

# IP 電話の通話品質評価と既存ソフトウェア電話の改良

2001MT025 平岡 翼

2001MT068 森岡 真生

指導教員 後藤 邦夫

## 1 はじめに

近年、VoIP 技術により音声を IP パケットに変換しインターネット網を使って送受信する IP 電話が大企業を中心に導入が進んでいる。IP 電話にはコスト削減の他にコンピュータと連携して高度な機能を実現できるという利点があり、これからは中小企業や一般家庭のユーザにも本格的に普及すると思われる。

しかし独自の IP 網ではなくインターネット全体を使うインターネット電話サービスでは通話品質がネットワークの状態に左右されるため音声途切れて聞き取りにくいなど、品質が保証されない。

そこで本研究では IP 電話による音声通信のパケットに遅延、損失などの障害を故意に与え、通話品質が実験ネットワークでどの程度の影響を受けるか評価する。

現在の IP 電話の呼制御プロトコルには H.323 と SIP(Session Initiation Protocol) の二つの系統があるが、本研究では電話回線を用いないインターネットの範囲で主に使用される SIP(RFC3261) を使用する。

SIP サーバとして VOCAL[1], IP 電話機能を持つルータと普通の電話機、電話機能を備えた PC 数台を使い実験ネットワークを構築し、通話品質を評価する。また、今後は PC を IP 電話機端末として使用するソフトウェア電話が一般に普及すると予想されるので、ソフトウェア電話として VOCAL に含まれる SIPset を使用する。しかし、本研究で使用した PC では SIPset の通話が不安定であった。そこでソフトウェア改良も行う。

平岡翼は主に品質評価を担当し、森岡真生はソフトウェア改良を担当した。

## 2 実験の環境

この節では本研究における実験環境について説明する。

### 2.1 音質劣化原因と発生箇所

IP ネットワークでは、音声のほかにアプリケーションなどのデータも伝送される。そのため、音声パケットが遅れて届いたり、あるいはネットワークの途中で廃棄されたりして音質が劣化する。場合によっては、電話が繋がらないということもある。

「パケット損失」「遅延」「ジッタ(遅れのゆらぎ)」が音質劣化の主な要因となる [2]。

### 2.2 実験環境

IP 電話は実時間通信であるため、リアルタイムアプリケーションのための転送プロトコルとして RTP(Real-time Transport Protocol) が使われている。SIP を使った IP 電話サービスでは SIP によってセッションの設定

を行い、端末同士がいったんつながると SIP サーバを介することなくユーザ同士での RTP(RFC1889 参照)、RTCP での通信となる。上記を図 1 に示す。

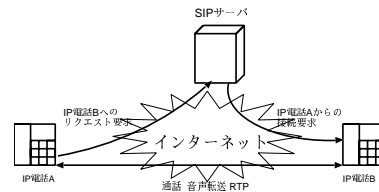


図 1 SIP を用いた IP 電話サービス

### ネットワークモデル

本研究では図 2 のような実験ネットワークを構築した。UA(ユーザー・エージェント)とは一般には IP 電話端末やソフトウェア電話である。SIP サーバにはオープンソースソフトウェアである VOCAL を用いた。また UA 同士のセッションが成功すると SIP サーバの PC は障害を発生させるルータとしても用いる。

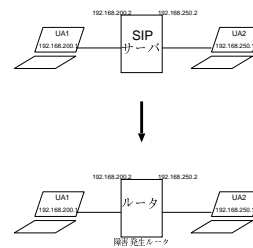


図 2 ネットワークモデル

使用機器、ソフトウェアは以下のとおりである。

- 使用 UA :
  1. YAMAHA ルータ RT57i + 電話機 (Pioneer TF-10-s)
  2. ノート PC (CPU : Intel Pentium 3 500MHz, メモリ : 192MB RAM, OS : VineLinux2.4.20, サウンドカード : intel8x0, ドライバ : ALSA-1.0.5) Linux カーネルに含まれる OSSFree は大抵のカードで入出力が同時にできないので、ALSA を使用
  3. SIPset : sip-1.5.0, Linphone : linphone-0.12.1
- SIP proxy / NISTNet ルータ : FMV (CPU : GenuineIntel 500MHz, メモリ : 320MB RAM, OS : VineLinux2.4.22)  
VOCAL : VOCAL-1.5.0

## NISTNet

本研究では NISTNet[3] というツールを使って故意に障害を起こした。NISTNet とはホスト間に流れるパケットに擬似的な通信障害を発生させるネットワークエミュレータである。NISTNet の特徴は、Linux のカーネルモジュールの拡張により実現しているので高速であり、X Window を用いた対話型の GUI インタフェースがあるので使いやすいことである。NISTNet で設定できる疑似障害を以下に示す。

表 1 NIST Net で実現できるパラメータ

パラメータ	意味	範囲
Delay	遅延時間	0 ~ 2,000,000 ms
4 Delsigma	遅延変動幅	0 ~ 2,000,000 ms
BandWidth	最大帯域幅	Bytes/Second
Drop	ドロップ率	0 ~ 100(%)
Dup	パケット重複率	0 ~ 100(%)

## 2.3 実験方法

本実験は会話ではなく、一方向からサンプル音声を流し評価する。実験手順は以下の通りである。

1. SIP サーバ用 PC 兼ルータで NISTNet を使用して UA1 と UA2 の間の双方の UDP 通信に遅延、UA1 から UA2 へパケット損失、パケット重複を設定する。
2. UA1 を録音したサンプル音声を再生するように設定する。
3. UA2 から UA1 に SIP サーバを介して電話をかけ、UA2 で受信したサンプル音声を複数の被験者が評価する。

UA2 には改良を加えた SIPset、UA1 には sipomatic を用いる。ルータで起こす障害はパケット損失、遅延、パケット損失 + 遅延、遅延 + パケット重複とする。これらの障害は IP 電話の品質クラス分類 (表 3) の評価基準に使用され、また 2.1 節で述べたように音質劣化の主な原因となっている。

実験で被験者に聞いてもらうサンプル音声は、同じでなければならないので事前に録音し、Linphone[4] 付属の自動応答プログラム sipomatic で再生する。本実験では mu-law 8000Hz サンプリングの音声データを使用する。サンプル音声は 22 歳男性が読み上げた約 1 分間のニュース記事である。

音声は、それぞれの障害を起こした状態で流す。被験者は各障害が起きている状態で音声を聞き、評価する。音声は約 1 分間なので評価が済み次第、障害パラメータを変更する。

## 3 ソフトウェア電話の改良

この節では、ソフトウェア電話の音声劣化の原因と改良点を説明する。ソフトウェア電話は、PC を電話機端

末として機能させるプログラムである。本研究で使った SIPset は音飛びが激しく、安定していない。そこで 3.3 節と 3.4 節で tcpdump コマンドを用いた分析と不安定の原因を示す。3.5 節では改良点とその結果を示す。

### 3.1 ソフトウェア電話

本研究では SIP サーバを使用するため、SIP が利用できる SIPset と Linphone を使用する。SIPset は GUI が使用でき、使いやすいが音飛びが激しく、通話が安定しない。Linphone は linphonec しか作れず、GUI が使用できない。そのため使いにくい、通話が安定している。他のソフトウェア電話は SIP に対応していない、Linux で使用できないなどの理由で用いない。

### 3.2 音声劣化原因の推測

Linphone 同士の通話は途切れることがなく、安定していた。そこで Linphone と SIPset での通話を試してみたところ、Linphone から送られた音声は SIPset で正常に聞き取れたのに対し、SIPset から送られた音声は Linphone でも音飛びが激しく、安定していなかった。この症状から、SIPset では通話時に使用されるパケットの送受信タイミングに問題があると推測した。

### 3.3 分析

RTP パケットが的確なタイミングで送られているか確認する。一般に IP 電話は 20 ミリ秒分、8bit 8000Hz の音声データを 160byte ずつ RTP ペイロードにして送信する。そこで、SIPset、Linphone と IP 電話の RTP タイムスタンプ (符号化時のシーケンス番号や時刻情報) も確認する。SIPset の通話時の RTP パケットの流れを調べるため tcpdump コマンドを用いて計測した。安定した通話ができる Linphone のパケットの流れを調べ、SIPset と比べることによって SIPset の音声劣化原因を分析した。以下に Linphone の結果を示す。

```
13:43:56.713969 01m068.G406.seto-private.7080 > 01m025.G406.seto-private.7078: udp/rtp 160 c0 25 4160 (DF)
13:43:56.753469 01m025.G406.seto-private.7080 > 01m068.G406.seto-private.7078: udp/rtp 160 c0 25 4320 (DF)
13:43:56.783671 01m068.G406.seto-private.7080 > 01m025.G406.seto-private.7078: udp/rtp 160 c0 27 4480 (DF)
13:43:56.773516 01m025.G406.seto-private.7080 > 01m068.G406.seto-private.7078: udp/rtp 160 c0 52 8480 (DF)
13:43:56.773592 01m025.G406.seto-private.7080 > 01m068.G406.seto-private.7078: udp/rtp 160 c0 53 8640 (DF)
13:43:56.783673 01m068.G406.seto-private.7080 > 01m025.G406.seto-private.7078: udp/rtp 160 c0 28 4640 (DF)
13:43:56.809965 01m068.G406.seto-private.7080 > 01m025.G406.seto-private.7078: udp/rtp 160 c0 29 4800 (DF)
13:43:56.810039 01m068.G406.seto-private.7080 > 01m025.G406.seto-private.7078: udp/rtp 160 c0 30 4960 (DF)
13:43:56.813528 01m025.G406.seto-private.7080 > 01m068.G406.seto-private.7078: udp/rtp 160 c0 31 5120 (DF)
13:43:56.813606 01m025.G406.seto-private.7080 > 01m068.G406.seto-private.7078: udp/rtp 160 c0 54 8800 (DF)
```

図 3 Linphone の tcpdump の結果

図 3 から Linphone のパケットは送信と受信がほぼ交互に行われていることが分かる。また RTP タイムスタンプは 1 ずつ増加し、tcpdump タイムスタンプはおよそ 20 ミリ秒ごとに送信されることが確認できた。160byte の音声データを 20 ミリ秒ごとに送信することが Linphone の安定性の理由と推察できる。

### 3.4 SIPset の音声劣化原因

SIPset の tcpdump の結果を以下に示す。

```
12:56:36.499473 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
12:56:36.501628 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
※12:56:36.973557 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
12:56:36.974074 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
12:56:36.974078 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
省略 ↓
12:56:36.982879 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
12:56:36.983027 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
※12:56:37.007897 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
12:56:37.008075 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
12:56:37.008233 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
```

図 4 SIPset の tcpdump の結果

図 4 から SIPset の RTP パケットは固め送りになっていることがわかる。また RTP タイムスタンプは 1 ずつ増加しているが、tcpdump タイムスタンプはおよそ 0.5 秒ごとに送信されている。これにより、受信側に届いても、遅すぎて再生に使用できないと推測される。このことが SIPset の音声劣化原因と考えられる。

### 3.5 改良点

3.4 節で推測された原因と考えられる音声送信ルーチンを LOG\_DEBUG\_STACK 出力で突き止めた。同時に SIPset 起動時に出力されるオーディオバッファサイズが大きいことに気づいた。最近のオーディオ装置とドライバは SIPSet が作られた 2002 年より高機能なので、default バッファサイズが 4096byte と大きすぎる。結果として 160byte の音声データが約 25 個 (0.5 秒分) ずつ固めて送信されている。

SIPSet は ALSA ドライバの OSS 互換機能しか使用していないので、OSS のプログラミングガイドでバッファサイズ調整方法 (ioctl) を調べ、SIPSet 内にバッファサイズを減らすコードを追加した。以下に修正したファイル (sip-1.5.0/libsoundcard/SoundCard.cxx) を示す。

```
12:56:36.499473 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
12:56:36.501628 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
※12:56:36.973557 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
12:56:36.974074 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
12:56:36.974078 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
省略 ↓
12:56:36.982879 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
12:56:36.983027 01mt068.G406.seto-private.10000 > 01mt025.G406.seto-private.10000: udp/rtp 160
※12:56:37.007897 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
12:56:37.008075 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
12:56:37.008233 01mt025.G406.seto-private.10000 > 01mt068.G406.seto-private.10000: udp/rtp 160
```

```
以下を 126 行目に追加した。
```

```
int flagment = 0x00140004;
if (ioctl (myFD, SNDCTL_DSP_SETFRAGMENT,
&flagment) == -1){
    perror ("SNDCTL_DSP_SETFRAGMENT");
    deviceMutex.unlock();
    myFD = -1;
    return -1; }
}
```

改良を加えた SIPset は、安定した通話が可能になる。改良前の SIPset は通話が途切れ途切れであったが、改良後は途切れることもなくなり、安定するようになった。SIPset は交互にパケットを送信できるようになり、また送信が 20 ミリ秒ごとに行われるようになった。

## 4 品質評価

IP 電話の通話品質評価には MOS(Mean Opinion Score), R 値, PSQM(Perceptual Speech Quality Measure), PESQ(Perceptual Evaluation of Speech

Quality)[5] が用いられる。

MOS は実際に人間が聞いて評価する方法であり、送信側に評価用音声サンプル音源を用意し、受信側で実験に参加する人 (被験者) が音を聞いて測定する。最終的な評価は、複数の人が 5 段階で評価した通話品質を平均値で表したものになる。MOS 評価のスコアは「楽に通話でき、努力を不要=5」から「努力しても、会話が理解できない=1」の 5 段階で表現され、スコアが高くなるほど音質が良い。

R 値は IP 電話番号の割り当て基準に使われる客観評価であり、音声を送る観点でネットワークの品質を表す数値である。ネットワークの品質に関する各種パラメータを計算式に代入し、音質の評価値を出す。揺らぎや損失に加え、遅延の影響も反映できる点が特徴である。

本実験では UA に改良した SIPset を使用し、MOS と R 値を用いて音声品質を評価する。RT57i を使用し高音質な IP 電話は実現できるが改良した SIPset は RT57i とかわらないので省略した。

### 4.1 主観評価 (MOS 評価)

実験ではパケット損失、遅延、パケットロス + 遅延、遅延 + パケット重複をそれぞれ起こした状態で行った。評価方法を以下に示す。

- 各通信障害を発生
- 一回の試験時間は約 1 分間
- 評価者は音質評価の非熟練者である 18 名
- 会話ではなく聞き取りのみ
- 5 段階尺度により評価

実験で行った MOS 評価の結果を以下に示す。

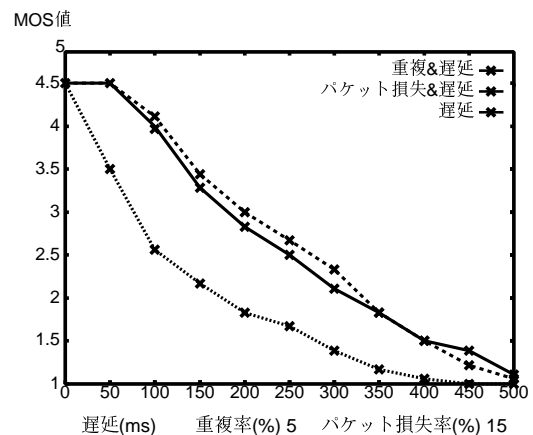


図 5 障害に対する MOS 値

図 5 の横軸は遅延 (ms) を表し、縦軸は評価値を示す。遅延、遅延 + 重複、パケット損失 + 遅延の 3 つでは、パケット損失 + 遅延が一番悪い結果になった。遅延と遅延 + 重複は 150 ミリ秒以下であれば MOS 値 3 を上回り、ほぼ同じ結果になった。このことから本実験で起こした複合障害では、遅延よりもパケット損失が品質劣化に大きく影響を与えていることがわかる。

#### 4.2 客観評価 (R 値)

伝送遅延時間やパケット損失以外にも音声品質劣化原因があるので、各種の音声劣化要因を総合的に考慮した音声品質の客観的評価方法が ITU-T 勧告 G.107[6] に記されている。

この評価方法は、雑音やエコーなどのネットワークの特性を表すパラメータをもとに計算して、通話品質を割り出したものである。E-model と呼ばれるアルゴリズム (計算式) に音声品質に関する回線の雑音や音量、エコーや遅延など、20 個のパラメータを入れて求める [2]。本来これらの値は測定器によって測定する必要があるが、一般に測定が困難なため、MOS 値から機器の影響を推定し、20 個のパラメータから R 値を計算するツール、emodel[7] を用いる。MOS の実験結果と遅延のパラメータを入力し、その他のパラメータは標準値を用いて R 値を導き出す。表 3 は現在、総務省が定めた IP 電話の品質クラス [2] である。「050」で始まる IP 電話の電話番号の割り当て基準にはクラス C 以上の品質が必要となる。emodel を用いた MOS 値から R 値への変換結果を表 2 に示す。

表 2 MOS 値の R 値への変換結果

遅延 (ms)	遅延		パケット損失 + 遅延		遅延 + 重複	
	MOS 値	R 値	MOS 値	R 値	MOS 値	R 値
50	4.5	93.8	3.5	71.5	4.56	93.3
100	4	87.3	2.56	47.7	4.11	92.0
150	3.28	64.4	2.17	38.1	3.44	68.7
200	2.83	52.9	1.83	29.1	3	57.0
250	2.5	44.8	1.67	24.4	2.67	48.7
300	2.11	35.1	1.39	15.7	2.33	40.3
350	1.83	27.7	1.17	7.5	1.83	27.7
400	1.5	18.3	1.06	2.4	1.5	18.3
450	1.39	14.5	1	0	1.22	8.6
500	1.11	3.9	1	0	1.06	1.7

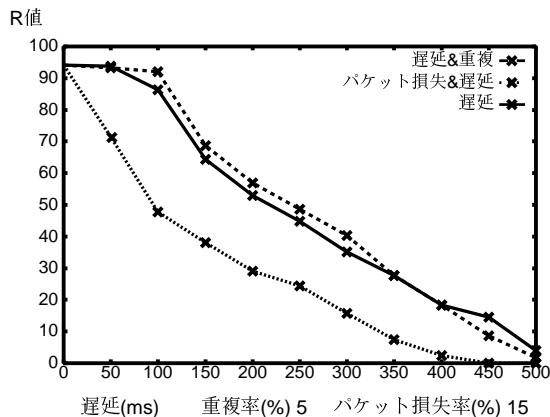


図 6 障害に対する R 値

図 6 は表 2 の R 値を示すが、この R 値は MOS から算出したので図 5 と同じ同様の結果になった。このこと

からツールを用いて R 値を算出する場合は、MOS 評価の高い精度が要求される。

表 3 IP 電話の品質クラス分類

	クラス A	クラス B	クラス C
R 値	> 80	> 70	> 50
遅延時間 (ミリ秒)	< 100	< 150	< 400
呼損率	≤ 0.15	≤ 0.15	≤ 0.15

#### 4.3 考察

MOS 評価と R 値から障害によって音声品質は大きな影響を受けることがわかった。図 5 と図 6 から遅延は 100ms を越えると品質に影響を与えることがわかった。また、パケット損失 + 遅延の場合はパケット損失の影響から遅延が少ない場合でも評価が低い。このことからパケット損失が音声品質に大きな影響を与えていることがわかる。しかし、遅延 + 重複と遅延の二つを比べるとあまり差がないことから重複はあまり影響を与えていないことがわかる。表 3 で定義されているクラスと比べると、パケット損失率 (%) 15 + 遅延 50ms 以下であればクラス C 以上なので、「050」で始まる IP 電話番号の割り当て基準に達し、実用域であると言える。

#### 5 おわりに

本研究では実験ネットワークを構築し、NISTNet を用いて通信障害を発生させ、インターネット網を介するインターネット電話と IP 電話の音声品質評価を行った。またソフトウェア改良を行うことによって SIPset を通話が安定して通話できるようにした。SIPset を用いて品質評価を行った結果、インターネット電話はネットワークの状態に影響を受けることがわかった。また、ソフトウェア電話はネットワークの状態がパケット損失率 (%)15、遅延 50ms 以下であれば携帯電話ほどの音声品質で通話ができるとわかった。

しかし、本研究で行った MOS 評価は同じ研究室の学生に聞いてもらったため、年齢が離れておらず、評価値の精度が高かったとは言えない。様々な年齢の被験者に協力してもらい評価値の精度を上げる必要がある。

また、SIPset に mpeg4 を組み込めば動画像も送信できるが、作るのが困難であり完成まではいかなかった。時間があれば試したかった。

#### 参考文献

- [1] Vocal : vovida.org, <http://vovida.org/>, 2002 .
- [2] 総務省 : IP ネットワーク技術に関する研究会報告書, <http://www.soumu.go.jp/s-news>, 2001 .
- [3] NISTNet:<http://www-x.antd.nist.gov/nistnet/>, 2002 .
- [4] Linphone : <http://www.linphone.org/>, 2004 .
- [5] 都丸敬介 : “IP 電話のすべて”, 電話新聞社, 2004 .
- [6] ITU-T : <http://www.itu.int/ITU-T/publications/recs.html>, 2003 .
- [7] Psytechnics : <http://www.psytechnics.com/>, 2004 .