

## ДОСЛІДЖЕННЯ РОЗПІЗНАВАННЯ МОВНИХ СИГНАЛІВ ЧОТИРИФОНЕМНИМ МЕТОДОМ

© Чорній А.М., 2009

**На основі запропонованої чотирифонемної моделі проведено експериментальні дослідження розпізнавання мовних сигналів. Встановлено, що цей метод дає змогу покращити ймовірність розпізнавання мовних сигналів порівняно з існуючими методами.**

**On the basis of the offered fourphonemes model experimental researches of recognition of linguistic signals are conducted. It is set that this method allows to improve probability of recognition of linguistic signals in comparing to the existent methods.**

### Вступ

Стрімкий розвиток цифрових технологій, без яких важко уявити діяльність людини у будь-яких сферах життя, відкриває безліч нових можливостей, наслідком чого постає завдання ретельного дослідження й аналізу існуючих методів розпізнавання мовних сигналів, з метою вдосконалення, покращання параметрів ефективності та надійності, побудованих на їх основі високопродуктивних цифрових систем та пристроїв.

Впровадження інформаційних комп'ютерних технологій у різноманітні сфери людської діяльності вимагає створення нових інтерфейсів, що підтримують мовний діалог "користувач-комп'ютер", і впровадження систем розпізнавання мови дасть змогу спростити взаємодію користувача з комп'ютером. На жаль, більшість систем розпізнавання мови орієнтована на англomовного користувача. Розроблення такої системи для української мови є трудомістким завданням через специфічні особливості української мови.

Звукова мова є не тільки найважливішим і найприроднішим засобом комунікації між людьми, але й однією із основних форм подання інформації в людино-машинних системах обробки інформації і управління.

Основними перевагами організації обміну інформацією за допомогою звукової мови є природність, відсутність вимог спеціальної підготовки та складних засобів відображення інформації, можливість одночасного обміну інформацією з декількома абонентами, використання як терміналу звичайного телефону. Також важливим достоїнством мовного каналу як засобу обміну інформацією є його достатньо висока продуктивність.

Для вирішення проблеми підвищення ймовірності розпізнавання мовних сигналів розробляються нові методи розпізнавання, які здатні враховувати більше аспектів варіації людської мови.

### Постановка завдання

Для розпізнавання мовних сигналів використовуються одно-, дво- та трифонемні моделі, однак вони не завжди забезпечують належну точність та ймовірність розпізнавання. Поставлено завдання розроблення й дослідження методу, який би покращував показники точності та ймовірності розпізнавання мовних сигналів порівняно з існуючими методами.

### Чотирифонемна модель розпізнавання мовних сигналів

Для покращання показників точності та ймовірності розпізнавання мовних сигналів запропоновано чотирифонемну модель розпізнавання мовних сигналів (рис. 1).

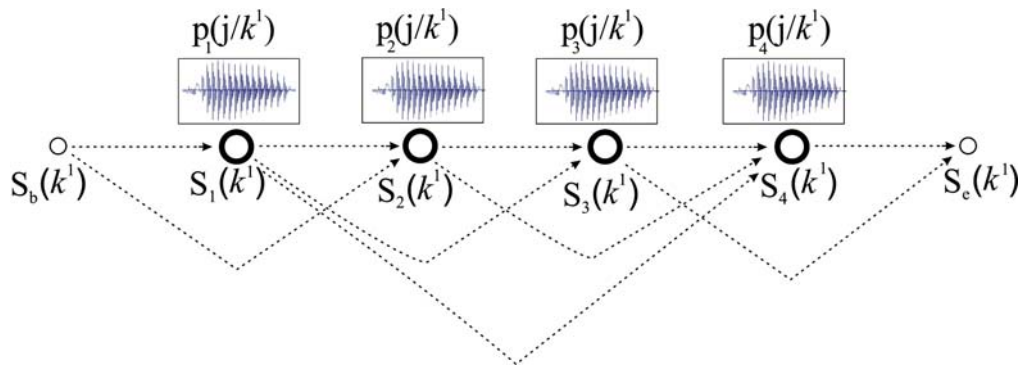


Рис. 1. Модель з чотирма станами:

$S_b(k^1)$  – початковий стан моделі;  $S_c(k^1)$  – кінцевий стан моделі;

$S(k^1)$  – основний стан моделі для фонем  $k^1$ ;

$p(j/k^1)$  – елемент, згенерований під час переходу на один такт

### Розпізнавання голосових команд

Дослідження проводилось на комп'ютері з такими технічними параметрами:

- процесор: Intel CeleronD 2,67 Гц;
- оперативна пам'ять: DDR 2 ГБ;
- операційна система: Windows Vista.

Дослідження швидкості розпізнавання голосових команд проводилось для словників різного обсягу (10, 20, 30, 50, 100, 125, 150 команд).

Результати дослідження розпізнавання голосових команд наведено у табл. 1.

Таблиця 1

### Результати розпізнавання голосових команд

№ досл.		Обсяг словника, команд						
		10	20	30	50	100	125	150
1	Час розпізнавання, с	0,23	0,25	0,40	0,75	2	2,62	3,75
2		0,21	0,26	0,41	0,76	2,11	2,65	3,78
3		0,20	0,24	0,42	0,78	2,31	2,32	3,68
4		0,21	0,25	0,41	0,78	2,11	2,55	3,95
5		0,23	0,26	0,43	0,79	2,15	2,89	3,55
6		0,21	0,27	0,45	0,75	2,14	2,45	3,74
7		0,25	0,25	0,42	0,72	1,95	2,32	3,72
8		0,23	0,26	0,46	0,71	1,95	2,65	3,65
9		0,24	0,25	0,47	0,75	2,01	2,66	3,84
10		0,21	0,21	0,48	0,71	2,05	2,58	3,70
<b>Сер.зн.</b>		<b>0,222</b>	<b>0,25</b>	<b>0,435</b>	<b>0,75</b>	<b>2,078</b>	<b>2,569</b>	<b>3,736</b>

На основі одержаних результатів на рис. 1 побудовано залежність часу розпізнавання голосових команд від розміру словника.

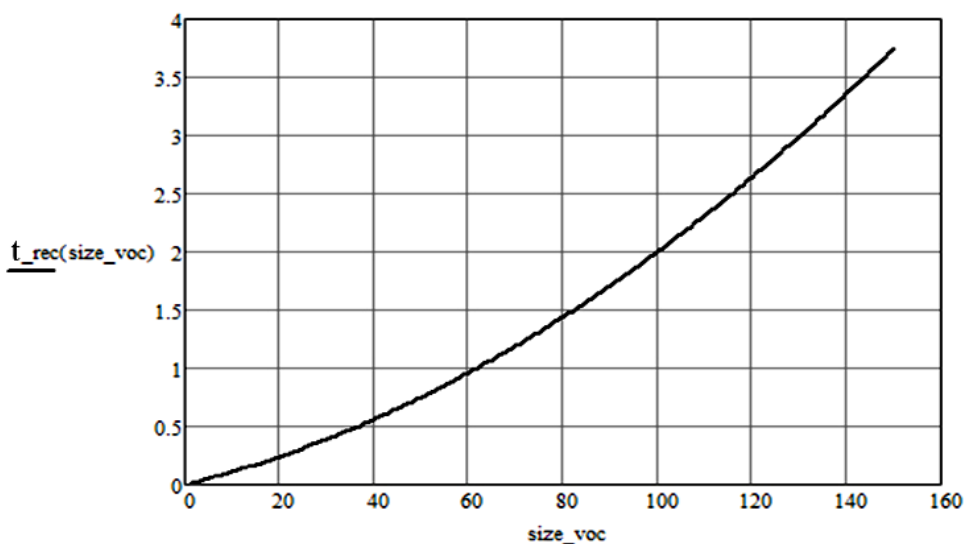


Рис. 1. Залежність часу розпізнавання команд від розміру словника

У результаті отриманих результатів можна сказати, що час розпізнавання голосових команд залежить від розміру словника, і чим більший розмір словника, тим більше часу потрібно на розпізнавання.

#### Розпізнавання злитого мовлення чотирифонемним методом

Для дослідження ймовірності розпізнавання злитого мовлення чотирифонемним методом було записано вибірку, яка складається зі 100 зразків: за допомогою 50 чоловіків та 50 жінок віком від 18 до 45 років. Для дослідження було вибрано фразу «сушка».

Під час розпізнавання злитого мовлення, розробленим програмним забезпеченням були отримані результати, наведені в табл. 2.

Таблиця 2

#### Результати дослідження розпізнавання злитого мовлення

№ досл.	Обсяг словника							
	25		50		75		100	
	Ймовір. роз. %	Час. розп. с.	Ймовір. роз. %	Час. розп. с.	Ймовір. роз. %	Час. розп. с.	Ймовір. фект. роз. %	Час. розп. с.
1.	94	1,4	93	1,8	94	2,1	95	2,5
2.	93	1,42	94	1,8	94	2,15	95	2,3
3.	94	1,41	94	1,85	94	2,3	95	2,8
4.	94	1,4	94	1,82	95	2,15	95	2,5
5.	93	1,4	93	1,83	93	2,45	95	2,7
6.	95	1,45	92	1,82	95	2,14	94	2,7
7.	93	1,5	94	1,7	95	2,15	95	2,5
8.	92	1,4	93	1,88	94	2,46	95	2,3
9.	93	1,5	92	1,9	95	2,5	95	2,4
10.	92	1,45	92	1,75	94	2,4	95	2,5
<b>Сер.</b>	93.3	1.433	93.1	1.815	94.3	2.28	94.9	2.52

На основі одержаних результатів розпізнавання злитого мовлення на рис. 2 побудовано залежність ймовірності розпізнавання від розміру словника, а на рис. 3 – залежність часу розпізнавання залежно від розміру словника.

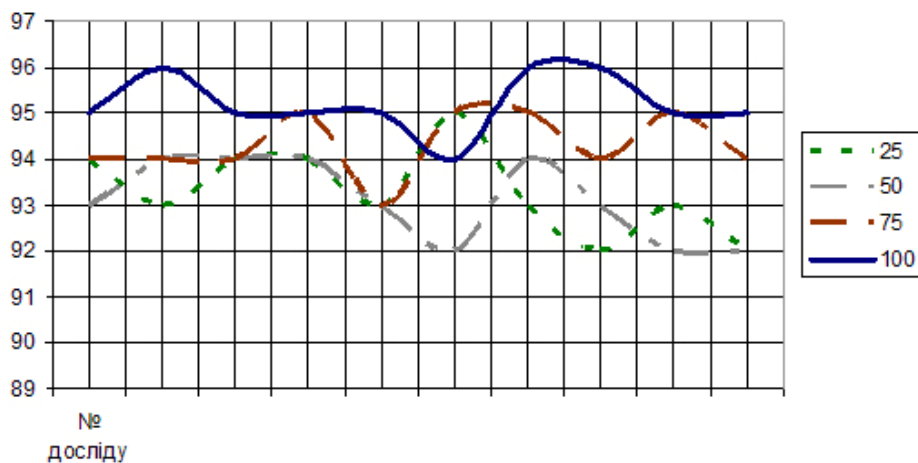


Рис. 2. Залежність ймовірності розпізнавання від розміру словника

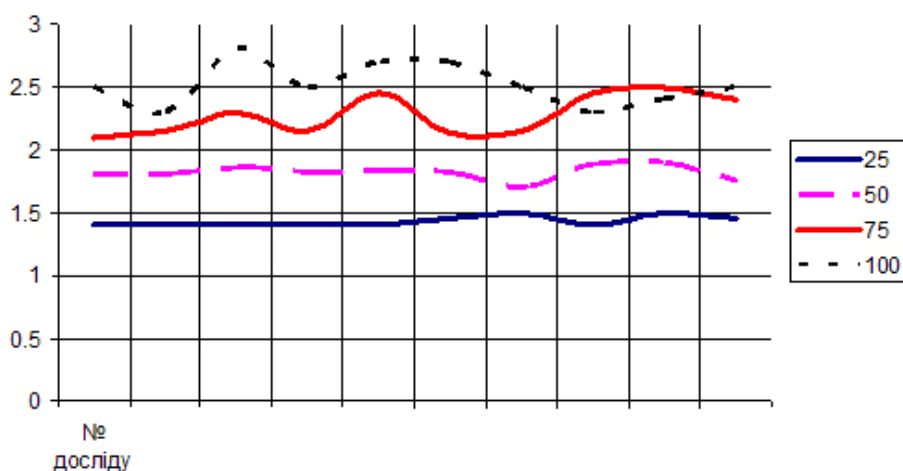


Рис. 3. Залежність часу розпізнавання від розміру словника

## Висновки

У роботі наведено чотирифонемну модель розпізнавання мовних сигналів, проведено дослідження ймовірності розпізнавання для голосових команд і для злитого мовлення і досягнуто ймовірність розпізнавання 95 %. На основі отриманих результатів можна зробити висновок, що ймовірність розпізнавання покращується зі збільшенням розміру словника, але разом з тим і збільшується час розпізнавання.

1. Винцюк Т.К. *Аналіз, розпознавание и интерпретация речевых сигналов*. – К.: Наук. думка, 1987. – 264 с. 2. Чорній А.М. Чотирифонемна модель розпізнавання мовних сигналів в телекомунікаційних мережах // *НМК «Сучасні проблеми телекомунікацій і підготовка фахівців в галузі телекомунікацій-2005»: Матеріали конференції 20–23 жовтня 2005 р.* – Львів, 2005. – С.31–33. 3. Чорній А.М. Сегментація в системах розпізнавання мовних сигналів: *Збірник наукових праць «Моделювання та інформаційні технології»*. – К., 2007. – Вип.42. – С.197–201. 4. Винцюк Т.К., Юхименко О.А. Оцінювання імовірнісних параметрів найпростіших моделей фонем // *Ймовірнісні моделі та обробка випадкових сигналів і полів: Міжнародний симпозіум*. – Тернопіль, 1993. – Т. 3, Ч. 2. – С.11–16. 5. Цыплихин А.И., Сорокин В.Н. Сегментація речи на кардинальные элементы // *Институт проблем передачи информации, Российская академия наук*. – М.: Информационные процессы, 2000. – Т. 6, № 3. – С. 177–207. 6. Бабкин В.В. Помехоустойчивый выделитель основного тона речи // *7-я Международная конференция и выставка «Цифровая обработка сигналов и ее применение DSPA-2005»*. – Москва, 16–18 марта 2005 г.