

УДК 681.43.06:389.14

Ю. В. Дубова, асп.

ТОЧНІСТЬ МОДЕЛІ КОНТРОЛЮ ЯКОСТІ КАНАЛІВ ПЕРЕДАВАННЯ ГОЛОСОВОЇ ІНФОРМАЦІЇ

Розглянуто сучасні підходи до проблеми контролю якості передавання голосової інформації телекомунікаційними мережами та їх недоліки. Детально проаналізовано модель контролю, що ґрунтується на оцінюванні якості розпізнавання переданої голосової інформації та запропоновано методику оцінювання її точності. Запропонований підхід до визначення точності моделі контролю якості каналів передавання голосової інформації безпосередньо визначає характеристики каналу і системи його тестування за його призначенням, тоді як інші методи визначають якість опосередковано, на основі спектральних, шумових та інших характеристик.

Задача оцінки якості передавання голосу є досить нетривіальною, оскільки кінцевою метою доставки голосової інформації є людина, яка і аналізує голосову інформацію за допомогою органів чуття, які досить важко, якщо й не зовсім неможливо, змоделювати. Звичайно, в цій ситуації природно використовувати так звану «експертну», або «суб'єктивну», тобто людську оцінку якості звуку, але такий сервіс є дуже дорогим та на регулярній основі його можуть дозволити собі лише декілька телекомунікаційних компаній у всьому світі, а саме такі «гіганти» ринку телефонії, як «AT&T Laboratories», «British Telecom», «Lucent Technologies Bell Labs» та декілька інших. Перед іншими компаніями постає проблема автоматизації оцінки якості звуку, над розв'язанням якої працюють науковці вже багато років.

Дослідження в галузі обробки, передавання та розпізнавання голосової інформації проводяться вітчизняними науковцями Інституту кібернетики АН України, Національного технічного університету України «Київський політехнічний інститут», Національного університету «Львівська політехніка», Харківського національного університету радіоелектроніки, Донецького інституту проблем штучного інтелекту, Вінницького національного технічного університету.

Серед закордонних наукових дослідників необхідно виділити міжнародну організацію International Telecommunication Union (ITU), а саме її підрозділ ITU Telecommunication Standardization Sector (ITU-T), який встановлює всесвітні стандарти та рекомендації в області телекомунікацій [2]. Необхідність створення такої організації була обумовлена тим, що з розвитком телекомунікацій почали з'являтися численні стандарти передавання, криптування та кодування інформації, тому узгоджена взаємодія телекомунікаційних програмних й апаратних засобів стала вкрай необхідною. Створення загальних рекомендацій в області телефонії дало можливість компаніям — виробникам обладнання та програмного забезпечення, реалізовувати стандартні алгоритми приймання/передавання даних із впевненістю в тому, що дані передаватимуться коректно незалежно від того, яке обладнання знаходиться на іншому боці каналу.

Взагалі, на сучасному рівні розв'язання проблеми контролю якості передавання голосової інформації телекомунікаційними мережами використовують три підходи:

- 1) апаратний контроль фізичних параметрів каналів зв'язку;
- 2) програмне тестування протоколів зв'язку;
- 3) експертне оцінювання якості прийнятої голосової інформації, про яке йшлося вище [3].

Отже, постає **задача** аналізу точності моделі контролю якості каналів передавання голосової інформації.

Кількісна оцінка якості системи передавання голосової інформації знаходиться на основі дослідження погіршення роздільної здатності системи розпізнавання на просторі прийнятої інформації у порівнянні з роздільною здатністю на просторі переданої інформації

$$K = \frac{1}{m} \sum_{i=1}^m \left[\rho_0(F_i, F_j) / \rho(F_i, F_j) \right], \quad (1)$$

де m — кількість фонем; $\rho(F_i, F_j)$ — відстань від i -ї фонемі до найближчої, які після передавання розпізнаються системою як різні; $\rho_0(F_i, F_j)$ — відстань від i -ї фонемі до найближчої, які до передавання розпізнавалися системою як різні.

Похибка оцінювання показника якості

$$\sigma_K = \sqrt{\left(\frac{\partial K}{\partial \rho_0} \sigma_{\rho_0}\right)^2 + \left(\frac{\partial K}{\partial \rho} \sigma_{\rho}\right)^2 - 2r_{\rho_0\rho} \sigma_{\rho_0} \sigma_{\rho}}, \quad (2)$$

де σ — с. к. в. відповідного параметра, r — коефіцієнт кореляції, який характеризує залежність якості розпізнавання від відстані між фонемами $\rho_0(F_i, F_j)$.

Коефіцієнт кореляції $r_{\rho_0\rho}$ оцінюватимо на основі адитивної моделі каналу

$$r_{\rho_0\rho} = \frac{\sigma_{\rho_0}^2 + \sigma_{\rho}^2}{\sigma_{\rho_0} \sigma_{\rho}}. \quad (3)$$

Підставляючи (3) в (2), отримуємо:

$$\sigma_K = \sqrt{\left[\left(\frac{\partial K}{\partial \rho_0}\right)^2 - 2\right] \sigma_{\rho_0}^2 + \left[\left(\frac{\partial K}{\partial \rho}\right)^2 - 2\right] \sigma_{\rho}^2}. \quad (4)$$

Підставляючи (1) в (4) остаточно отримуємо вираз для оцінювання похибки моделі

$$\sigma_K = \sqrt{\left[\left(\frac{1}{m} \sum_{i=1}^m [1/\rho(F_i, F_j)]\right)^2 - 2\right] \sigma_{\rho_0}^2 + \left[\left(\frac{1}{m} \sum_{i=1}^m [\rho_0(F_i, F_j)/\rho^2(F_i, F_j)]\right)^2 - 2\right] \sigma_{\rho}^2}. \quad (5)$$

Оцінювання відстані $\rho_0(F_i, F_j)$ здійснюється відповідно до методики [4], яка ґрунтується на описі фонем у багатовимірному просторі ознак за допомогою прихованої марковської моделі (ПММ). Відстань між фонемами є функцією від параметрів ПММ фонем F_i і F_j

$$\rho_0(F_i, F_j) = f(\overline{m_{ik}}, \overline{\sigma_{ik}}, P_{ik}, \overline{m_{jk}}, \overline{\sigma_{jk}}, P_{jk}; k = 1 \dots n), \quad (6)$$

де n — кількість станів ПММ; $\overline{m_{ik}}, \overline{\sigma_{ik}}$ — вектори математичних сподівань і с. к. в. багатовимірних розподілів ймовірності ознак у кожному стані ПММ фонем F_i ; $\overline{m_{jk}}, \overline{\sigma_{jk}}$ — вектори математичних сподівань і с. к. в. багатовимірних розподілів ймовірності ознак у кожному стані ПММ фонем F_j ; P_{ik}, P_{jk} — ймовірності потрапляння у k -й стан ПММ фонем F_i і F_j відповідно,

$$P_{ik} = \prod_{a_{uv} \in L_{ik}} a_{uv}, \quad (7)$$

L_{ik} — шлях до k -го стану ПММ фонем F_i ; a_{uv} — ймовірність переходу зі стану u у стан v ПММ.

Оцінювання відстані $\rho(F_i, F_j)$ здійснюється на базі статистичної обробки результатів тестування. Відстань $\rho(F_i, F_j)$ характеризує якість розпізнавання фонем.

Методика оцінювання похибки передбачає:

— статистичний аналіз показників σ_{ρ} і r на основі даних експериментальних досліджень каналу передавання голосової інформації;

— оцінювання показника σ_{ρ_0} аналітичним методом на основі виразу відстані через характеристики прихованої марковської моделі фонем [4].

Основними джерелами похибок моделі контролю якості каналів передавання голосової інформації є:

— похибки оцінювання статистичних характеристик ПММ, зумовлені обмеженням обсягом тестів;

— похибки вимірювання значень ознак у багатовимірному просторі;

— похибка апроксимації залежності ймовірності розпізнавання від відстані між фонемами $P_r(\rho)$;

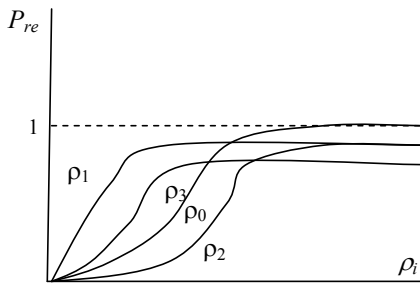
— обчислювальна похибка знаходження відстані на основі експериментальної оцінки P_{re} і розв'язання рівняння $P_r(\rho) = P_{re}$.

Величини складових похибки залежать від обраної метрики (6) у просторі фонем. Найпоширеніші метрики та відповідні похибки цих метрик наведені в таблиці [5, 6].

Найпоширеніші метрики та відповідні похибки цих метрик

Назва метрики	Метрика	Похибка, %
Гранична відстань	$\rho_3 = \sup_{t,j} P_j(t) \cdot G_j - P_j'(t) \cdot G_j' $	0,20
Метрика «міських кварталів» (Манхеттенська відстань)	$\rho_2 = \sum_{t=1}^{\infty} \sum_{j=2}^{N-1} \left[\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} P_j(t) \cdot G_j - P_j'(t) \cdot G_j' dx \right]$	0,18
Метрика Бхатахарія (Bhattacharyya)	$\rho_1 = \sum_{t=1}^{\infty} \sum_{j=2}^{N-1} \left[\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \left \frac{d(P_j(t) \cdot G_j)}{dx} \cdot \frac{d(P_j'(t) \cdot G_j')}{dx} \right dx \right]$	0,11
Метрика, запропонована у [4]	$\rho_0 = \sum_{t=1}^{\infty} \sum_{j=2}^{N-1} \left[\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} (P_j(t) \cdot G_j - P_j'(t) \cdot G_j')^2 dx \right]^{\frac{1}{2}}$	0,02

Дані у таблиці розраховані для фонему $F_i = \langle \text{а} \rangle$ як базової і фонем $F_j = [\langle \text{ха} \rangle, \langle \text{га} \rangle, \langle \text{да} \rangle, \langle \text{та} \rangle, \langle \text{ас} \rangle, \langle \text{ах} \rangle]$ одновимірних гаусівських розподілів G_j , ознака — частота першої гармоніки, $\sigma = 20$ Гц, $\sigma' = 30$ Гц, ПММ з $N = 5$ вершинами, $t = 1 \dots 300000$ (час промовляння фонему 0,3 с, частота дискретизації 1 МГц).



Експериментальні залежності ймовірності розпізнавання голосової інформації для різних метрик

Експериментальні залежності $P_{re}(\rho)$ показані на рис. Експерименти полягали у статистичному обробленні результатів розпізнавання корпусу наговорів слів, які містили фонему «а» та інші фонему, і розрахунках відповідної відстані між фонемами за чотирма метриками. Експериментальні дослідження доводять, що використання запропонованої метрики [4] дають можливість отримати більшу ймовірність розпізнавання, а отже її можна використовувати для контролю якості каналів передавання голосової інформації з більшою достовірністю.

Запропонована методика застосована для синтезу тестів, які забезпечують необхідну повноту і достовірність контролю якості каналів передавання голосової інформації.

Висновки

Запропонований підхід до визначення точності моделі контролю якості каналів передавання голосової інформації має перевагу перед відомими, яка полягає у прямому визначенні характеристик каналу і системи його тестування за його призначенням (інші методи визначають якість опосередковано, на основі спектральних, шумових та інших характеристик). Недолік підходу – більша складність, яка потребує повної автоматизації.

СПИСОК ЛІТЕРАТУРИ

1. Perceptual Evaluation of Speech Quality (PESQ). <http://microtronix.ca/pesq-disc.html>
2. ITU-T Recommendation H.225.0. Call signaling protocols and media stream packetization for packet-based multimedia communication systems. — Geneva, 1998.
3. Гольдштейн Б. С. Сигналізація в сетях связи. Том 1. — М.: Радио и связь, 1998.
4. Грузман М. З., Дубова Ю. В. Розпізнавання голосових команд як точок метричного простору // Вісник Вінницького політехнічного інституту. — 2003. — № 6. — С. 262—268.
5. A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions // Bull. Calcutta Math. Soc. — 1943. — 35:99—109.
6. H. Chernoff. Measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations // Ann. Math. Stat. — 1952. — 23. — P. 493—507.

Дубова Юлія Володимирівна — аспірантка кафедри автоматичної та вимірювальної техніки.

Вінницький національний технічний університет