G.723.1 MP-MLQ	6,3		30	3,9
G.723.1 ACELP	5,3		30	3,65

Pada Tabel 2.2 menampilkan data beberapa teknik kompresi dengan beberapa parameter yang mencerminkan kinerja dari teknik kompresi tersebut.

Kolom kbps memperlihatkan berapa lebar *bandwidth* yang digunakan untuk mengirim suara yang dikompres menggunakan teknik kompresi tertentu. MIPS (*Mega Instruction Per Second*) memperlihatkan berapa kebutuhan waktu pemrosesan data pada saat melakukan kompresi suara dalam juta instruksi per detik. Mili-detik (ms) adalah waktu yang dibutuhkan untuk melakukan kompresi. *Mean Opinion Score* (MOS) [3] adalah nilai opini pendengar di ujung pesawat penerima.

Selain standar *codec* yang sudah ada, terdapat *codec* yang biasa digunakan dalam teknologi VoIP, seperti iLBC, LPC, Speex, AMR-NB, GSM-FR, dan sebagainya. Sebagai bahan perbandingan, lebih jelas terdapat pada Tabel 2.3 di bawah ini.

Tabel 2.3. Perbandingan Fitur *Codec* [6]

Codec	Rate (kHz)	Bit-rate (kbps)	Delay frame+ Lookahead (ms)	Multi -rate	Embed-ded	VBR	PLC	Bit-robust	License
Speex	8, 16, 32	2.15-24.6 (NB) 4-44.2 (WB)	20+10 (NB) 20+14 (WB)	yes	yes	yes	yes		open-source/ free software
iLBC	8	15.2 or 13.3	20+5 or 30+10				yes		no charge, but not open-source
AMR-NB	8	4.75-12.2	20+5?	yes			yes	yes	Proprietary
AMR-WB (G.722.2)	16	6.6-23.85	20+5?	yes			yes	yes	Proprietary
G.722.1 (Siren7)	16	(16) 24, 32	20+20	yes			yes	yes	no charge, but not open-source
G.729	8	8	10+5				yes	yes	Proprietary
GSM-FR	8	13	20+?				?	?	patented?
G.723.1	8	5.3 6.3	37.5				yes	?	Proprietary
G.728	8	16	0.625						Proprietary
G.722	16	48 56 64	?		yes			?	?

Definisi :

Multi-rate, mengijinkan *codec* untuk mengubah *bit-rate* secara dinamis pada setiap waktu.

Embedded, codec yang menempelkan *narrowband bit-streams* pada *wideband bit-streams*.

VBR : *Variable Bit-Rate*.

PLC : *Packet Loss Concealment*.

Bit-robust, kekuatan terhadap kerusakan pada tingkat *bit*, sebagaimana yang ditemukan pada jaringan *wireless.*

Speex *codec* [6] (http://www.speex.org) merupakan pengembangan *open source codec* berdasarkan teknik *encoding* CELP [3] (karena telah terbukti dapat diandalkan untuk *encoding* baik *bit* rate rendah maupun yang tinggi) dengan karakteristik yang

ditawarkan antara lain *Voice Active Detection* (VAD) dan *Discontinuous Transmission* (DTX). Tidak seperti *codec* suara yang lain, Speex tidak hanya didesain untuk *mobile phones* melainkan lebih untuk jaringan paket dan aplikasi VoIP. Kompresi berdasarkan *file* juga didukung oleh Speex. *Codec* ini didesain menjadi sangat fleksibel dan mendukung area yang luas dari kualitas percakapan dan *bit rate*. Dukungan untuk kualitas percakapan yang sangat bagus berarti bahwa Speex dapat meng-*encode* percakapan *wideband* (16 kHz *sampling rate*) sebagai tambahan percakapan *narrowband* (kualitas telepon, 8 kHz *sampling rate*). Desain khusus untuk VoIP ini berarti bahwa Speex kuat terhadap kemungkinan kehilangan paket.

2.8. Pengukuran Kualitas Suara

Untuk memberikan penilaian kualitas suara telepon internet (VoIP) diperlukan pengukuran kuantitatif yang mencerminkan persepsi pengguna akan kualitas layanan. Secara umum, teknik pengukuran kualitas suara [3] mencakup dua hal sebagai berikut:

1. Pengukuran subyektif yaitu teknik pengukuran yang didasarkan kepada opini pendengar secara langsung. Contoh teknik pengukuran yang sangat populer pada sistem komunikasi suara adalah MOS.
2. Pengukuran obyektif yaitu teknik pengukuran yang didasarkan kepada pengukuran sinyal suara. Dua tipe pengukuran obyektif yaitu pengukuran intrusif, misalnya PESQ (*Perceptual Evaluation of Speech Quality*) dan pengukuran non-intrusif (misalnya *E-model* dan *Extended E-model*).

Teknik pengukuran dengan MOS dan PESQ pada prinsipnya menghasilkan nilai yang memenuhi skala mulai dari 1 (*bad*) sampai dengan 5 (*excellent*).