

## LAN 環境における負荷適応制御を用いた 低遅延リアルタイム音声通信システム

星 徹† 谷川 桂子† 松井 進†  
岩見 直子† 寺田 松 昭†\*

PCのネットワーク化とマルチメディア化の進展により、PCとLANなどのパケットネットワークを用いる双方向リアルタイム音声通信システムが実現されつつある。このようなシステムでは、データ通信の仕組を維持しつつ、実時間性の保証が難しいパケットネットワークの時間特性を克服して、エンドツーエンドの通信遅延時間を削減するとともに、遅延ゆらぎを吸収し音切れのない双方向音声通信を行うことが課題である。本論文では、IPプロトコルを用いベストエフォートで通信するLAN環境において低遅延の双方向リアルタイム音声通信を実現する音声通信システムを提案する。提案方式の特徴は、(1) エンドツーエンドでの遅延変動に動的に適応して受信側PCに設けたジッタ吸収バッファを音声品質を維持しつつサイズを最小になるように制御し、遅延時間を短縮する負荷適応制御を行う、(2) PCでの音声通信遅延を削減するためにアルゴリズム遅延が小さいPCM/ADPCM方式とし、さらにハードウェアコーデックを用いる方式とする、(3) データ通信環境との統合を図るためにTCP(UDP)/IP通信を用いる、である。システムを試作し実験的評価を行うことにより、エンドツーエンドの遅延時間をITU-Tがリアルタイム双方向通信アプリケーションとして満足できる値として勧告している150msを大幅に下回る100ms以下にする低遅延双方向リアルタイム通信システムが実現できること、ネットワークの負荷変動に影響されにくく音声の途切れを意識させない通話品質を実現できることを確認している。

### LAN Based Low Delay Real Time Voice Communication System with Adaptive Jitter Control

TOHRU HOSHI,† KÈIKO TANIGAWA,† SUSUMU MATSUI,†  
NAOKO IWAMI† and MATSUAKI TERADA†\*

Progress being made in the development of multimedia network environments with networked multimedia PCs is leading to the day when LAN-based real time voice communication systems can be achieved. The major problems with LANs are delay, jitter and packet loss due to their best effort transfer characteristics; accordingly, the major issues of LAN-based systems are how to reduce end-to-end delay, absorb jitter and recover packet loss in order to provide high quality bi-directional real time voice communication. In this paper, a LAN-based real time voice communication system using IP protocols with adaptive jitter control is proposed. The advantageous characteristics of this system are the introduction of an adaptive jitter control mechanism that minimizes end-to-end delay by optimizing jitter buffer size, the adoption of small algorithm delay hardware codecs, and the use of a TCP(UDP)/IP environment for voice communication in order to integrate data and voice applications. Results obtained in tests on a prototype system show that delay of less than 100 ms is achieved, which satisfies the ITU-T G.114 Recommendation of 150 ms as an acceptable range for bi-directional realtime voice communication. In addition, stable voice quality is achieved that is little affected by the disturbance caused by dynamic changes in network load.

#### 1. はじめに

PCのネットワーク化とマルチメディア化の進展に

より、音声通信機能を実装したクライアントPCをLANに接続して、リアルタイムで音声通信するLAN環境でのリアルタイム音声通信システムの構築が可能となってきた。また、このようなLAN、WANを含むIPネットワークにおけるベストエフォート網を対象としたリアルタイム音声通信方式、リアルタイム音声通信プロトコルなどの研究開発と標準化活動が活発に進められている<sup>1)~5)</sup>。

† 株式会社日立製作所システム開発研究所  
Systems Development Laboratory, Hitachi, Ltd.

☆ 現在、東京農工大学工学部  
Presently with Department of Computer Science,  
Tokyo University of Agriculture and Technology

リアルタイム音声通信においては、(1) 遅延を極力小さくし会話をしやすくすることと、(2) 遅延のゆらぎ(ジッタ)を抑え、かつ途切れのない音声を再生することが要求される。特に LAN 環境(LAN: イントラネットなどのベストエフォート型の IP ネットワーク)で PC によるリアルタイム音声通信を行う場合、

- エンドツーエンド遅延が大きい
- ジッタが発生する
- パケットロスが発生する
- メールなどの通信アプリケーションとの連携を図る

などの課題を克服する必要がある。

以下、これらについて、背景、課題と関連研究について述べる。

### (1) エンドツーエンド遅延

送受信端末までを含めたエンドツーエンドでの遅延が大きい IP ネットワークにおける遅延は負荷状態により変動し、またルートによっても変動する。端末側においては、音声圧縮伸長処理、パケット化処理などにより遅延が発生する。エンドツーエンド遅延と音声品質の関係は、通話遅延に関する ITU-T 勧告 G.114 において、0~150 ms 間では、多くのユーザアプリケーションで許容できる、150~400 ms 間では、管理側が考慮していれば許容範囲である、400 ms 以上は、一般的には許容不可、と勧告している<sup>6)</sup>。また音声通信のユーザビリティという視点で電話網、衛星網、CB (Citizen Band) 間で相対的に比較論じた報告がされている。これによると遅延時間約 100 ms 以下である電話網は、ユーザビリティ  $\approx 0.9 \sim 1.0$  としている。ここでユーザビリティが大きいほど良質である。ちなみに、衛星網は 0.2~0.5、CB 網は、0.1 程度としている<sup>7)</sup>。

エンドツーエンドの遅延を削減するには、端末(PC)とネットワーク装置(ルータなど)それぞれでの遅延時間を削減する必要がある。

端末での遅延に関しては、音声符号化復号化時間の削減が大きな課題であり、これは CODEC の符号化方式と実現方式に大きく依存する。符号化方式に関して、G.723.1、G.729A などの高圧縮のパケット型 CODEC は、従来の PCM (G.711)/ADPCM (G.726) CODEC に比べ、帯域削減の効果は大きい<sup>8)</sup>。また、音声符号化処理の実現方式に関しては、PC ベースの端末を経済的に実現するためにソフトウェア CODEC を用いたシステムが多数報告されているが、ハードウェア CODEC を採用したシステムに比べて処理遅延が大きくなる<sup>2)</sup>。さらに、処理性能の低い PC では、他

の高負荷のアプリケーションを動かしたときに音声処理が間に合わなくなり音切れが発生するなどの課題を含んでいる。

音声統合ルータなど、ネットワーク装置における遅延削減に関しては、上記課題に加えて、複数の音声ストリームを同時に扱うこと、音声以外のデータトラフィックが共存することが端末と異なる点である。ルータ処理での音声の遅延を抑えるために、定期的に音声パケットの転送処理ができるように長いパケットは分割して、その間に音声パケット処理を行う Weighted Fair Queuing 方式などが提案されている<sup>9)</sup>。一方で IP ネットワークの転送プロトコル自体を音声などのリアルタイム性の高いストリームに対応できるようにするために、IP パケットの TOS フィールドを用いて優先制御を行う Differentiated Service、帯域を予約する RSVP などの QoS を保証する新しいプロトコルが提案されて一部実用化されている<sup>10),11)</sup>。しかし、これらは、普遍的に利用できるまでに普及していない。

音声通信システムにおける遅延時間削減に関して、端末とネットワーク両者で対応する方式、端末のみで対応する方式、ネットワークで対応する方式の 3 つのアプローチがある。本論文では、ネットワーク側での対応は行わず端末の処理のみで対応する方式とする。これにより、現状のベストエフォート型の IP ネットワークで適用できるのみならず、今後普及すると想定される RSVP などの帯域予約型ネットワーク、Differentiated Service などの優先制御型ネットワークなどの新たな IP ネットワークまで対応することが可能となる。

### (2) エンドツーエンドのジッタ

IP ネットワークでは遅延とともにジッタが発生する。これはネットワーク内および端末における瞬時の負荷変動、パケット単位でのルート変更による遅延変動などに起因する。途切れのない通話を実現するためには、ジッタの振舞いの把握と適切なジッタ吸収制御が要求される。

パケットネットワークにおけるジッタの振舞いに関して多くの報告がなされており、これらは解析的手法によるアプローチ、シミュレーションによるアプローチ、実験的アプローチに大きく分類される。当初よりマルチメディアリアルタイム通信を行うことを目的とした ATM ネットワークに関して多くの研究がされている。ATM ノードで CBR トラフィックの多重化時におけるジッタ特性の解析、複数段の ATM ノードでのエンドツーエンドの ATM セルのジッタ特性などが報告されている<sup>12)~14)</sup>。実験的アプローチなどによる精度向上などが今後の課題である。また ATM ネット

ワークの実現の観点からエンドツーエンドの ATM レイヤの QoS パラメータが ATM Forum で規定されており、ジッタに関しては peak-to-peak CDV として規定されている<sup>15)</sup>。

一方、ベストエフォート網である IP ネットワークにおけるエンドツーエンドの遅延とジッタに関しては、解析的アプローチとして、それぞれのノードのキューが独立であるという仮定を持たせることによりノードを M/M/1, M/D/1 キューなどに単純化しキューイングネットワークモデルとした解析的アプローチがパケットネットワークの初期に提案されて利用されてきた<sup>16)</sup>。複雑なトラフィック特性を示す IP ネットワークでの精度を高めるために実測データを基にした実験、シミュレーションを併用したアプローチなどが報告されている<sup>17)</sup>。音声通信に関連する研究として UDP パケットを一定間隔で発生して、ラウンドトリップ遅延を計測して遅延時間の分布などを求めたものが報告されており、パケット遅延が負の指数分布で近似できることが示されている<sup>17)</sup>。これは、本論文で前提としている ARPANET におけるエンドツーエンド遅延測定結果を基に鈴木らが解析した結果と一致している<sup>18)</sup>。

エンドツーエンドのジッタ吸収制御に関して、ジッタ吸収制御をネットワークノードのみで行うアーキテクチャ、端末のみで行うアーキテクチャ、両者で行うアーキテクチャがそれぞれ提案されている<sup>19)~22)</sup>。ATM ネットワーク、IP ネットワークにおける RSVP, Differentiated Service などはメッセージを中継する個々のネットワークノードでの制御を前提としている。一方、現状の IP ネットワークでは端末のみでジッタ吸収制御を行う必要がある。これに対して、固定長のジッタ吸収バッファを設ける方式と、ジッタ特性により動的にジッタ吸収バッファサイズを制御する方式があげられる。固定長方式では、初期の LAN 電話システムである Etherphone ではトークスパートごとに固定のジッタ吸収時間を決定する方式を採用している。ただし決定アルゴリズムは明確にされていない<sup>20)</sup>。また、最近の PC ベースのインターネット電話端末の実験評価において、ジッタ吸収処理が十分できていないものがあることが報告されている<sup>23)</sup>。

動的ジッタ吸収バッファ制御方式では、ジッタ吸収時間を小さくしすぎると、遅延時間は小さくできるが遅れてきた音声パケットの取りこぼしが生じて音切れが生じ音質が悪化する。逆にジッタ吸収時間を大きくしすぎると、音声パケットの取りこぼしなくなり音質を向上させることができるが、その分遅延時間が長くなり実時間での会話に支障をきたす。放送型やオン

デマンド型のような片方向通信の場合は大きなジッタ吸収時間をとることができるが、電話のようなリアルタイム性の強い双方向通信の場合はジッタ吸収時間を短くする必要がある。したがって、音切れを起こさずかつできるだけジッタ吸収バッファを小さくし遅延時間を小さくする、きめ細かなジッタ吸収制御を行うことが重要な課題である。ジッタ吸収制御の具体的提案が、広域パケットネットワークを対象にした遅延適応制御の提案が鈴木らによりされている<sup>18)</sup>。これは遅延時間とジッタがともに大きいネットワークを対象としたものである。実測データを基にモデル化し、解析とシミュレーションによる評価を行いリアルタイム通信への有効性を示しているが、実システムでの評価は行われていない。

### (3) パケットロス

パケットロスの多いネットワークでは、パケットロスへの対処が必要である。一方、音声通信の場合 2~5% のパケットロスがあっても会話に支障をきたさないという研究結果が報告されている<sup>24)</sup>。この許容値を超えるパケットロスが発生した場合に音切れが目立ち会話に支障をきたす。パケットロスの多いインターネットを対象に、受信側でのロスパケットの補正方式、送信側での FEC (Forward Error Correction) の付与などのエラー回復処理が提案されている<sup>21)</sup>。しかし、LAN 環境においては、パケットロス率は小さく、エラー回復処理の必要性はインターネットの場合ほど多くはない。

### (4) メールなどの通信アプリケーションとの連携

オフィスにおいては、電子メール、データカンファレンスなどのリアルタイムグループウェア、ワークフローなどのコミュニケーション、コラボレーションアプリケーションと電話系の通信アプリケーションが独立に提供されてきている。PC を用いる音声通信システムはこれらのアプリケーションの統合を行いやすくする。このようなシステムとして、リアルタイム型の通信アプリケーションと蓄積型の通信アプリケーションのシームレス化を図るシステムなどが提案されている<sup>25),26)</sup>。統合システム実現のうえでデータ通信と整合性の良い通信方式が求められる。

本論文では、LAN 環境における低遅延リアルタイム音声通信システムを提案し、試作システムを開発し実験的に評価を行う。提案するシステムは以下の特徴がある。

(1) ジッタ吸収制御は端末のみで行う。網と端末を合わせたエンドツーエンドでの音声パケットの到着時間間隔を計測して、これよりジッタの分

布を求めジッタ吸収制御を行う。以下これを負荷適応制御と呼ぶ。端末とネットワークを含めた系全体での負荷状態に動的に適應することができる。

- (2) CODECはアルゴリズム処理遅延の少なく圧縮音声品質の良いPCM/ADPCM方式を採用し、かつ専用のハードウェアで実現する。
- (3) データ通信環境との統合を図るためにTCP(UDP)/IP通信を利用する。専用のリアルタイム通信プロトコルを提案するのではなく、TCP(UDP)/IPプラットフォーム上でリアルタイム音声通信を実現する。音声ストリームはUDPを用いて通信し、呼制御はTCPを用いて通信する。

本論文では、鈴木らが提唱した遅延適応制御基本方式<sup>18)</sup>をベースに、新たにLANとPCによる環境での低遅延音声通信システムの実現方式を提案し、実適用することにより、有用性を実験的に評価する。

以下、2章で、LAN環境でのリアルタイム音声通信システムの課題とアプローチを述べる。3章では提案する負荷適応制御を用いたLAN環境における低遅延リアルタイム音声通信システムの実現方式を、4章で実験的評価について述べる。

## 2. LAN環境におけるリアルタイム音声通信システム

### 2.1 対象とするパケットネットワーク

音声通信の観点から代表的なベストエフォート型通信を行うパケット交換網であるLAN、インターネット、X.25を表1で比較する。広域ネットワークにおいてはエラー再送機能があるX.25のような高信頼ネットワークではパケットロスが発生しないが、ノード間で再送などのエラー回復処理を行うため、遅延、ジッタともに増大しリアルタイム音声通信に適さない。同じく広域ネットワークであるインターネットの場合、ネットワーク層(IP層)での再送機能がないので遅延はX.25に比べ少なくなるがパケットロスが多くなり逆に音声通信の許容限度を超える可能性がある。一方、

表1 ネットワーク特性比較

Table 1 Comparison of network characteristics.

項目	X.25パケット網	インターネット(IP)	LAN/イントラネット(IP)
パケットロス	無	大	小~中
遅延	大	中~大	小~中
ジッタ	大	中~大	小~中

インターネットと同様、IPプロトコルを用いたLANおよびLANを専用線で接続したイントラネットでは、インターネットに比較して遅延、ジッタ、パケットロスともに比較的少ない。提案するシステムは、このようなLAN環境を対象としている。

### 2.2 LANベースリアルタイム音声通信システムの構成

提案するリアルタイム音声通信システムは、PCにリアルタイム音声通信機能を実現し、LANで接続されたPC間でIPプロトコルを用いて通話を可能とすること、PCと既存電話の相互通話を実現すること、メール、グループウェアなどと音声通信の融合化を図りオフィスにおけるコミュニケーション・コラボレーション機能を統一的に実現することである。

LANベースリアルタイム音声通信システムの構成を図1に示す。リアルタイム音声通信機能を搭載したPC、LAN、および既存電話網接続用ゲートウェイなどにより構成される。これにより、PC-PC間、PC-電話間の音声通信を可能とする。

図2に音声通信の機能構成を示し動作を説明する。発信者からのGUI操作で発信側PCの電話接続制御の呼制御から着信側PCへ呼がTCP/IPを用いて設定される。ハンドセットから入出力部に取り込まれた音声デジタル化されCODECで符号化される。符号化された音声ストリームは音声通信制御で一定の間隔で音声パケット化され、シーケンス番号などのヘッダ情報、呼設定先のPCのIPアドレスなどを付与され、UDP/IP通信処理とNIC(Network Interface Card)を経てLAN上へ送出される。受信側PCでは、NIC、UDP/IPを経て受信した音声パケットは音声通信処理のジッタ吸収バッファにいったん取り込まれる。ここで負荷適応制御によりジッタ吸収を行い、デパケット化された音声ストリームとしてCODECに渡される。CODECで復号化後、アナログ変換され出力される。

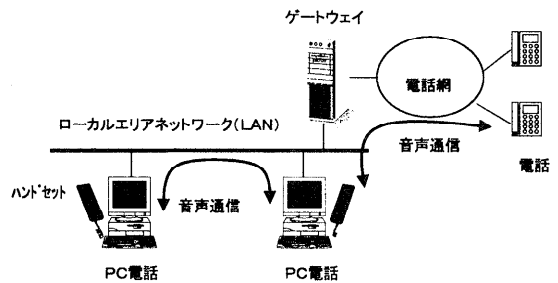


図1 LANベースリアルタイム音声通信システム

Fig. 1 LAN based realtime voice communication system.

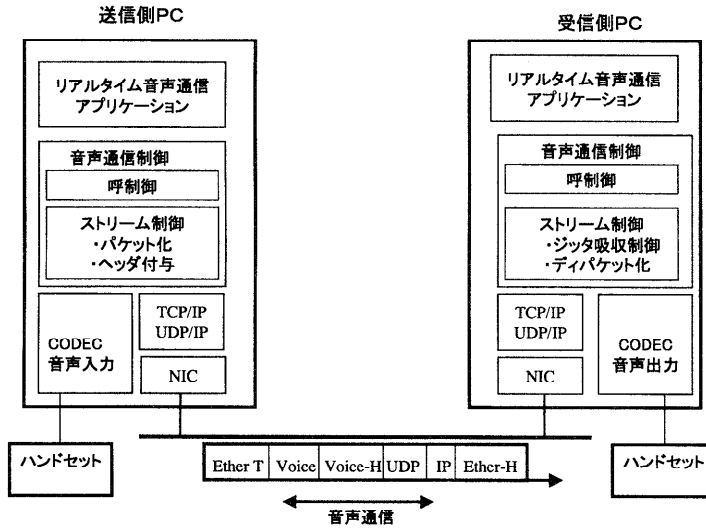


図 2 音声通信構成

Fig. 2 Realtime voice communication configuration.

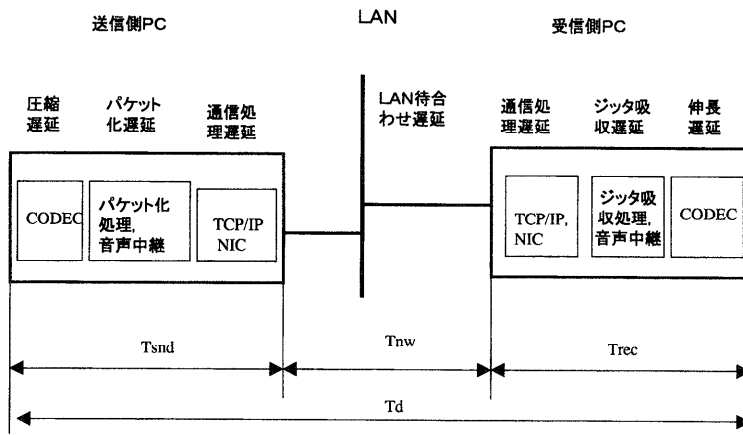


図 3 リアルタイム音声通信における遅延発生要因

Fig. 3 Delay in realtime voice communication.

### 2.3 低遅延リアルタイム音声通信の課題とアプローチ

LAN 環境での低遅延リアルタイム音声通信の課題を述べる。

#### (1) エンドツーエンドでのトータル遅延時間の短縮

図 3 に本音声通信システムにおける遅延とジッタの要因を示す。エンドツーエンドでの遅延時間は、送信側端末での遅延時間 ( $T_{snd}$ )、LAN での遅延時間 ( $T_{nw}$ ) および受信側端末での遅延時間 ( $T_{rec}$ ) が加算される。すなわち、全体の遅延時間  $T_d$  は、 $T_d = T_{snd} + T_{nw} + T_{rec}$  である。

送信側端末における遅延 ( $T_{snd}$ ) の 2 大要因は、CODEC における符号化遅延、およびデジタル化され

た音声を一定時間取り込んで蓄積してパケット化することによるパケット化遅延である。ここで、パケット化時間は G.729A などのフレーム型 CODEC において、フレーム化時間と同じか整数倍となる。PCM/ADPCM 方式ではパケット化時間に関して方式上の制限はない。遅延時間はパケット化間隔を縮めることで短縮できるが、逆にパケット数が増大して端末の処理負荷とネットワーク負荷が増えてしまう。したがって、許容できるパケット化遅延時間の範囲でパケット化間隔は長くすることが望ましい。

受信側端末での遅延 ( $T_{rec}$ ) の大きな要因は、ジッタ吸収バッファでの遅延、CODEC での復号化遅延である。前者については、(2) で述べる。CODEC に対

する考え方は、送信側端末と同様である。

LANにおける遅延 ( $T_{nw}$ ) の主要因はアクセス待ち遅延である。これが、ジッタの要因の1つとなる。これらの対応について(2)で述べる。

### (2) ジッタおよび中・長期で変動する遅延への対応

LANで発生するジッタ、すなわち瞬時的な遅延変動に対して、ジッタ吸収を行い会話中の音声の途切れを発生しないようにする制御が必要である。さらに、1日の昼と夜での遅延変動など、中・長期での遅延時間の変動が発生する。平均遅延が大きいときと小さいときをともにカバーでき、かつジッタを吸収する遅延制御を行うことが望ましい。

### (3) パケットの廃棄

到着した音声パケットをすべて取り込むためには、大幅に遅れてきたパケットが到着するのを待つ必要が生じて遅延時間が増大する。5%程度のパケットロスが発生しても会話に支障はないという特質を利用して、大幅に遅れてきた音声パケットを積極的に廃棄することにより遅延時間の短縮を図る。このために音声パケットの到着率の累計を求め、一定の到着率に達した以降に到着した音声パケットを廃棄する。

### (4) 同期

半2重通信の場合は、送話者と受話者が切り替えながら会話を行うので同期制御が必要となる。全二重通信方式にし、エンドツーエンドで双方向の音声チャンネルをつねにアクティブにしておくことで、遅延の大きい場合でも送受話者間での会話で補うことで同期の課題を解決する。

## 3. 低遅延リアルタイム音声通信システムの実現

### 3.1 設計方針

端末制御のみでエンドツーエンドのジッタを吸収し、エンドツーエンドでの遅延時間をITU-T G.114 勧告のガイドライン 150ms を十分満たす値以下に抑える。これを実現するために以下の方針でシステムを設計する。

- (1) PC内での遅延を抑えるために、ハード CODEC を採用する。CODEC にはアルゴリズム遅延、処理遅延が小さく、音質が良いPCM/ADPCM 方式とする。
- (2) 到着音声パケットの遅延変動からジッタ吸収制御を行う負荷適応制御方式の実現を図る。これによりジッタ吸収を図るとともに、中・長時間遅延変動への対応を図り、遅延時間が少なく、かつ音切れがないリアルタイム音声通信システム

を実現する。

### 3.2 負荷適応制御方式

#### (1) 方式の概要と特徴

一定のパケットロスが許容できるという特質を活かし、大幅に遅れてきたパケットを廃棄することによりエンドツーエンド遅延を削減する。これを定量化するために許容パケット廃棄率を新たに定義する。音声パケットの到着遅延分布は負の指数分布を仮定する<sup>17)</sup>。受信側で音声パケットの到着間隔を定期的に計測し平均遅延時間を求める。平均遅延時間と許容パケット廃棄率からジッタ吸収時間を決定する。負の指数分布の仮定から、遅延の分散と平均が同じになることの性質を利用することにより求めることが可能となる。ジッタ吸収時間からジッタ吸収バッファサイズを決定する。これにより許容パケット廃棄率を保って音声パケットが受信されスムーズな音声再生ができる。この操作を数秒間隔で一定時間ごとに繰り返し行う。これにより、中・長期的遅延変動にも適応した制御が行われ、その時点で達成できる最小遅延時間で通信を行うことが可能となる。さらに受信側での局所的な計測のみでシステム全体のエンドツーエンドの遅延時間変動が把握できる点が特徴である。

#### (2) アルゴリズム

図4に音声通信の流れを示す。送信側端末で、音声はデジタル化、符号化されたあと一定長にパケット化され、順次LAN上に送出される。LANで転送された音声パケットは受信側端末で順次受信され、ジッタ吸収バッファを経てCODECへ送られ復号化され再生される。ここで、受信側ではジッタ吸収バッファを設け、設定されたジッタ吸収時間遅延させてから音声を再生する。また、これを越えて遅れてきた音声パケットは廃棄する。

以下、アルゴリズムの概要を示す。

図5は、パケット受信間隔の計測点を示している。音声パケットの到着間隔はポアソン分布に従うと仮定しており、したがって、遅延時間の分布は負の指数分布に従う。平均遅延時間を  $D_e$  とすると、音声パケットの遅延時間の確率分布関数  $F(D)$  は次式で定義される。

$$F(D) = 1 - e^{-D/D_e} \quad (1)$$

ジッタ吸収時間を  $D_a$  とすると、パケット廃棄率  $k$ 、平均遅延時間  $D_e$  間には次式が成り立つ。

$$D_a = -\ln k \times D_e \quad (2)$$

したがって、平均遅延時間  $D_e$  を求めることができれば、ジッタ吸収時間の決定が可能となる。

送信側では一定間隔  $\Delta T$  で音声パケットを送信して

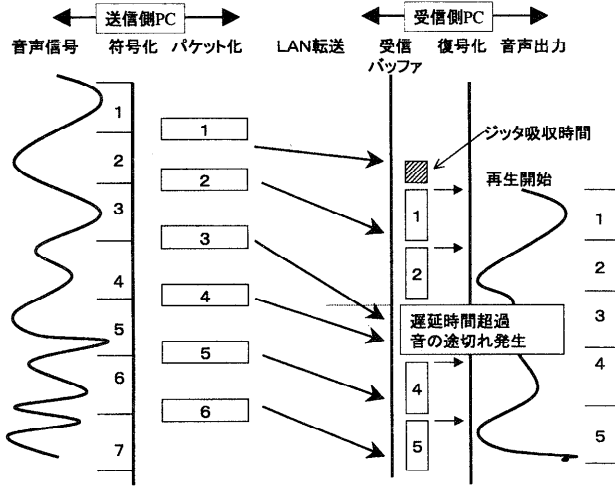


図4 音声パケット通信  
Fig. 4 Voice packet communication method.

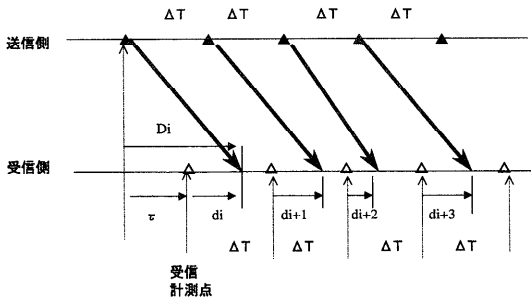


図5 受信音声パケットの遅延ゆらぎ時間計測  
Fig. 5 Jitter measurement of arrival packets.

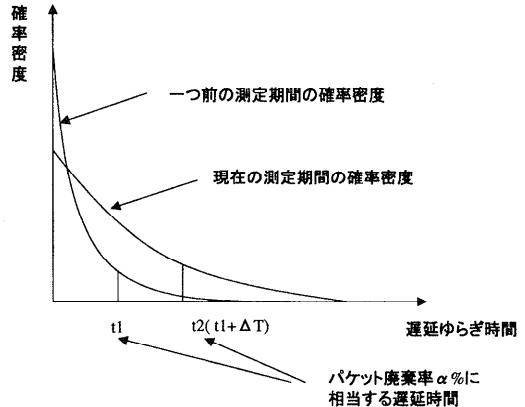


図6 受信音声パケットの遅延ゆらぎ分布とジッタ吸収時間の計算  
Fig. 6 Probability density of packet arrivals vs. jitter.

いる。しかし、送信側と受信側の時刻を一致させることは一般的に難しいため、受信側では音声パケット遅延を直接観測することができない。そこで、受信側では、送信間隔と同じ間隔  $\Delta T$  のパケット受信計測点を基準とし、各音声パケットの受信時刻の基準点からの変位  $d_i$  を計測する。各パケットの送信側時刻と受信側基準点時刻の差を  $\tau$  とすると、実際の各音声パケットの遅延時間  $D_i$  とすると、 $D_i$  は次式で表される。

$$D_i = \tau + d_i \quad (3)$$

ここで、遅延時間は負の指数分布と仮定しているから、音声パケットの平均遅延時間  $D_e$  と遅延時間の標準偏差  $\sigma(D_i)$  が等しい。

$$D_e = \sigma(D_i) = \sigma(d_i + \tau) = \sigma(d_i) \quad (4)$$

式(4)より、受信側での観測のみで、音声パケットの平均遅延時間が求められることが示された。これを、式(2)に代入することにより音声パケット廃棄率を  $k$  としたときのジッタ吸収時間を求められる。

図6は、ジッタ吸収時間値を示している。ここで

は、パケット廃棄率  $\alpha$  に相当するジッタ吸収時間が、1つ前の計測時刻で  $t_1$  だったものが次の計測時刻で  $t_2(t_1 + \Delta T)$  になったことを示している。

### 3.3 ジッタ吸収時間の再設定

ジッタ吸収時間は数秒単位の周期で再計算しジッタ吸収バッファサイズを再設定する。再設定時に遅延が減少傾向にあると、図7(a)に示すように、ジッタ吸収時間は  $t_1$  から  $t_2$  へと減少することになり、ジッタ吸収バッファサイズを減少させる。時間短縮に対応して、音声パケット4を廃棄するなどして調整する。再設定時に遅延が増大傾向にあると図7(b)に示すようにジッタ吸収時間は  $t_1$  から  $t_2$  へと増大することになり、ジッタ吸収バッファを増加させる。アンダーランに対応して、ホワイトノイズなどの補完パケットを挿入する。以上のようにして、ジッタ吸収時間変更

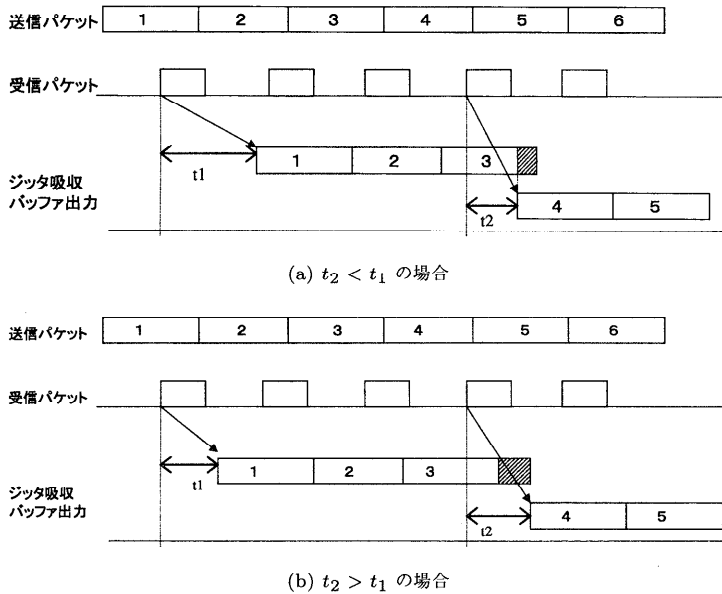


図7 ジッタ吸収時間の再設定  
Fig.7 Jitter absorption time set up.

表2 CODECの比較  
Table 2 Comparison of CODECs.

CODEC 種別	PCM (G.711)	ADPCM (G.726)	CS-ACELP (G.729A)	MPC-MQL (G.723.1)	ACELP (G.723.1)	備考
速度	64 kbps	32/16 kbps	8 kbps	6.3 kbps	5.3 kbps	
遅延	0.1 ms	0.1 ms	25 ms	67.5 ms	67.5 ms	送信側+受信側

出展: *IEEE Communication Magazine*, pp.40-47, September, 1997.

従した音声再生が可能となる。

### 3.4 中・長期遅延変動への適応制御

3.3節で示したジッタ吸収時間の再設定方式は、一方で中・長期遅延変動への対応ができる。すなわち、遅延が中・長期的に減少または増加傾向にあるときでも、ジッタ吸収時間を一定間隔で再計算して、ジッタ吸収バッファサイズの再設定を行い続ける。これにより、網の遅延負荷変動に適応したジッタ吸収制御が可能になる。

### 3.5 音声符号化復号化方式

表2に使用帯域と遅延時間の観点からCODECを比較したものを示す。アルゴリズム遅延を1ms以下にでき、他の高圧縮符号化方式より処理遅延が大幅に小さく、音声品質の良いPCM/ADPCM方式を採用する。LAN環境での使用を前提としているので帯域はインターネットに比べて潤沢に確保できるので、帯域削減よりも遅延削減と音声本質向上を狙う。圧縮レートはPCMで64kbps, ADPCMで32kbps, または16kbpsである。環境によって選択可とする。さらにCODECは専用ハードウェアにすることにより、PC

本体のCPU負荷の低減を図り処理遅延を削減する。

### 3.6 音声パケット長

音声のパケット化間隔は短くすると遅延時間は短縮されるがパケット数が増大して端末の処理負荷、ネットワーク負荷がともに増大してしまい、かつパケット長が短くなるのでプロトコルヘッダのオーバーヘッドが増大しネットワーク利用効率が低下する。逆に音声のパケット化間隔を長くするとPCの処理負荷、ネットワーク負荷、オーバーヘッドは減少するが、パケット化遅延時間が増大する。これらを考慮して、本システムではパケット化間隔を実験的に決定して50msにしている。これにより端末では送受信で毎秒20パケットずつの音声パケット送受信を行う。すなわち音声パケット長はヘッダを除き、64kbps時40バイト、32kbps圧縮時20バイト、16kbps時10バイトとなる。

### 3.7 低遅延音声通信システム

提案方式の有用性を検証するために低遅延音声通信システムを試作した。表3に試作システムの諸元を示す。



表3 試作リアルタイム音声通信システムの仕様

Table 3 Specification of LAN based realtime voice communication prototype system.

項目	仕様
端末	PC
OS	Windows <sup>☆</sup>
LAN	Ethernet <sup>☆☆</sup>
通信プロトコル	制御: TCP/IP 音声: UDP/IP
音声符号化復号化	ADPCM (32 kbps, 16 kbps) PCM (64 kbps)
音声パケット化間隔	50 ms
ジッタ吸収制御	負荷適応制御

#### 4. 低遅延音声通信システムの評価

試作システムを、音声パケット廃棄率、LAN 負荷変動の影響、エンドツーエンド遅延時間の観点で実験的に評価する。

##### 4.1 音声パケット廃棄率

5%の音声パケット廃棄率が会話に影響がないということの妥当性を実験的に検証し設計の指針とするために、試作システムにより実際の会話を行うことにより主観的な評価を実施する。音声は 32 kbps の ADPCM/PCM 圧縮を用い、音声パケット化間隔は 50 ms で行っている。研究者 20 名を被験者として、強制的に音声パケットを廃棄していくことによる音声品質への影響をみるために、音質、音の途切れ、遅延などの内容を含む通常の会話を行い主観的に評価してもらった。音声パケットの廃棄パターンとして以下の 3 パターンを用いた。1 パケットの廃棄は 50 ms の音声瞬断になる。

- (a) 一定間隔ごとに 1 パケット廃棄
- (b) 一定間隔ごとにバースト的にパケットを廃棄 (2 パケットおよび 5 パケット連続廃棄)
- (c) ランダムにパケットを廃棄

評価結果を表 4 に示す。

廃棄率が 5% 以下の場合、廃棄パターンにかかわらず音の途切れはほとんど気にならない。

廃棄率が 5% から 10% の間では廃棄パターンにより評価が異なる。5 パケット連続廃棄といったバースト的なパケット廃棄の場合、音の途切れが非常に気になるが、他の場合は音の途切れは気にならないという結果となった。廃棄率が 10% 以上になると、すべての廃棄パターンにおいて音の途切れが非常に気になるという結果となった。この結果は文献 18) の結果とほぼ一

表4 パケット廃棄率による音質評価

Table 4 Evaluation of voice quality vs. packet loss rate.

廃棄パターン	廃棄率			
	2%	2~5%	5~10%	10%以上
一定間隔で廃棄 (1 パケット)	○	○	△	×
一定間隔で廃棄 (2 パケット連続)	○	○	△	×
一定間隔で廃棄 (5 パケット連続)	○	○	×	×
ランダムに廃棄	○	○	△	×

- : 音の途切れにほとんど気が付かない
- △: 音の途切れが気にならない
- ×: 会話に支障をきたす

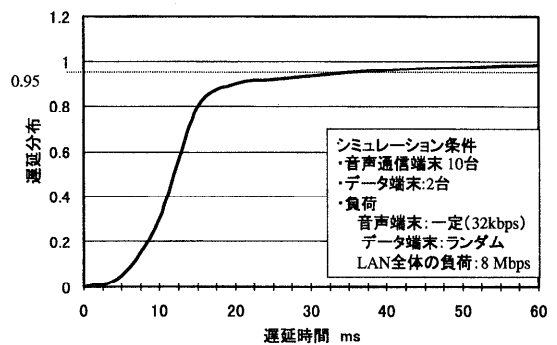


図8 遅延時間分布シミュレーション

Fig. 8 Probability distribution of delay by simulation.

致することが実験的に検証された。

##### 4.2 LAN 高負荷時の音質への影響と遅延時間の関係

LAN の負荷増大による音声品質への影響を調べるために、LAN にトラフィックジェネレータを接続し、擬似負荷トラフィックをかけた状態で実際に会話を行い、擬似負荷トラフィックを変動させて音声品質の評価を行った。音声は 32 kbps ADPCM で圧縮を行っている。LAN 負荷が高くなるのはファイル転送が行われたときと想定し、パケット長は Ethernet の最大フレーム長 (1500 byte) としている。また、ジッタ吸収時間は 50 ms に固定設定している。実験の結果、擬似負荷トラフィックが 80% 程度 (8 Mbps) までは音切れは識別できず会話に影響は出なかった。しかし、擬似負荷トラフィックが 90% 近くになるにつれパケット廃棄による音切れが急速に増大し、会話が難しい状態になった。

このような LAN が高負荷時の通信遅延時間への影響について、シミュレーションによる評価をあわせて行った。図 8 に負荷が 80% の高負荷時のシミュレーション結果を示す。負荷条件は、音声端末 10 台 (5 通

☆ Windows は、米国およびその他の国における米国 Microsoft Corp. の登録商標です。

☆☆ Ethernet は、米 Xerox Corp. の商品名称です。

話相当)をそれぞれ 32 kbps で通信し、データ端末 2 台間でランダムトラフィックを流す。シミュレーションは GPSS を用いている。図 8 の結果から、パケット廃棄率 5% (到着率 95%) に対応する通信遅延時間は 40 ms である、したがって、50 ms のジッタ吸収時間を持たせれば、95%以上のパケットは到着することが示されており、上記実験結果が傍証される。実際のネットワーク設計では、このような高負荷で LAN を用いることはなく、一般に 50%以下で使用されると考えられるので、音声パケット廃棄率は 5%を超えることはほとんどないと想定される。

#### 4.3 エンドツーエンドの遅延時間

上記の評価において、LAN 負荷が高い場合でも、ジッタ吸収時間を 50 ms 程度に設定すればパケット廃棄率が 5%以下になり、音声の途切れはほとんど気にならないことが判明したので、このような条件でエンドツーエンドでの遅延時間の影響につき実験評価した。実験システムではパケット化間隔を 50 ms に設定している。ジッタ吸収時間を 50 ms 程度 (LAN が高負荷時) とすると、エンドツーエンドの遅延は 100 ms 程度となる。本システムを 100 名の被験者に実際に使用してもらい主観評価を行った。その結果、90%の被験者が遅延は気にならずスムーズな会話ができたと報告している。なお、実際の LAN での負荷トラフィックは一般に 50%以下で使用されており、ジッタ吸収時間はこれに対応して小さな値に自動的に設定さる。たとえば、前節と同様なシミュレーションにより、LAN 利用率 40%時で、ジッタ吸収時間は 20 ms 程度となり、エンドツーエンドの遅延は 70 ms 程度にさらに減少する。

## 5. む す び

LAN 環境における音声通信システムを対象に、遅延時間の短縮、ジッタの吸収、パケットロス、データ通信との共存方式の検討を行い、パケットの到着間隔を測定することによりジッタ吸収時間を決定することができる負荷適応制御方式および、ハードウェア CODEC を採用し、TCP (UDP)/IP 通信を活用して音声通信を行う低遅延リアルタイム音声通信システムを提案した。

提案した音声通信システムの有用性を把握するために複数の PC と LAN などから成るシステムを開発し、実験評価をシミュレーションとあわせて実施した。実験の結果、遅延時間、音質に関して次のように当初の結果が得られていることが明らかになった。

(1) 実験評価より提案システムにおいてパケット廃

棄率が 5%以下であれば音の途切れにほとんど気づかず会話に支障をきたさない。

- (2) 負荷適応制御の有効性評価をシミュレーションと主観評価を行い、実装した負荷適応制御方式によるジッタ吸収時間値の妥当性を確認した。
- (3) LAN が高負荷時 (80%利用率, ジッタ吸収時間 50 ms, エンドツーエンド遅延約 100 ms) での評価実験において 100 名の被験者の 9 割が会話に支障がないと報告された。
- (4) 上記により、エンドツーエンド遅延時間を 100 ms 以下とする低遅延リアルタイム音声通信システムが実現できることを示した。
- (5) 一般的な LAN の使用状態において (利用率 40%時), ジッタ吸収時間は約 20 ms となり、エンドツーエンド遅延は約 70 ms になり、リアルタイム性はさらに向上する。

提案した低遅延音声通信方式を実システムに適用して問題なく稼働している。今後、広域 IP ネットワークへの適用拡大の研究を進める。

謝辞 本研究を進めるにあたり、ご指導いただいている (株) 日立製作所システム開発研究所片岡雅憲所長、同システム開発研究所主管研究長坂東忠秋博士、中根啓一部長に感謝いたします。

## 参 考 文 献

- 1) Zellweger, H.M., Swinehart, D.C. and Rangan, P.V.: Multimedia Conferencing in the Ethernet Environment, *IEEE Computer*, Vol.24, No.10, pp.69-79 (1991).
- 2) Press, L.: Net Speech: Audio Comes to the Net, *Comm. ACM*, Vol.38, No.10, pp.25-31 (1995).
- 3) Clark, D.D., Shenker, S. and Zhang, L.: Supporting Real-Time Application in an Integrated Services Packet Network: Architecture and Mechanism, *Proc. SIGCOMM'92* (1992).
- 4) ITU-T Recommendation H.323: Packet Based Multimedia Communication Systems (1998).
- 5) IMTC Voice over IP Forum: IMTC Voice over IP Forum Service Interoperability Implementation Agreement 1.0 (1998).
- 6) ITU-T Recommendation G.114: One Way Transmission Time (1996).
- 7) Oran, D.: Voice Quality: An End-to-End Problem, *Voice on the Net (VON) conference records*, September 26, Boston (1997).
- 8) Cox, R.V.: Three New Speech Coders from ITU Cover a Range of Applications, *IEEE Communications Magazine*, Vol.35, No.9, pp.

- 40-47 (1997).
- 9) Lefelhocz, C., Lyles, B., Schenker, S. and Zhang, L.: Congestion Control for Best-Effort Service, *IEEE Network Magazine*, January/February pp.10-19 (1996).
  - 10) Blake, S., Black, D., Carlson, M., Davies, E., Wng, Z and Weiss, W.: An Architecture for Differentiated Services, IETF RFC2475, December (1998).
  - 11) Braden, R., Zhang, L., Berson, S., Herzog, S. and Jamin, S.: Resource ReSerVation Protocol (RSVP), IETF RFC2205, September (1997).
  - 12) Privalov, A. and Sohraby, K.: Per-Stream Jitter Analysis in CBR ATM Multiplexors, *IEEE/ACM Trans. Netwroking*, Vol.6, No.2, pp.141-149 (1998).
  - 13) Matragi, W., Sohraby, A. and Bisdikian, C.: Jitter Calculus in ATM Netwroks: Multiple Nodes, *IEEE/ACM Trans. Networking*, Vol.5, No.1, pp.123-133 (1997).
  - 14) Cidon, I., Khamisky, A. and Sidi, M.: Dispersed Messeges in Discrete-Time Queues: Delay, Jitter and Threshold Crossing, *Proc. IEEE INFOCOM*, Toronto, Canada, pp.903-907 (1994).
  - 15) ATM Forum: Traffic Management Specification Version 4.0 af-tm-0056.000 (1996).
  - 16) Kleinrock, L.: Queueing Systems Vol.2, *Computer Applications*, John Wiley and Sons, (1976).
  - 17) Bolot, J.: Characterizing End-to-End Packet Delay and Loss in the Internet, *Proc. SIGCOMM'93*, pp.289-298 (1993).
  - 18) 鈴木三知男, 櫻尾次郎, 土岐隆一, 加藤孝雄: 音声パケット交換システムにおける遅延適応制御方式の提案, *電子情報通信学会論文誌*, Vol.J66-B, No.7, pp.893-899 (1983).
  - 19) Ferrari, D.: Delay jitter control scheme for packet-switching netwroks, *Computer Communications*, Vol.15, No.6, pp.367-373 (1992).
  - 20) Vin, H.M., Aellweger, P.T., Swinehart, D.C. and Rangan, V.: Multimedia Conferencing in the Etherphone Environment, *IEEE Computer*, Vol.24, No.10, pp.69-79 (1991).
  - 21) Bolot, J. and Vega-Garcia, A.: Control Mechanism for Packet Audio in the Internet, *IEEE INFOCOM'96*, San Francisco, pp.232-239 (1996).
  - 22) Kostas, T.J., Borella, M.S., Sidhu, I., Schster, G.M., Grabiec, J. and Mahler, J.: Real-Time Voice Over Packet-Switched Networks, *IEEE Network*, pp.18-27 (1998).
  - 23) 山崎哲朗, 井合 知, 住本純一: インターネット電話の受聴品質測定結果, 信学会総合大会, SB-11-6 (1998).
  - 24) Jayant, N.S.: Effect of Packet Losses in Waveform-coded speech, *Proc. ICC*, pp.275-280 (1980).
  - 25) 星 徹, 松井 進, 高田 治, 岩見直子, 高原桂子: リアルタイム・蓄積統合マルチメディアコミュニケーションシステム, 情報処理学会研究会報告, GW8-5 (1994).
  - 26) Hoshi, T., Mori, K., Takahashi, Y., Nakayama, Y. and Ishizaki, T.: B-ISDN Multimedia Communication and Collaboration Platform Using Advanced Video Workstations to Support Cooperative Work, *IEEE Journal on Selected Areas in Communications*, Vol.10, No.9, pp.1403-1412 (1992).

(平成 10 年 9 月 11 日受付)

(平成 11 年 4 月 1 日採録)



星 徹 (正会員)

1969年東京工業大学電気工学科卒業。同年(株)日立製作所入社。1975年カリフォルニア大学ロサンゼルス校(UCLA)大学院コンピュータサイエンス学科修了。交換システム, マルチメディア LAN, CSCW, CTI, IP テレフォニ等の研究に従事。現在同社システム開発研究所第4部主任研究員。IEEE, 電子情報通信学会各会員。



谷川 桂子 (正会員)

1991年宇都宮大学情報工学科卒業。同年(株)日立製作所入社。マルチメディア通信, IP テレフォニの研究に従事。現在同社システム開発研究所第4部研究員。



松井 進 (正会員)

1980年大阪大学大学院基礎工学研究科物理系修了。同年(株)日立製作所入社。LAN, IP テレフォニ, モバイルコンピューティング等の研究に従事。現在, 同社システム開発研究者第4部主任研究員。IEEE, 電子情報通信学会各会員。



岩見 直子（正会員）

1983年東京女子大学数理学科卒業。同年（株）日立製作所入社。LAN, CSCW, IPテレフォニ, Webサーバ等の研究に従事。現在同社システム開発研究所第6部研究員。



寺田 松昭（正会員）

1970年岡山大学工学部電気工学科卒業。同年（株）日立製作所入社。同社システム開発研究所において、制御用分散処理システム, LAN, プロトコル高速処理, VOIP, 次世代インターネットの研究に従事。工学博士。著書「制御用計算機におけるリアルタイム技術」（共著, コロナ社）, 「デジタルサービス革命」（共著, 日刊工業新聞社）。1991年4月より東京農工大学工学部情報コミュニケーション工学科教授。IEEE, ACM, 電子情報通信学会各会員。

---