

УДК 621.391

ОБ ОДНОМ ПОДХОДЕ К РАСПОЗНАВАНИЮ ДИСКРЕТНЫХ РЕЧЕВЫХ
СООБЩЕНИЙ С ЖЕСТКИМ ПОРЯДКОМ СЛЕДОВАНИЯ СЛОВ

Р.Л.Кремень

Важнейшей характеристикой информационно-поисковых систем с устным запросом, речевых терминалов, является надежность распознавания речевого сообщения (РС). Под РС подразумевается фраза, предложение и, как частный случай, отдельное слово. Так как признаки, используемые при распознавании слов, в общем случае неинвариантно описывают образ, важное значение приобретают процедуры, позволяющие искусственно повышать надежность распознавания РС. Как показывает анализ литературных источников, в большинстве экспериментальных систем распознавания дискретной речи процедура распознавания РС состоит в последовательном распознавании в процессе произношения каждого слова слева направо. При этом как отказ от распознавания, так и ложное распознавание любого слова сообщения приводят к нарушению взаимодействия между оператором и распознающей системой. В первом случае требуется повторное произнесение слова, во втором — исправление ошибки. Отказ от распознавания или ложное распознавание возникают вследствие пересечения исходных множеств эталонов, порождаемых различными словами. При формировании списков эталонов распознающей системы элемент подмножества, являющегося пересечением исходных множеств эталонов, либо вовсе исключается, либо включается в список эталонов того или иного слова, при этом данный элемент из списков эталонов других слов исключается. В противном случае однозначное распознавание становится невозможным.

Существует, однако, иной подход. Пусть списки эталонов слов представлены настолько полно, что вероятность отказа от распознавания существенно мала. При этом никаких ограничений на пересекае-

мость множеств эталонов слов не накладывается, т.е. допускается неоднозначное распознавание. Процедура принятия решения об имеющемся месте РС пусть выглядит следующим образом. После произнесения первого слова РС определяется принадлежность его к одному из возможных РС. Исключая вариант отказа от распознавания, здесь возможны следующие два случая: а) слово распознано однозначно, б) слово распознано неоднозначно. В первом случае можно сделать заключение о всем РС, во втором случае анализируется второе по порядку слово РС, причем анализ ведется на множествах эталонов тех РС, которые распознались неоднозначно в результате анализа первого слова. Процесс может продолжаться до n , где n – число слов в РС.

Постановка задачи. Пусть имеется набор из $N+1$ речевых сообщений, каждое из которых состоит из n слов. Будем говорить, что неоднозначность распознавания имеет степень i , если в результате произнесения какого-либо слова неоднозначно распознано $i+1$ слово, включая произнесенное. Максимально возможная степень неоднозначности равна μ , где $\mu \leq N$. Требуется определить вероятность неоднозначного распознавания РС $P(n, \mu, N)$. Можно показать, что $P(n, \mu, N)$ определяется из следующего выражения:

$$P(n, \mu, N) = \sum_{i_1=1}^{\mu} C_N^{i_1} \sum_{i_2=1}^{i_1} C_{i_1}^{i_2} \sum_{i_3=1}^{i_2} C_{i_2}^{i_3} \dots \sum_{i_{n-1}=1}^{i_{n-2}} C_{i_{n-2}}^{i_{n-1}} P_N^{i_1}(k_1) P_{i_1}^{i_2}(k_1, k_2) \dots$$

$$\dots P_{i_{n-1}}^{i_n}(k_1, k_2, \dots, k_n), \quad (1)$$

где $P_N^{i_1}(k_1)$, $P_{i_1}^{i_2}(k_1, k_2)$, ..., $P_{i_{n-1}}^{i_n}(k_1, k_2, \dots, k_n)$ – вероятности событий, заключающихся в появлении неоднозначностей распознавания степеней i_1, i_2, \dots, i_n , возникающих в результате произнесения соответственно первого, второго, ..., n -го слова РС; k_1, k_2, \dots, k_n – порядковые номера соответствующих событий; $C_N^{i_1}, C_{i_1}^{i_2}, \dots, C_{i_{n-1}}^{i_n}$ – число сочетаний соответственно из N по i_1 , из i_1 по i_2, \dots , из i_{n-1} по i_n .

Для иллюстрации поведения функции $P(n, \mu, N)$ положим

$$P_N^{i_1}(k_1) = P_{i_1}^{i_2}(k_1, k_2) = \dots = P_{i_{n-1}}^{i_n}(k_1, k_2, \dots, k_n) = P.$$

Тогда формула (1) преобразуется к виду:

$$P(n, \mu, N) = P^n \sum_{i_1=1}^{\mu} \sum_{i_2=1}^{i_1} \dots \sum_{i_n=1}^{i_{n-1}} C_N^{i_1} C_{i_1}^{i_2} \dots C_{i_{n-1}}^{i_n}. \quad (2)$$

В таблице приведены значения функции $P(n, \mu, N)$ при $P = 0,001$ и различных n, μ, N , рассчитанные по формуле (2).

Т а б л и ц а

μ	$P(1, \mu, N) \cdot 10^{-1}$ при N			$P(2, \mu, N) \cdot 10^{-4}$ при N			$P(3, \mu, N) \cdot 10^{-7}$ при N		
	10	15	18	10	15	18	10	15	18
1	0,10	0,15	0,18	0,10	0,15	0,18	0,10	0,15	0,18
2	0,55	1,20	1,71	1,45	3,30	4,77	2,35	5,40	7,83
3	1,75	5,75	9,87	9,85	35,15	61,89	25,15	91,85	162,9

Из таблицы видно, что уже при $n = 2$ вероятность неоднозначного распознавания РС становится существенно малой величиной, что подтверждает высокую эффективность изложенного здесь подхода.

Поступила в ред.-изд.отд.
12 апреля 1984 года