

УДК 004.934.2

МАРКОВСКИЕ МОДЕЛИ В ЗАДАЧАХ ПАРАМЕТРИЗАЦИИ РЕЧЕВЫХ СИГНАЛОВ

Д. Е. Прозоров, К. В. Плетнев

Вятский государственный университет
Вятский государственный гуманитарный университет

Статья получена 2 июня 2014 г.

Аннотация. В статье исследована возможность применения методов параметризации речевых сигналов простыми и сложными (многосвязными) цепями Маркова в системах автоматического распознавания команд (АРК). Показано, что параметризация речевых сигналов многосвязными цепями Маркова позволяет достичь компромисса между эффективностью распознавания речевых команд и вычислительной сложностью алгоритма параметризации. Низкая вычислительная сложность алгоритмов марковской параметризации свидетельствуют о возможности значительного снижения требований к вычислительным ресурсам систем АРК.

Ключевые слова: распознавание речи, параметризация, многосвязная цепь Маркова.

Abstract: The article investigates the possibility of using of the speech parameterization methods by simple and complex (multivariable) Markov chains in systems of automatic recognition of commands (ARC). It is shown that the speech parametrization by multivariable Markov chains allows to reach a compromise between the efficiency of speech recognition commands and computational complexity of the algorithm parameterization. Low computational complexity Markov parameterization indicate the possibility of a significant reduction in requirements for computational complexity of ARC.

Key words: speech recognition, parameterization, multivariate Markov chain.

1. Введение

К настоящему времени разработано большое количество моделей и методов параметризации речевых сигналов, обеспечивающих высокое качество работы систем автоматического распознавания речевых команд (АРК), среди которых наиболее распространены методы, основанные на расчете кепстральных коэффициентов и коэффициентов линейного предсказания речи [1-3].

Однако нельзя сказать, что задача параметризации полностью решена. Производительность систем АРК еще далека от «производительности» человеческой слуховой системы. Например, в задаче распознавания цифр, когда словарь мал и существенная часть ресурсов тратится на акустическое моделирование, производительность систем АРК на порядок ниже производительности человека [4]. Отчасти это связано с большим количеством скрытых и явных переменных состояния современных систем АРК. Таким образом, задача разработки моделей и методов параметризации речевых сигналов, позволяющих найти компромисс между производительностью, требованиям к ресурсам и качеством работы систем АРК остается актуальной.

Первые работы, посвященные марковским моделям речевых сигналов появились еще в 1970-х годах. Так, в монографии [5] проведен подробный анализ марковских моделей дельта-модулированных речевых сигналов. Тем не менее, малое количество публикаций, посвященных исследованию методов марковской параметризации речевых сигналов и такие достоинства марковских моделей как относительно небольшое количество параметров и линейная вычислительная сложность алгоритмов расчета параметров, продолжают вызывать интерес к указанным методам.

В данной статье рассмотрены марковские модели в задачах параметризации речевых сигналов.

2. Постановка задачи

Пусть фрагмент речевого сигнала представлен массивом дискретных отсчетов

$$\mathbf{s} = \{s^{(1)} s^{(2)} \dots s^{(L)}\}, \quad (1)$$

где $s^{(k)}$ – k -я двоичная выборка вида

$$s^{(k)} = \sum_{b=0}^{N-1} ({}^b s^{(k)} \cdot 2^{N-b-1}), \quad k = \overline{1, L}, \quad {}^b s^{(k)} = \overline{0, 1}, \quad (2)$$

L – длина фрагмента.

Фрагмент (1) можно рассматривать как суперпозицию битовых последовательностей ${}^b \mathbf{s}$

$$\mathbf{s} = \sum_{b=0}^{N-1} ({}^b \mathbf{s} \cdot 2^{N-b-1}), \quad (3)$$

где $\{\cdot\}$ – операция скалярного умножения, ${}^b \mathbf{s}$ – бинарная последовательность, образованная b -м битом двоичных выборок фрагмента речевого сигнала.

Ряд экспериментов [5,6] показывает, что для речевых сигналов справедливо допущение о практически монотонном уменьшении корреляционных связей между выборками $s^{(n)}$ и $s^{(n-k)}$ речевого сигнала с увеличением интервала k между выборками. Следовательно, можно указать интервал m , за пределами которого корреляционные связи практически не распространяются. С учетом этого допущения, для описания и анализа коротких фрагментов речевых сигналов можно использовать математический аппарат простых или сложных цепей Маркова.

Требуется разработать модель и метод параметризации фрагмента речевого сигнала (1).

3. Многосвязная цепь Маркова

Рассмотрим многосвязную (связности m) однородную цепь Маркова с $l = 2^N$ состояниями и вероятностями перехода вида

$$P\left(s_n^{(k)} \mid s_i^{(k-1)} s_j^{(k-2)} \dots s_r^{(k-m)} s_q^{(k-m-1)}, \dots\right) = P\left(s_n^{(k)} \mid s_i^{(k-1)} s_j^{(k-2)} \dots s_r^{(k-m)}\right), \quad (4)$$

где $i, j, r, q, n = \overline{1, l}$, $k = \overline{1, L}$.

Перейдем от многосвязной цепи Маркова к простой, формируя вектор $\vec{s}^{(k-1)} = \left(s_i^{(k-1)} s_j^{(k-2)} \dots s_r^{(k-m)}\right)$ длины m [7].

Тогда

$$P\left(\vec{s}^{(k)} \mid \vec{s}^{(k-1)} \dots \vec{s}^{(k-m)}\right) = P\left(s_n^{(k)} s_i^{(k-1)} \dots s_v^{(k-m+1)} \mid s_i^{(k-1)} s_j^{(k-2)} \dots s_r^{(k-m)}\right), \quad (5)$$

где $i, j, n, r, v = \overline{1, l}$.

Число состояний полученной таким образом простой цепи Маркова равно l^m .

При известных условных вероятностях (4) можно определить вероятности перехода

$$\begin{aligned} P\left(\vec{s}^{(k)} \mid \vec{s}^{(k-1)}\right) &= \frac{P\left(\left\{s_n^{(k)} \dots s_v^{(k-m+1)}\right\}, \left\{s_i^{(k-1)} \dots s_r^{(k-m)}\right\}\right)}{P\left(s_i^{(k-1)} \dots s_r^{(k-m)}\right)} = \\ &= \frac{P\left(s_n^{(k)} s_i^{(k-1)} \dots s_r^{(k-m)}\right)}{P\left(s_i^{(k-1)} \dots s_r^{(k-m)}\right)} = P\left(s_n^{(k)} \mid s_i^{(k-1)} \dots s_r^{(k-m)}\right), \end{aligned} \quad (6)$$

где $i, j, n, r, v = \overline{1, l}$.

Например, при $m = 2$ и $l = 2$ переходные вероятности (6) принимают вид

$$\begin{aligned} P\left(\vec{s}^{(k)} \mid \vec{s}^{(k-1)}\right) &= \frac{P\left(\left\{s_n^{(k)} s_i^{(k-1)}\right\}, \left\{s_i^{(k-1)} s_j^{(k-2)}\right\}\right)}{P\left(s_i^{(k-1)} s_j^{(k-2)}\right)} = \\ &= \frac{P\left(s_n^{(k)} s_i^{(k-1)} s_j^{(k-2)}\right)}{P\left(s_i^{(k-1)} s_j^{(k-2)}\right)} = P\left(s_n^{(k)} \mid s_i^{(k-1)} s_j^{(k-2)}\right), \end{aligned} \quad (7)$$

где $i, j, n = \overline{1, 2}$.

Переходные вероятности (7) образуют матрицу вида (8) для четырех комбинаций состояний $\{s_1 s_1\}$, $\{s_1 s_2\}$, $\{s_2 s_1\}$, $\{s_2 s_2\}$:

$$\Pi = \begin{matrix} & \begin{matrix} s_1 s_1 & s_1 s_2 & s_2 s_1 & s_2 s_2 \end{matrix} \\ \begin{matrix} s_1 s_1 \\ s_1 s_2 \\ s_2 s_1 \\ s_2 s_2 \end{matrix} & \begin{bmatrix} \pi_{111} & \pi_{112} & 0 & 0 \\ 0 & 0 & \pi_{121} & \pi_{122} \\ \pi_{211} & \pi_{212} & 0 & 0 \\ 0 & 0 & \pi_{221} & \pi_{222} \end{bmatrix} \end{matrix}, \quad (8)$$

где $\pi_{ijn} = P(s_n^{(k)} s_i^{(k-1)} | s_i^{(k-1)} s_j^{(k-2)})$, $i, j, n = \overline{1, 2}$.

Нулевые элементы матрицы (8) соответствуют вероятностям невозможным событиям.

Подобным образом можно получить матрицы переходных вероятностей для общего случая ($m > 2$, $l > 2$). Для элементов матриц вида (8) должны соблюдаться модифицированные условия нормировки

$$\sum_{n=1}^l \pi_{rv\dots in} = 1, \quad i, j = \overline{1, l} \quad (9)$$

и согласованности

$$p_{v\dots in} = \sum_{r=1}^l p_{rv\dots i} \pi_{rv\dots in}, \quad v, \dots, i, n = \overline{1, l}, \quad (10)$$

где $p_{v\dots in} = P(s_n^{(k)} s_i^{(k-1)} \dots s_v^{(k-m+1)})$ – безусловная вероятность комбинации из m состояний $\{s_v \dots s_i s_n\}$.

В работе [7] показано, что для введенной таким образом простой цепи Маркова соблюдаются разностные уравнения

$$\mathbf{p}^{(k)} = \mathbf{p}^{(k-1)} \Pi = \mathbf{p}^{(0)} \Pi^k, \quad (11)$$

где $\mathbf{p}^{(k)}$ – вектор безусловных вероятностей всех возможных комбинаций из m состояний на k -м шаге. Например, для $m = 2$

$$\mathbf{p}^{(k)} = \left[P(s_1^{(k)} s_1^{(k-1)}), P(s_1^{(k)} s_2^{(k-1)}), \dots, P(s_i^{(k)} s_j^{(k-1)}), \dots, P(s_l^{(k)} s_l^{(k-1)}) \right]. \quad (12)$$

Рассмотрим метод марковской параметризации фрагментов речевых сигналов, использующий описанную модель.

4. Метод параметризации цепью Маркова

Метод параметризации фрагмента речевого сигнала цепью Маркова связности m содержит следующую последовательность шагов.

1. Битовые последовательности ${}^i \mathbf{s}$ ($i = \overline{1, N}$) объединяются в непересекающиеся массивы $\left[{}^b \mathbf{s}, {}^{b+1} \mathbf{s}, \dots, {}^{b+N_b-1} \mathbf{s} \right]^T$ размера $N_b \times L$.

Обозначим

$${}^b \mathbf{S}^{(k)} = \begin{pmatrix} {}^b s^{(k)} \\ \vdots \\ {}^{b+N_b-1} s^{(k)} \end{pmatrix}, \quad b = \overline{1, N / N_b}. \quad (13)$$

Будем полагать, что последовательность

$${}^b \mathbf{S} = \left\{ {}^b \mathbf{S}^{(1)}, {}^b \mathbf{S}^{(2)}, \dots, {}^b \mathbf{S}^{(L)} \right\} \quad (14)$$

является многосвязной цепью Маркова связности m с количеством состояний равным 2^{N_b} . Преобразуем многосвязную цепь (14) в простую цепь Маркова, с количеством состояний равным 2^{mN_b} формируя векторы (5)

$${}^b \vec{\mathbf{S}}^{(k)} = \left({}^b s^{(k)}, {}^b s^{(k-1)}, \dots, {}^b s^{(k-m+1)} \right) \quad (15)$$

и векторные последовательности

$${}^b \vec{\mathbf{S}} = \left\{ {}^b \vec{\mathbf{S}}^{(k)}, {}^b \vec{\mathbf{S}}^{(k-1)}, \dots, {}^b \vec{\mathbf{S}}^{(L)} \right\}. \quad (16)$$

2. Последовательности ${}^b \vec{\mathbf{S}}$ разделяется на p пересекающихся сегментов ${}^b_1 \vec{\mathbf{S}}, \dots, {}^b_p \vec{\mathbf{S}}$ с коэффициентом перекрытия h .

3. Для каждого q -го сегмента последовательности ${}^b \vec{\mathbf{S}}$ производится оценка элементов ${}^b_q \pi_{rv\dots in}$ матрицы вероятностей переходов ${}^b_q \Pi$

$$\begin{aligned} {}^b_q \pi_{rv\dots in} &= P \left({}^b_q S_n^{(k)} \mid {}^b_q S_i^{(k-1)}, \dots, {}^b_q S_r^{(k-m)} \right) = \\ &= \frac{\hat{P} \left({}^b_q S_n^{(k)}, {}^b_q S_i^{(k-1)}, \dots, {}^b_q S_r^{(k-m)} \right)}{\hat{P} \left({}^b_q S_i^{(k-1)}, \dots, {}^b_q S_r^{(k-m)} \right)} = \frac{P \left({}^b_q \vec{\mathbf{S}}^{(k)}, {}^b_q \vec{\mathbf{S}}^{(k-1)} \right)}{P \left({}^b_q \vec{\mathbf{S}}^{(k-1)} \right)}, \end{aligned} \quad (17)$$

где $b = \overline{1, N/N_b}$, $q = \overline{1, p}$, $k = \overline{1, L}$, $r, v, \dots, i, n = \overline{1, 2^{N_b}}$; $\hat{P} \left({}^b S_n^{(k)}, {}^b S_i^{(k-1)}, \dots, {}^b S_r^{(k-m)} \right)$ – относительная частота комбинаций $\left\{ {}^b S_n^{(k)}, {}^b S_i^{(k-1)}, \dots, {}^b S_r^{(k-m)} \right\}$ в q -м сегменте последовательности ${}^b S$; $\hat{P} \left({}^b S_i^{(k-1)}, \dots, {}^b S_r^{(k-m)} \right)$ – относительная частота комбинации состояний $\left\{ {}^b S_i^{(k-1)}, \dots, {}^b S_r^{(k-m)} \right\}$ в q -м сегменте последовательности ${}^b S$.

Полученные матрицы вероятностей переходов ${}^b \Pi$ объединяются в матрицу параметров сигнала

$$\mathbf{M} = \begin{bmatrix} {}^0_1 \Pi & {}^0_2 \Pi & \dots & {}^0_p \Pi \\ \dots & \dots & \dots & \dots \\ {}^{N/N_b}_1 \Pi & {}^{N/N_b}_2 \Pi & \dots & {}^{N/N_b}_p \Pi \end{bmatrix}. \quad (18)$$

Поскольку матрица (18) является разреженной, в качестве параметров фрагмента речевого сигнала достаточно использовать только ее значащие элементы.

Пример 1. $m = 2, N = 4, N_b = 1$.

1. Из фрагмента речевого сигнала выделяются бинарные битовые последовательности ${}^b s$, $b = \overline{0, 3}$ (рис.1).

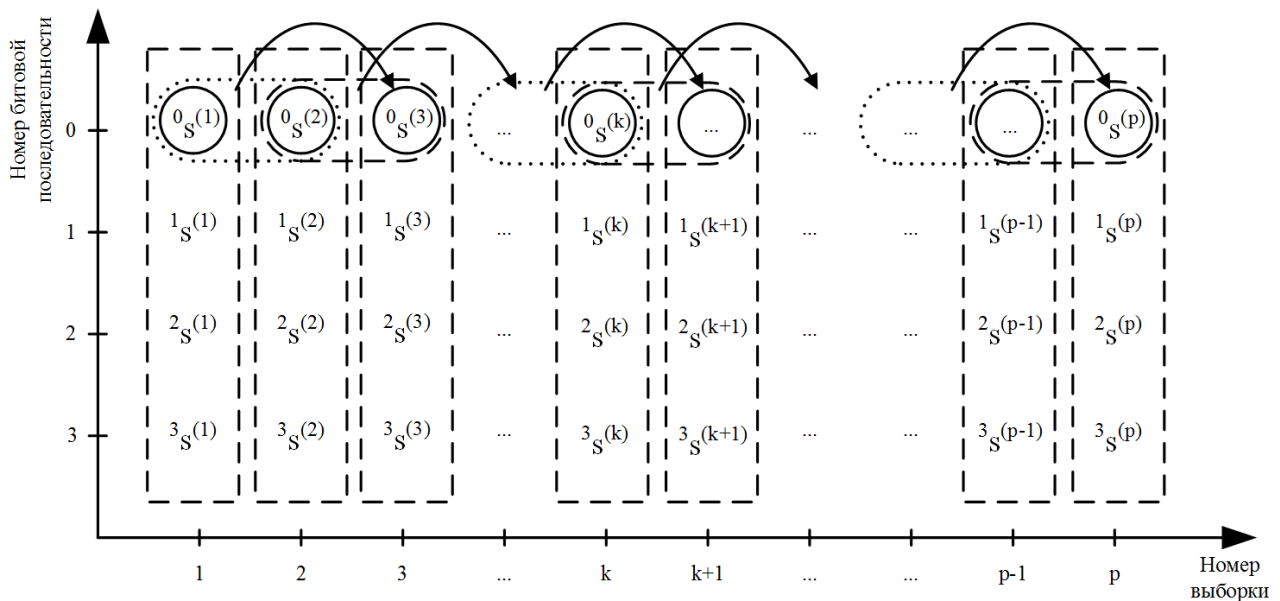


Рис.1. Эпюры переходов в многосвязной цепи Маркова при $m = 2$ и $N_b = 1$.

Полагая, что последовательности ${}^b s$ являются цепями Маркова связности $m = 2$ с количеством состояний равным двум, преобразуем их в простые цепи Маркова, с количеством состояний равным четырем, формируя векторы

$${}^b \vec{S}^{(k)} = ({}^b S^{(k)}, {}^b S^{(k-1)}), \quad (19)$$

где ${}^b S^{(k)} \equiv {}^b s^{(k)}$ и векторные последовательности ${}^b \vec{S}$ (16).

Диаграмма возможных переходов между метасостояниями в таких последовательностях представлена на рис.2.

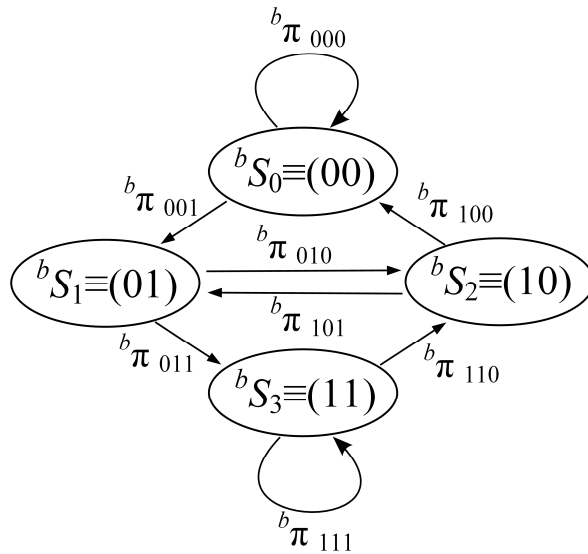


Рис. 2. Диаграмма переходов в последовательностях ${}^b \vec{S}$ при $m = 2$.

2. Последовательности ${}^b \vec{S}$ разделяются на p пересекающихся сегментов с коэффициентом перекрытия h : ${}^b \vec{S}_1, \dots, {}^b \vec{S}_p$.

3. Для каждого q -го сегмента последовательности ${}^b \vec{S}$ производится оценка элементов ${}^b \pi_{jin}$ матриц вероятностей переходов ${}^b \Pi$

$$\begin{aligned} {}^b \pi_{jin} &= \frac{P\left({}^b \vec{S}^{(k)}, {}^b \vec{S}^{(k-1)}\right)}{P\left({}^b \vec{S}^{(k-1)}\right)} = P\left({}^b S_n^{(k)} \mid {}^b S_i^{(k-1)}, {}^b S_j^{(k-2)}\right) = \\ &= \frac{\hat{P}\left({}^b S_n^{(k)}, {}^b S_i^{(k-1)}, {}^b S_j^{(k-2)}\right)}{\hat{P}\left({}^b S_i^{(k-1)}, {}^b S_j^{(k-2)}\right)}, \end{aligned} \quad (20)$$

где $b = \overline{1, 4}$, $q = \overline{1, p}$, $k = \overline{1, L}$, $i, j, n = \overline{1, 2}$.

$${}^b_q\Pi = \begin{bmatrix} \boxed{{}^b_q\pi_{111}} & {}^b_q\pi_{112} & 0 & 0 \\ 0 & 0 & {}^b_q\pi_{121} & \boxed{{}^b_q\pi_{122}} \\ \boxed{{}^b_q\pi_{211}} & {}^b_q\pi_{212} & 0 & 0 \\ 0 & 0 & {}^b_q\pi_{221} & \boxed{{}^b_q\pi_{222}} \end{bmatrix}. \quad (21)$$

Нулевые элементы матрицы (21) соответствуют вероятностям невозможных переходов. Из примера нетрудно заметить, что невозможны переходы: ${}^b\vec{S}_0 \rightarrow {}^b\vec{S}_2$, ${}^b\vec{S}_0 \rightarrow {}^b\vec{S}_3$, ${}^b\vec{S}_0 \rightarrow {}^b\vec{S}_3$, ${}^b\vec{S}_1 \rightarrow {}^b\vec{S}_0$, ${}^b\vec{S}_1 \rightarrow {}^b\vec{S}_1$, ${}^b\vec{S}_2 \rightarrow {}^b\vec{S}_2$, ${}^b\vec{S}_2 \rightarrow {}^b\vec{S}_3$, ${}^b\vec{S}_3 \rightarrow {}^b\vec{S}_0$, ${}^b\vec{S}_3 \rightarrow {}^b\vec{S}_1$, где ${}^b\vec{S}_0 = ({}^bS_0^{(k)}, {}^bS_0^{(k-1)})$, ${}^b\vec{S}_1 = ({}^bS_1^{(k)}, {}^bS_0^{(k-1)})$, ${}^b\vec{S}_2 = ({}^bS_0^{(k)}, {}^bS_1^{(k-1)})$, ${}^b\vec{S}_3 = ({}^bS_1^{(k)}, {}^bS_1^{(k-1)})$.

Поскольку элементы матрицы вероятностей переходов (21) удовлетворяют условию нормировки

$$\sum_{n=1}^2 {}^b_q\pi_{jin} = 1, \quad j, i = \overline{1, 2}, \quad (22)$$

в качестве параметров сегмента ${}^b_q\vec{S}$ достаточно использовать только значения элементов ${}^b_q\pi_{jii}$.

В результате каждый сегмент ${}^b_q\vec{S}$ характеризуется матрицей параметров

$${}^b\mathbf{M} = \begin{bmatrix} {}^b_1\pi_{111} & {}^b_2\pi_{111} & \dots & {}^b_p\pi_{111} \\ {}^b_1\pi_{122} & {}^b_2\pi_{122} & \dots & {}^b_p\pi_{122} \\ {}^b_1\pi_{211} & {}^b_2\pi_{211} & \dots & {}^b_p\pi_{211} \\ {}^b_1\pi_{222} & {}^b_2\pi_{222} & \dots & {}^b_p\pi_{222} \end{bmatrix}. \quad (23)$$

4. Матрицы ${}^b\mathbf{M}$ (23) объединяются в итоговую матрицу параметров фрагмента речевого сигнала

$$\mathbf{M} = \begin{bmatrix} {}^0\mathbf{M} \\ \dots \\ {}^3\mathbf{M} \end{bmatrix}. \quad (24)$$

Пример 2. $m = 2, N = 4, N_b = 2.$

1. Бинарные битовые последовательности ${}^0\mathbf{s}, \dots, {}^3\mathbf{s}$ объединяются в последовательности ${}^b\mathbf{S}$ (рис.3), такие что

$$\begin{aligned}
 {}^0\mathbf{S} &= \left\{ \begin{pmatrix} {}^0s^{(1)} \\ {}^1s^{(1)} \end{pmatrix}, \begin{pmatrix} {}^0s^{(2)} \\ {}^1s^{(2)} \end{pmatrix}, \dots, \begin{pmatrix} {}^0s^{(L)} \\ {}^1s^{(L)} \end{pmatrix} \right\} = \\
 &= \{ {}^0s^{(1)}, {}^0s^{(2)}, \dots, {}^0s^{(L)} \}, \\
 {}^1\mathbf{S} &= \left\{ \begin{pmatrix} {}^2s^{(1)} \\ {}^3s^{(1)} \end{pmatrix}, \begin{pmatrix} {}^2s^{(2)} \\ {}^3s^{(2)} \end{pmatrix}, \dots, \begin{pmatrix} {}^2s^{(L)} \\ {}^3s^{(L)} \end{pmatrix} \right\} = \\
 &= \{ {}^1s^{(1)}, {}^1s^{(2)}, \dots, {}^1s^{(L)} \}.
 \end{aligned}
 \tag{25}$$

Количество метасостояний ${}^bS^{(k)}$ таких последовательностей равно $2^{N_b} = 4.$

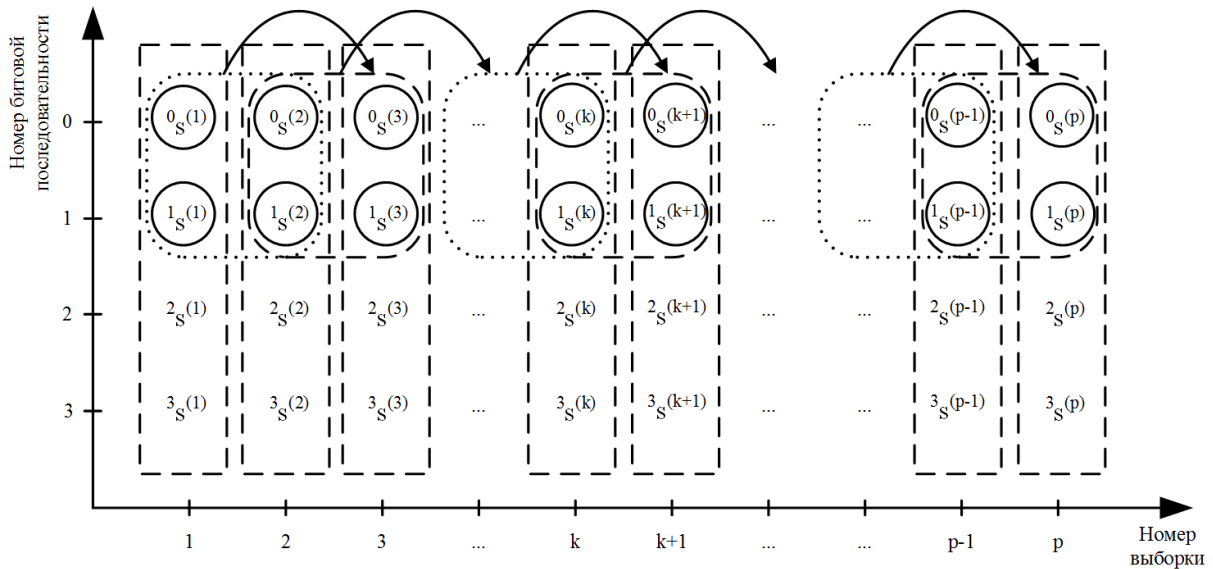


Рис.3. Эпюры переходов в многосвязной цепи Маркова при $m = 2$ и $N_b = 2.$

Полагая, что последовательности ${}^b\mathbf{S}$ являются многосвязными цепями Маркова связности $m = 2$ и количеством состояний равным четырем

$${}^bS_1 = \begin{pmatrix} {}^bS_1 \\ {}^bS_1 \end{pmatrix}, {}^bS_2 = \begin{pmatrix} {}^bS_1 \\ {}^bS_2 \end{pmatrix}, {}^bS_3 = \begin{pmatrix} {}^bS_2 \\ {}^bS_1 \end{pmatrix}, {}^bS_4 = \begin{pmatrix} {}^bS_2 \\ {}^bS_2 \end{pmatrix},
 \tag{26}$$

преобразуем их в простые цепи Маркова, с количеством состояний равным $2^{mN_b} = 16$ формируя векторы

$${}^b \vec{S}^{(k)} = \left({}^b S^{(k)}, {}^b S^{(k-1)} \right) \quad (27)$$

и векторные последовательности ${}^b \vec{S}$ (16).

2. Последовательности ${}^b \vec{S}$ разделяются на p пересекающихся сегментов с коэффициентом перекрытия h : ${}^1 \vec{S}, \dots, {}^p \vec{S}$ и ${}^2 \vec{S}, \dots, {}^p \vec{S}$.

3. Для каждого q -го сегмента последовательности ${}^b \vec{S}$ производится оценка элементов ${}^b \pi_{jin}$ матриц вероятностей переходов ${}^b \Pi$

$$\begin{aligned} {}^b \pi_{jin} &= \frac{P\left({}^b \vec{S}^{(k)}, {}^b \vec{S}^{(k-1)}\right)}{P\left({}^b \vec{S}^{(k-1)}\right)} = P\left({}^b S_n^{(k)} \mid {}^b S_i^{(k-1)}, {}^b S_j^{(k-2)}\right) = \\ &= \frac{\hat{P}\left({}^b S_n^{(k)}, {}^b S_i^{(k-1)}, {}^b S_j^{(k-2)}\right)}{\hat{P}\left({}^b S_i^{(k-1)}, {}^b S_j^{(k-2)}\right)}, \end{aligned} \quad (28)$$

где $b = \overline{1, 2}$, $q = \overline{1, p}$, $k = \overline{1, L}$, $j, i, n = \overline{1, 4}$.

4. Ненулевые элементы ${}^b \pi_{jin}$ матриц вероятностей переходов ${}^b \Pi$ объединяются в матрицу параметров сигнала \mathbf{M} .

5. Эксперимент

Рассмотренные в статье модели речевых сигналов использованы для оценки эффективности методов марковской параметризации при решении задачи дикторозависимого распознавания речевых команд. В качестве альтернативного метода параметризации речевых сигналов применен метод параметризации мел-кепстральными коэффициентами (MFCC).

Для эксперимента сформирована авторская коллекция из 150 речевых команд с частотой дискретизации 8 кГц.

Эксперимент заключался в выполнении следующих шагов:

1) оценке среднего времени параметризации одного фрагмента речевого сигнала;

2) оценке вероятности правильного распознавания команд в результате выполнения 1000 опытов для каждой модели (табл.1).

Результаты эксперимента представлены в табл.1. Для обозначения модели используемой при параметризации введены следующие сокращения: общее название модели dtmc (Discrete Time Markov Chain), следующие две цифры обозначают порядок цепи Маркова – m и количество группируемых битовых последовательностей – N_b . В скобках указаны номера старших бит, используемых при параметризации.

Таблица 1 – Вероятность распознавания и время параметризации

	Метод параметризации	Среднее время параметризации одного фрагмента, мкс.	Относительное время параметризации одного фрагмента	Вероятность распознавания, %
Простая цепь	dtmc11(0)	37,7	0,052	76,3
	dtmc11(0,1)	40,4	0,055	83,5
	dtmc11(0,1,2)	42,5	0,058	84,2
	dtmc11(0,1,2,3)	45,0	0,062	84,7
	dtmc12(0,1)	37,4	0,051	85,3
	dtmc12(0,1,2,3)	40,3	0,055	89,7
Сложная цепь	dtmc21(0)	37,9	0,052	83,4
	dtmc21(0,1)	40,4	0,055	92,4
	dtmc21(0,1,2)	42,7	0,059	94,8
	dtmc21(0,1,2,3)	44,9	0,062	95,8
	dtmc22(0,1)	37,9	0,052	85,6
	dtmc21(0,1,2,3)	40,4	0,055	93,9
	MFCC	728,1	1	99,7

Эксперимент показал (табл.1), что компромисс между эффективностью распознавания речевых команд и вычислительной сложностью алгоритма параметризации достигается при использовании (для оценки параметров аппроксимирующей цепи Маркова) от одного до четырех старших бит равномерно квантованного речевого сигнала.

Полученные результаты (табл. 1) свидетельствуют о значительном (в 16-19 раз) снижении временных затрат на параметризацию фрагментов речевых сигналов при сопутствующем уменьшении вероятности распознавания команд на 5-8% относительно классического метода параметризации мелкепестральными коэффициентами.

Выводы

В работе обобщены варианты представления фрагментов речевых сигналов простыми и сложными (многосвязными) цепями Маркова. Представлены частные случаи реализации данных моделей при параметризации речевых сигналов многосвязными цепями Маркова связностью $m = 1$ и $m = 2$.

Показано, что при использовании методов марковской параметризации в задачах автоматического дикторозависимого распознавания речевых команд (АРК) целесообразно применение моделей многосвязных цепей Маркова. Рассмотренные модели речевых сигналов позволяют значительно снизить требования к вычислительным ресурсам систем АРК.

Литература

1. Рабинер, Л.Р. Цифровая обработка речевых сигналов / Рабинер Л.Р., Шафер Р.В. // Пер. с англ. Под ред. Прохорова Ю.Н., Назарова М.В. – М.: Радио и связь, 1981. – 496 с.
2. Huang, X. Spoken Language Processing: A guide to theory, algorithm, and system development / X. Huang, A.Acerо, H.Hon. // Prentice Hall. – 2001.

3. Zheng. Comparison Of Different Implementations Of MFCC / F. Zheng, G. Zhang, Z. Song // Computer Science & Technology, 16(6): 2001. - pp. 582-589.
4. Picone, J.W. Signal modeling techniques in speech recognition / proceedings of the IEEE, September 1993, pp. 1215-1247.
5. Венедиктов М.Д. Дельта-модуляция. Теория и применение / Венедиктов М.Д., Женеvский Ю.П., Марков В.В. - М.: Связь, 1976. С. 104-114.
6. Плетнев К.В. Анализ метода марковской параметризации речевых сигналов / Плетнев К.В., Прозоров Д.Е. // Информационные системы и технологии, 2014. – №1(81). – С. 24-29.
7. Яншин В.В. Многосвязные цепи Маркова и их свойства // Радиотехника и электроника, Наука. 1993. – Том 38. – № 6 – С. 1081-1091.