

# 公衆無線 LAN におけるマルチレート特性を考慮した VoIP の音声品質改善アルゴリズムの評価

M2011MM073 田財 彰久

指導教員：奥村 康行

## 1 はじめに

近年、映像や音声サービスを IP ネットワークに統合する無線の LTE のサービス展開に伴い、固定電話回線の経済的代案として注目される無線の Voice over Internet Protocol(以下 VoIP)の商業展開が進んでいる。既に、LTE で利用可能な VoIP として、VoLTE のサービスが世界各国で商用段階に突入している。しかし、LTE 過渡期は従来の回線交換方式の通話が主流で、輻輳に弱く災害発生時の利用には向いていない。

よって、本研究では災害発生時のライフラインとなる公衆無線 LAN の利用を想定した環境をシミュレートし、LTE 過渡期や災害発生時にも通話可能な WLAN VoIP を研究対象とする。

当研究分野の先行研究の一例である“An Adaptive Codec and Frame Size Modification Based QoS Algorithm over Multi-rate WLANs”[1]において、VoIP を無線 LAN 環境で使用するさいの混雑を想定する音声品質改善アルゴリズムが考案された。しかし、先行研究のシミュレーションでは、無線ノードの電波減衰に伴う通信帯域変化、すなわちマルチレート特性の考慮が不十分である。さらに、E-model[2]と呼ばれる計算モデルを使用しており、音声品質を 0 から 100 で客観的評価する R 値を算出している。この音声評価手法から算出される R 値で主観品質を推定することは難しい。

そこで、本研究では先行研究よりも再現度の高いマルチレート環境を想定したネットワークをシミュレータ上で構築し、先行研究と同じく E-model で客観的評価する。そして、マルチレート環境による音声品質劣化要因を反映させた実音声サンプルを作成した上で、その音源を評価する PESQ ツール [3] を用いて主観品質を推定する。

2つの評価値を比較することで、先行研究の音声品質改善アルゴリズムに対して、より信頼性の高い評価をすることを本研究の目的とする。

## 2 シミュレーションモデル

本節では VoIP シミュレーションの対象となるモデルと、次節から取り扱う 2つのシミュレータの利用目的を説明する。

### 2.1 利用想定

本研究では1節で述べたように、VoIP ユーザが LTE 過渡期や災害発生時のようにトラフィック過密状態で VoIP を使用する図 1 の状況を想定する。

図 1 では、IEEE802.11g の公衆無線 LAN と一般家庭、企業ネットワークが VoIP を利用するインフラストラクチャーネットワークを想定している。公衆無線 LAN ではスマートフォンのソフトウェア電話を利用し、一般家庭、

企業側では有線のデスクトップ PC でソフトウェア電話を利用する。主な留意点は以下の 2つとなる。

また、呼び出しや切断処理は広域インターネットに繋がる SIP Server を経由して実行する。

#### 1. VoIP ユーザのランダム移動

公衆無線 LAN のようなアウトドア系の VoIP の場合、通話中に VoIP ユーザは常に移動する。よって、VoIP ユーザの初期配置はランダムとし、公衆無線 LAN 圏内で常に移動する。

#### 2. トラフィックの過密状態

1つのアクセスポイントが許容できる複数の VoIP 通話が同時にあり、更に TCP や UDP といった他のトラフィックも混在するマルチレート環境を想定する。

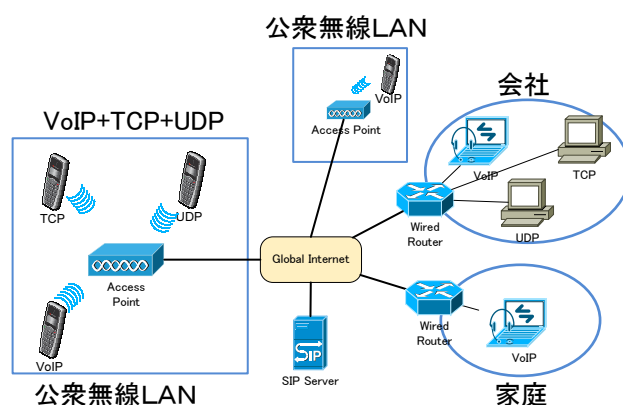


図 1 VoIP の利用想定

### 2.2 NS2 と OMNET++ の利用目的

1節で述べたように、先行研究 [1] で考慮していない「マルチレート環境での VoIP 利用」と「ユーザの主観品質を推定する実音声サンプルの生成」を、本研究では十分に考慮する。

しかし、この 2つの機能を両立したネットワークシミュレータが現時点で存在しない。本研究では 2つの要件を満たすために、マルチレート環境をシミュレート可能な NS2 と、実音声サンプルを生成可能な OMNET++ の 2つを並行して使用する。

## 3 NS2 によるマルチレート VoIP シミュレーション

本節では、NS2 によるマルチレート VoIP シミュレーションで使用する先行研究及び本研究のモデルと、ネットワーク構成を説明する。

### 3.1 NS2 で使用するモデル

NS2 によるマルチレート VoIP シミュレーションでは、先行研究 [1] の音声品質評価・改善モデルと、本研究で新たに扱うマルチレート環境の模倣に必要な 3 つのモデルを使用する。

- 音声品質評価モデル (E-model)[2]

音声品質に関係する計 20 個のパラメータから、音声品質を 0 から 100 で客観的評価する R 値を算出する式。TTC(情報通信技術委員会)の策定に準拠し、20 個のパラメータの内 18 個に TTC 標準値を用いる。他 2 個 (パケット損失率, 遅延) に関しては適宜計測が必要不可欠である。

- 音声品質改善アルゴリズム 1(音声 Codec 切り替え)

VoIP の音声コーデックを無線リンク状態に応じて切り替える。VoIP 通信開始時は、高ビットレートの音声コーデック G.711 を使用する。パケット損失率や遅延の増加により、R 値が ITU-T (国際電気通信連合電気通信標準化部門) が定める音声品質の許容範囲 (R 値 > 70) を満たさなくなった場合、低ビットレートの音声コーデック G.729 に切り替える。

- 音声品質改善アルゴリズム 2(フレーム化間隔変更)

音声信号を切り出す時間単位を示す、フレーム化間隔を変更する。フレーム化間隔を長くし、Voice ペイロードを大きくすることで、オーバーヘッド問題を回避し伝送効率を向上させる。なお、フレーム化間隔は伝送レートの高い無線ノードを優先して長くする。

- 移動モデル (RWP : Random Waypoint)

RWP モデルは本研究の想定通り、無線ノードのランダム初期配置をし、指定したシミュレーション時間だけランダム移動を繰り返すといった移動性を持たせるモデルである。

- 電波伝搬モデル (大地反射モデル)

RWP モデルによる無線ノードの移動に伴い、受信信号電力の変動を模倣するモデル。大地反射モデルはアクセスポイントから発生する直接波と地面の反射波を考慮したものである。

- リンクアダプテーション機能

大地反射モデルで与えられる受信信号電力に伴い、適切な伝送レートの切り替えをする機能。ACK フレームの連続送信成功/失敗回数でレート変化させる ARF(Automatic Rate Fallback) 方式と、RTS(Request to Send) パケットを受信した無線端末の受信信号電力に基づいてレート変化させる RBAR(Receiver Based Autorate) 方式がある。

NS2 に上記の 6 つのモデルを組み込んだマルチレート VoIP シミュレーションの流れを、以下の図 2 にまとめる。

NS2 では図 2 のように、マルチレートモデルと音声品質評価・改善モデルの 2 つを組み合わせ、先行研究 [1] の 2 種類の音声品質改善アルゴリズムを再評価することを目的とする。

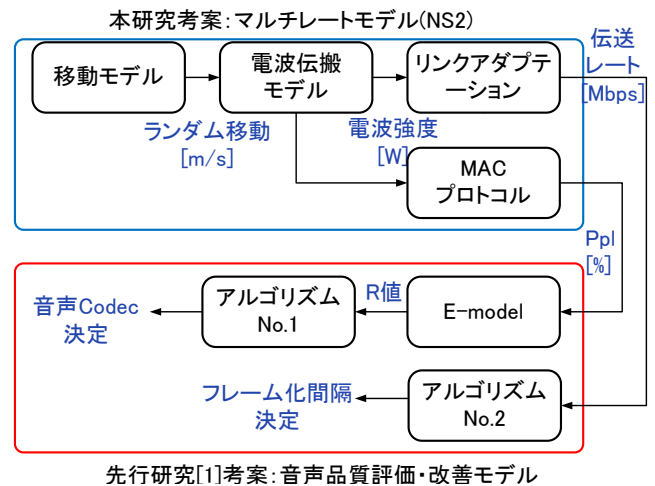


図 2 NS2 : マルチレート VoIP シミュレーションの流れ

### 3.2 NS2 におけるネットワーク構成

2.1 節に基づいて、VoIP ユーザのランダム移動とトラフィック過密状態を想定したネットワークを構成する。

IEEE802.11g における同時通話の限界許容数に合わせて、24 台の有線ノード、24 台の無線ノードの構成で VoIP 通話を開始する。さらに、1000Byte の UDP トラフィック、TCP トラフィックを 1 つずつ追加し、NS2 のネットワーク構成は 26 台の有線ノード、26 台の無線ノード、1 台の AP(IEEE802.11g) とする。

また、公衆無線 LAN の利用可能伝送距離が平均 100m であることを考慮し、アクセスポイントを中心に 200m × 200m の平面エリアをシミュレートする。公衆無線 LAN 圏内の各無線ノードは、この平面エリア内を移動する。

## 4 OMNET++による実音声シミュレーション

本節ではネットワークシミュレータ OMNET++における、音声品質劣化要因を反映させた実音声サンプルの生成方法を説明する。

### 4.1 実音声サンプルの生成

OMNET++による実音声シミュレーションは、OMNET++パッケージの配布先で提供されている VoIPTool 2.0 と、INET Framework を OMNET++ に導入することで可能になる。VoIPTool のサンプルプログラムである、以下の VoIPToolTest.ned を直接編集することで、NS2 で計測した遅延やパケット損失率を反映させた劣化音声サンプルを作成できる。

```

遅延・パケット損失率の変更 (VoIPToolTest.ned)
channel line1 extends ned.DatarateChannel {
parameters:
delay = 10ms;
per = 0.1;
}

```

しかし、OMNET++に遅延パラメータを与えても、実音声サンプルに反映されない仕様になっている。そこで、

OMNET++にパケット損失率のみ与え、次節の改良版 PESQ に遅延を代入する。パケット損失率を反映させた音源を生成した後、遅延による劣化予測値を改造版 PESQ による音声評価段階で適宜減算するといった手順を踏む。

なお、OMNET++の実音声シミュレーションのネットワーク構成は、NS2 で模倣したマルチレート環境で計測した遅延とパケット損失率をそのまま反映させるために、無劣化の有線区間の 1 対 1 通信をしている。

## 4.2 音声品質評価 (PESQ) の改良

PESQ[3] と呼ばれる音声品質評価ツールを用いて音声評価値を求める。PESQ は ITU-T P.862 で勧告された客観的評価手法であり、原音と評価対象音声を入力とし、雑音や歪み等の妨害値を計算し主観品質評価値である MOS 値にマッピングする。

本研究では E-Model の遅延計算部分を PESQ に引用することで、遅延を反映させた音声品質評価値を算出する改良を施した。以下に、改良版 PESQ の実行例を示す。

改良版 PESQ の実行例

```
./PESQ +8000 sent.wav results.wav 100.0 //delay
Reading reference file sent.wav...done.
Reading degraded file results.wav...done.
Level normalization...
IRS filtering...
Variable delay compensation...
Acoustic model processing...
P.862 Prediction (Raw MOS, MOS-LQO): = 3.173
3.080
MOS-LQO → Rfactor: = 59.612
Delay(100.0ms) → Rfactor: = 57.212
Rfactor(add delay) → MOS-CQE: = 2.955
```

原音の sent.wav と、劣化音源の results.wav を入力した実行例である。コマンドライン引数の最後に、遅延パラメータを入力することで、遅延反映させた MOS 値と R 値に変換している。

## 5 評価

本節では以下の 3 段階の評価を経て、先行研究の音声品質改善アルゴリズムについて考察する。

1. NS2 にマルチレート特性を加え、音声品質改善アルゴリズム使用時の音声品質を評価 (5.2 節)
2. NS2 で測定したパケット損失率、遅延を OMNET++ に与え、実音声サンプルの音声品質を評価 (5.3 節)
3. 両者の評価値を比較した上で、最終的に音声品質改善アルゴリズムを評価 (5.4 節)

### 5.1 評価方法

- NS2 : マルチレート VoIP シミュレーション

実際の VoIP 通話では遅延・パケット損失率を含むフィードバック情報を、RTCP パケットを用いて最短周期 5 秒で受信ノードから送信ノードへと送出する。よって、本研究では 5 秒間隔で E-model の計算を繰り返す。また、シ

ミュレーション時間を 100 秒とし、計 20 回全 VoIP 通話の R 値を測定する。

- OMNET++ : 実音声シミュレーション

本研究では 2.2 節で述べた通り、マルチレート環境による音声品質劣化要因を反映させた実音声サンプルを作成した上で、その音源を評価することを目的とする。先に NS2 でパケット損失率・遅延を計測した後、OMNET++ でも同じ値を使用する。

### 5.2 NS2(マルチレート VoIP シミュレーション) の音声品質評価

リンクアダプテーション機能のうち、ARF 方式と RBAR 方式を使用したシミュレーション想定で 10 回実行し、R 値・パケット損失率 (Ppl)・遅延 (T) を計算した結果を、以下の表 1 に示す。

表 1 マルチレート VoIP シミュレーションの音声品質評価

ARF 方式		
アルゴリズム [1]	平均 R 値	信頼区間
適用無	70.68	70.68 ± 6.62
適用有	78.91	78.91 ± 2.35
アルゴリズム [1]	平均 Ppl[%]	信頼区間 [%]
適用無	2.41	2.41 ± 1.44
適用有	0.76	0.76 ± 0.71
アルゴリズム [1]	平均遅延 [ms]	信頼区間 [ms]
適用無	79.96	79.96 ± 0.42
適用有	105.42	105.42 ± 4.06
RBAR 方式		
アルゴリズム [1]	平均 R 値	信頼区間
適用無	74.46	74.46 ± 4.93
適用有	80.13	80.13 ± 2.63
アルゴリズム [1]	平均 Ppl[%]	信頼区間 [%]
適用無	1.82	1.82 ± 0.81
適用有	0.58	0.58 ± 0.42
アルゴリズム [1]	平均遅延 [ms]	信頼区間 [ms]
適用無	79.78	79.78 ± 0.19
適用有	103.26	103.26 ± 5.57

ARF 方式使用時は音声品質改善アルゴリズム [1] を適用することで、平均 R 値は 8.23 向上した。また、分散が小さくなり、信頼区間下限を見ても R 値の音声許容値である 70 を下回ることが無くなっていることが分かる。それに伴い、パケット損失率 (Ppl) も 2% 未満に収まる。一方、遅延は適用後に音声コーデックの原理遅延分増加してしまい、パケット損失率ほどではないが R 値を下げる原因になると考えられる。

次に、RBAR 方式使用時は音声品質改善アルゴリズム [1] を適用することで、平均 R 値は 5.67 向上した。ARF 方式の結果と比較すると、RBAR 方式は平均 R 値の伸び代は少ないものの、R 値・パケット損失率・遅延全てにおいて ARF 方式よりも優れた結果を出した。

ARF 方式より RBAR 方式が優れる理由としては、RBAR 方式の場合、送信時に受信信号電力を基に伝送レートを設定し、ARF よりも無線リンク状態に適した伝送レート選択が可能であり、R 値やパケット損失率も安定しやすいと考えられる。

### 5.3 OMNET++(実音声シミュレーション)の音声品質評価

5.2 節のシミュレーションで得られたパケット損失率と遅延を用いて、10 回分実音声サンプルを OMNET++ で生成し、R 値の平均・分散・95%信頼区間を推定した結果を以下の表 2 に示す。

表 2 実音声サンプルの音声品質評価

ARF 方式 (男性)		
アルゴリズム [1]	平均 R 値	信頼区間
適用無	75.44	75.44 ± 5.50
適用有	84.37	84.37 ± 2.62
ARF 方式 (女性)		
アルゴリズム [1]	平均 R 値	信頼区間
適用無	71.77	71.77 ± 6.36
適用有	82.45	82.45 ± 3.11
RBAR 方式 (男性)		
アルゴリズム [1]	平均 R 値	信頼区間
適用無	78.48	78.40 ± 3.88
適用有	86.38	86.38 ± 2.58
RBAR 方式 (女性)		
アルゴリズム [1]	平均 R 値	信頼区間
適用無	75.74	75.74 ± 4.67
適用有	85.26	85.26 ± 2.86

5.2 節の E-model による計算式から求めた R 値と比較すると、女性音源使用時は近い値になった。一方、男性音源使用時は平均 R 値で 5.67 も差が生まれた。この原因としては、音声解析は性別や録音者本人の発音の明瞭度等、録音者の発音の特徴に依存するところが大きく、PESQ による実音声分析アルゴリズムでは音質劣化の要因が検出されなかったことが考えられる。これにより、男性音源は 5.2 節の NS2 の R 値とは異なる結果になった。

### 5.4 音声品質改善アルゴリズムの評価

最後に、NS2 と OMNET++ による 2 つの音声品質評価を基に、音声品質改善アルゴリズム [1] を評価する。

表 1, 表 2 の通り、本研究のマルチレート環境を想定した VoIP シミュレーションの場合、全 VoIP 通話の平均値をとると、大きく R 値が改善されたようには見受けられない。しかし、R 値が許容値である 70 を大きく下回る VoIP 通話に対して効果が大きく、R 値が 40 を下回る場合も、70 以上に改善する結果が出た。その例を図 3 に示す。

図 3 の場合、シミュレーション時間 40 秒経過時に、音声品質改善アルゴリズム [1] を適用しない場合は R 値が

31.23 であったのが、適用することで 82.44 まで改善された。よって、音声品質改善アルゴリズム [1] は全体的な音声品質の改善よりも、極端に音声品質が悪い VoIP 通話に対して部分的に機能することが分かる。

これにより、先行研究の音声品質改善アルゴリズム [1] はマルチレート環境による VoIP 使用時にも有効であると言える。

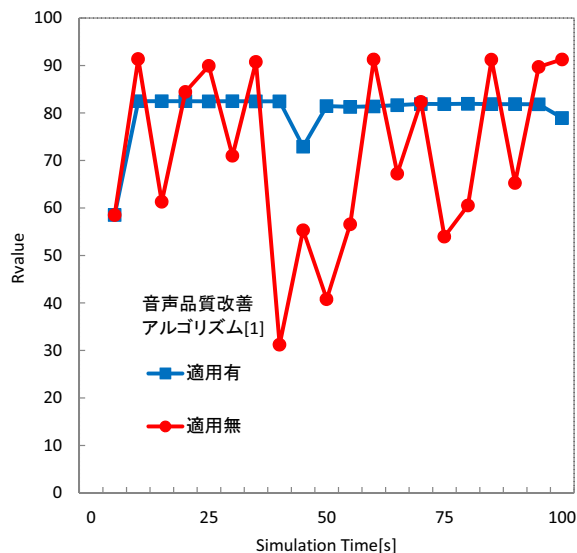


図 3 NS2 : RBAR 方式使用時の R 値変動の一例

## 6 おわりに

本研究では、NS2 でマルチレート環境を模倣し、先行研究の音声品質改善アルゴリズム [1] を評価した。そして、OMNET++ でパケット損失率を反映させた音源を生成し、実音声サンプルを用いた音声評価にも取り組んだ。2 つの評価で、マルチレート環境による VoIP 使用時にも音声品質改善ができるアルゴリズムであることを示した。

今後の課題としては、評価に使用する実音声サンプル数を増やすこと、実音声サンプルに使用する音声評価ツールをより精度の良いものに変えることが考えられる。

## 参考文献

- [1] M. F. Tuysuz, and H. A. Mantar, "An Adaptive Codec and Frame Size Modification Based QoS Algorithm over Multi-rate WLANs," In Proc of *Second International Conference, WiMo 2010, Ankara, Turkey*, Jun. 2010, pp. 137-147.
- [2] International Telecommunication Union, "G.107: The E-model: a computational model for use in transmission planning," accessed Aug. 2012, <http://www.itu.int/rec/T-REC-G.107/>.
- [3] International Telecommunication Union, "P.862 : Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," accessed Oct. 2012, <http://www.itu.int/rec/T-REC-P.862/>.