

А.А. Рыболовлев, О.О. Басов

## АЛГОРИТМ ИДЕНТИФИКАЦИИ КАНАЛА СВЯЗИ, ОПИСЫВАЕМОГО МОДЕЛЬЮ ГИЛБЕРТА

Рассмотрена возможность оценивания параметров канала связи, описываемого моделью Гилберта, в паузах речевого сигнала. Получены выражения, связывающие параметры потока ошибок, действующего в канале, с их оценками, полученными в периодических паузах речевого сигнала. С целью проверки проведено имитационное моделирование, по результатам которого предложен алгоритм идентификации канала связи в паузах речевого сигнала. Погрешность определения параметров канала связи относительно истинных значений не превышает 13,2 %.

**Введение.** Одной из причин, снижающих качество речевого сигнала (РС) при передаче его по декаметровым каналам связи, являются канальные ошибки. Для таких каналов, как это следует из отечественных и зарубежных работ, достаточную точность описания при приемлемых вычислительных затратах можно получить с использованием модели Гилберта [1].

Адаптивная система кодирования (АСК) РС, согласно предложенной методике проектирования [2], в процессе обмена речевой информацией должна идентифицировать канал связи, описываемый выбранной моделью, и изменять свои параметры в зависимости от текущих оценок параметров модели.

Целью настоящей работы является разработка алгоритма оценивания с заданной точностью параметров дискретного канала связи (ДКС), описываемого моделью Гилберта, в паузах РС.

### 1. Формулировка гипотезы о возможности идентификации дискретного канала связи Гилберта в паузах РС

Для оценки качества ДКС в реальном времени в адаптивных системах кодирования и передачи цифровой информации наиболее предпочтительным является тестовый метод [3]. Так в работе [4] предложен алгоритм оценки параметров модели Гилберта по статистике средних длин состояний канала, обеспечивающий лучшую точность идентификации модели по сравнению с известными алгоритмами. Определен необходимый объем испытаний для получения заданной точности оценок:  $Q > 500$  смен состояний канала или  $Q_{дз} \approx 50000$  двоичных элементов. Оценка времени  $t_{ид}$  идентификации канала при различных скоростях  $R$  эффективного кодирования для предложенного алгоритма представлена на рисунке 1.

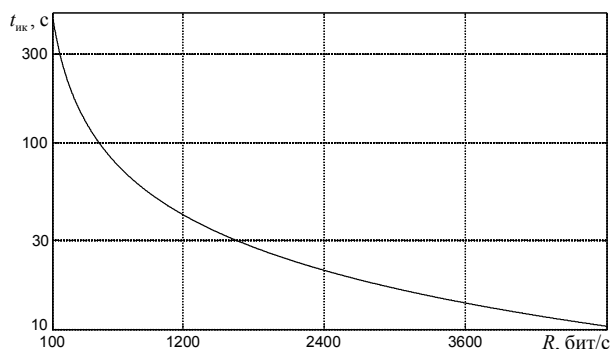


Рисунок 1

Нетрудно видеть, что даже при скорости кодирования  $R = 4.8$  кбит/с для идентификации канала необходимо  $t_{ид} \approx 10$  с. Данный факт указывает на недостаток тестового метода – влияние на пропускную способность.

Как показано в работах [5...8], относительным постоянством обладает процесс появления пауз речи длительности  $t_{п} = 10 - 60$  мс каждые  $t_{мп} = 200 - 250$  мс. Наличие такой периодичности позволяет сформулировать гипотезу о возможности идентификации ДКС Гилберта в паузах РС.

Пусть в паузах длительностью  $t_{п}$  передаются известные приемной части АСК тестовые последовательности длиной  $x$  двоичных элементов. Период  $t_{мп}$  следования пауз данной длительности обозначим как  $Z$  двоичных элементов.

Внутри каждой тестовой последовательности  $x$  элементы идут подряд, и следовательно, на данную последовательность будет влиять поток ошибок с исходными параметрами модели Гилберта. Последний элемент текущей тестовой последовательности и первый элемент следующей разделены промежутком в  $Z - x$  единичных интервалов, что аналогично перемежению с глу-

биной  $Z - x + 1$ , а значит для этой пары будут справедливы следующие выражения [4]:

$$P_{gb}^* = P_b (1 - M^\mu), \quad (1)$$

$$P_{bg}^* = P_g (1 - M^\mu), \quad (2)$$

где  $P_{gb}^*, P_{bg}^*$  – переходные вероятности модифицированного в результате перемежения канала связи;  $P_g, P_b$  – предельные вероятности нахождения исходного канала в хорошем и плохом состоянии соответственно;  $M$  – коэффициент группирования;  $\mu$  – глубина перемежения.

В среднем на каждую  $(x - 1)$  пару элементов исходной последовательности, идущих подряд, приходится одна пара элементов, разделенных промежутком  $Z - x$ .

Окончательно можно сформулировать следующую гипотезу: с вероятностью  $(x - 1)/x$  выходной процесс будет описываться исходными параметрами, а с вероятностью  $1/x$  – модифицированными:

$$\begin{cases} P_{gb}^* = \frac{x-1}{x} P_{gb} + \frac{1}{x} P_b (1 - M^{Z-x+1}) \\ P_{bg}^* = \frac{x-1}{x} P_{bg} + \frac{1}{x} P_g (1 - M^{Z-x+1}) \end{cases}, \quad (3)$$

где  $P_{gb}, P_{bg}$  – переходные вероятности исходного канала.

Значимость выражений, аналогичных (3), для случая посимвольного временного разделения каналов и посимвольного перемежения оценена в [4].

В нашем случае модифицированные параметры модели Гилберта могут быть определены по статистике средних длин состояний просеянного потока ошибок. Под просеянным будем понимать поток ошибок, полученный из исходного потока ошибок путем взятия окон длительностью  $x$  через каждые  $Z$  элементов.

Для определения истинных параметров исходного канала необходимо выразить их через модифицированные  $P_{gb}^*$  и  $P_{bg}^*$ .

Рассмотрим множитель  $(1 - M^{Z-x+1})$ , входящий в выражения (3). Оказывается, даже при большом коэффициенте группирования  $M$ , для различных значений  $t_{п}$  и  $t_{мп}$ , имеющем место в речевом сигнале, в большинстве случаев:

$$(1 - M^{Z-x+1}) \approx 1. \quad (4)$$

Для строгости вывода введем следующее условие:

$$Z - x > 40, \quad (5)$$

приводящее к выполнению равенства (4).

С учетом (4) выражения (3) примут следующий вид:

$$\begin{cases} P_{gb}^* = \frac{x-1}{x} P_{gb} + \frac{1}{x} P_b, \\ P_{bg}^* = \frac{x-1}{x} P_{bg} + \frac{1}{x} P_g. \end{cases} \quad (6)$$

Раскрыв предельные вероятности  $P_g, P_b$  [1] и введя обозначение  $P = P_{bg} + P_{gb}$ , преобразуем выражение (6) к следующему виду:

$$\begin{cases} xP_{gb}^* = P_{gb} \left( x - 1 + \frac{1}{P} \right), \\ xP_{bg}^* = P_{bg} \left( x - 1 + \frac{1}{P} \right). \end{cases} \quad (7)$$

Выразив  $P$  через известные переменные, для параметров исходного канала окончательно получаем:

$$\begin{cases} P_{gb} = \frac{xP_{gb}^*}{x - 1 + \frac{1}{x(P_{gb}^* + P_{bg}^*)} - 1}, \\ P_{bg} = \frac{xP_{bg}^*}{x - 1 + \frac{1}{x(P_{gb}^* + P_{bg}^*)} - 1}. \end{cases} \quad (8)$$

В связи с тем, что модифицированные переходные вероятности  $P_{gb}^*, P_{bg}^*$  будут определяться по статистике средних длин состояний, целесообразно выразить переходные вероятности исходного канала  $P_{gb}, P_{bg}$  (8) через средние длины состояний  $D_g^*, D_b^*$  [1] просеянного потока ошибок:

$$\begin{cases} P_{gb} = \frac{x/D_g^*}{x - 1 + \frac{1}{x(1/D_g^* + 1/D_b^*)} - 1}, \\ P_{bg} = \frac{x/D_b^*}{x - 1 + \frac{1}{x(1/D_g^* + 1/D_b^*)} - 1}. \end{cases} \quad (9)$$

Проведем проверку полученных соотношений (9), используя имитационное моделирование.

## 2. Схема и результаты имитационного моделирования

**Шаг 1.** Задавая параметры исходного дискретного канала, генерируем случайный поток ошибок. Один из вариантов алгоритма генерации ошибок в соответствии с моделью Гилберта достаточно подробно рассмотрен в [1].

**Шаг 2.** В соответствии с алгоритмом, основанным на методике оценки параметров модели Гилберта [4], находим средние значения длин

хорошего  $D_g^{исх}$  и плохого  $D_b^{исх}$  состояний исходного потока ошибок.

**Шаг 3.** Просеиваем исходный поток ошибок через окна размером  $x$  и периодом  $Z$ .

**Шаг 4.** Для просеянного потока ошибок в соответствии с известным алгоритмом [4] определяем средние длины состояний  $D_g^*$  и  $D_b^*$ .

**Шаг 5.** По выражениям (9) находим параметры  $P_{gb}$  и  $P_{bg}$  исходного потока ошибок и выражаем через них средние длины

$$D_g^{расч} = 1/P_{gb}; D_b^{расч} = 1/P_{bg}. \quad (10)$$

**Шаг 6.** При совпадении рассчитанных параметров  $D_g^{расч}$  и  $D_b^{расч}$  с исходными  $D_g^{исх}$  и  $D_b^{исх}$ , выражения (8), (9) можно использовать для идентификации исходного канала по просеянному потоку ошибок.

Результаты имитационного моделирования по указанной схеме для различных значений  $P_{gb}, P_{bg}, x, Z$  и объема исходного потока ошибок  $Q_{дз} \approx 50000$  двоичных элементов представлены в табл. 1, 2.

**Таблица 1**

$x$	$Z$	$P_{gb} = 0.005; P_{bg} = 0.07$			
		$D_g^{исх}$	$D_g^{расч}$	$D_b^{исх}$	$D_b^{расч}$
50	100	176	180	13.7895	13.8931
75	120	201	198	14.0212	13.9022
100	150	198	196	14.1367	14.0988
125	170	210	207	14.285	14.2363
150	200	202	200	14.2903	14.2815
175	220	187	183	14.2873	14.2785
200	250	198	195	14.2886	14.2903
225	270	200	199	14.3154	14.2992
250	300	195	193	14.2891	14.2786
300	350	205	203	14.333	14.2745
$\bar{S}_x^2$		70.0444		0.0260	
$\bar{S}_{хост}^2$		9.5		0.0041	
$F$		7.3731		6.3674	

Для получения количественной оценки значимости предложенных выражений используем  $F$ -критерий Фишера [9]. В соответствии с данным критерием качество аппроксимации определяется путем сравнения остаточной дисперсии  $\bar{S}_{хост}^2$  с общей  $\bar{S}_x^2$ :

$$\bar{S}_{хост}^2 = \frac{1}{n-2} \sum_{i=1}^n \left[ (D_{...}^{расч})_i - (D_{...}^{исх})_i \right]^2; \quad (11)$$

$$\bar{S}_x^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n (D_{...}^{расч})_i^2 - \frac{1}{n} \left( \sum_{i=1}^n (D_{...}^{расч})_i \right)^2 \right], \quad (12)$$

**Таблица 2**

$x$	$Z$	$P_{gb} = 0.14; P_{bg} = 0.11$			
		$D_g^{исх}$	$D_g^{расч}$	$D_b^{исх}$	$D_b^{расч}$
50	100	7.6974	7.5238	10.207	10.181
75	120	7.4581	7.3994	9.8943	9.8713
100	150	7.3117	7.2932	9.7334	9.7197
125	170	7.2986	7.2614	9.691	9.4909
150	200	7.2428	7.0618	9.6168	9.5608
175	220	7.2035	7.1004	9.7232	9.6580
200	250	7.2367	7.1843	9.618	9.5479
225	270	7.1914	7.1585	9.7026	9.6683
250	300	7.2044	7.1264	9.5189	9.5038
300	350	7.1837	7.1061	9.7012	9.66
$\bar{S}_x^2$		9.7654		0.0431	
$\bar{S}_{хост}^2$		0.0118		0.0071	
$F$		825.5404		6.0663	

т.е. отношением

$$F = \bar{S}_x^2 / \bar{S}_{хост}^2. \quad (13)$$

Считается, что уравнение аппроксимации предсказывает результаты опытов лучше среднего, если  $F > F_p(f_1, f_2)$ , где  $F_p(f_1, f_2)$  - критическое значение  $F$ -статистики на уровне значимости  $p$  при  $f_1 = n-1$  и  $f_2 = n-2$  степенях свободы ( $n$  - количество точек сравнения).

При  $p = 0.99$  из аппроксимации для  $F$ -распределения [9] находим  $F_{0.99}(9,8) = 5.9538$ , следовательно, для всех рассмотренных случаев уравнения (8) описывают значения параметров исходного канала с более чем 99 %-ной достоверной вероятностью.

**3. Алгоритм идентификации канала связи, описываемого моделью Гилберта**

Вывод о значимости выражений (8) позволяет модифицировать предложенный в [4] алгоритм оценки параметров модели Гилберта по статистике средних длин состояний канала следующим образом (рисунок 2).

В алгоритме использованы следующие обозначения:

$\tilde{k}$  - максимальное расстояние между ошибками, принадлежащими одному пакету;

$\rho$  - величина, позволяющая исключить методическую погрешность в определении длины плохого состояния;

$g$  - указатель индекса в массивах длин состояний  $D_g$  и  $D_b$ ;

$S$  - длина плохого состояния;

$k$  - указатель длины блока безошибочных элементов;

$S_{er}$  – поток (массив) ошибок;  
 $l$  – указатель начального состояния в массиве ошибок;  
 $last(\dots)$  – индекс последнего элемента массива;  
 $M\{\dots\}$  – математическое ожидание случайной величины.

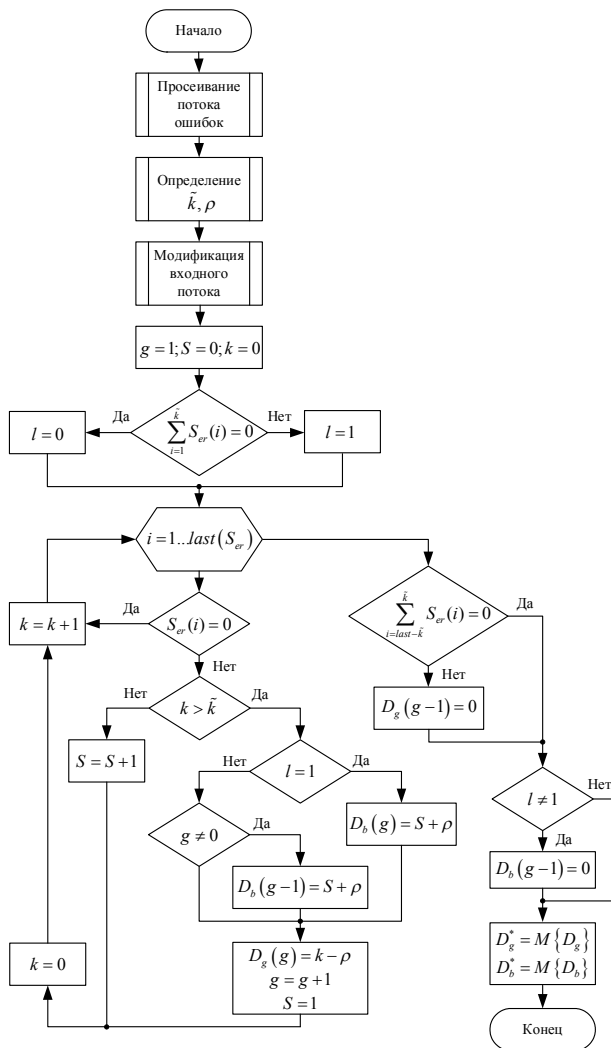


Рисунок 2

Сущность процесса просеивания исходного потока ошибок пояснена выше. Принципиальным моментом является нахождение границ состояний в анализируемом потоке [4].

Под пакетом ошибок следует понимать участок потока, начинающийся и оканчивающийся ошибочным элементом, внутри которого расстояние между ошибками не превышает величины  $\tilde{k}$ .

Если плохое состояние канала Гилберта рассматривать как пакет, то возможно искусственное занижение величины  $D_b$ , что, в свою очередь, приводит к искажениям оценок всех остальных параметров модели. Для исключения

данной методической погрешности необходимо увеличивать длину каждого пакета на величину  $\rho$ .

Для определения значений величин  $\tilde{k}$  и  $\rho$  необходимо рассмотреть зависимость вероятности  $P(0^k | 1)$  того, что после ошибки будет не менее  $k$  безошибочных элементов. Данная зависимость характеризуется наличием двух участков: начального – характеризующего распределение длин серий безошибочных элементов в пакетах, и пологого – отражающего расстояния между пакетами. Указанные участки хорошо аппроксимируются прямыми, а точка их пересечения довольно удовлетворительно определяет величину

$$\tilde{k} = \lceil (b_2 - b_1) / (a_1 - a_2) \rceil, \quad (14)$$

где  $a_1, b_1, a_2, b_2$  – коэффициенты уравнений соответствующих прямых.

Величина  $\rho$  определяется выражением

$$\rho = \lfloor (b_2 - b_1) / (a_1 - a_2) \rfloor. \quad (15)$$

Модификация входного потока ошибок осуществляется в соответствии со следующими правилами:

- а) если сумма последних  $\tilde{k}$  элементов массива  $S_{er}$  равна 0 (ошибок нет), необходимо дописать одну единицу за последним элементом;
- б) если в последних  $\tilde{k}$  элементах есть хотя бы одна ошибка, то дописывается еще  $\tilde{k} + 1$  нулей и одна единица.

Далее в результате поэлементного анализа просеянный поток ошибок разделяется на хорошие и плохие состояния.

Результатом работы алгоритма (рисунок 2) являются средние значения длин состояний просеянного потока ошибок: хорошего  $D_g^*$  и плохого  $D_b^*$ , по которым определяются переходные вероятности исходного потока ошибок (9).

Проведем оценку погрешностей и доверительных интервалов найденных величин  $D_g^{расч}$  и  $D_b^{расч}$  (10) при значении доверительной вероятности  $\beta = 0,95$  [4].

Пусть величина  $h$  представляет собой оценку вероятности события по частоте в  $\varphi$  независимых опытах. Доверительный интервал в данном случае будет определяться выражением

$$\varepsilon_\beta = t_\beta \sqrt{h(1-h)/\varphi}. \quad (16)$$

Величина  $t_\beta$  определяется известным выражением [10]:

$$t_{\beta} = \arg \left[ \Phi \left( \frac{1 + \beta}{2} \right) \right], \quad (17)$$

и при  $\beta = 0,95$  составляет  $t_{\beta} = 1,96$ .

Для большей наглядности доверительный интервал (16) представим в процентах:

$$\Delta = \frac{\varepsilon_{\beta}}{\tilde{m}_{\dots}} \cdot 100\%, \quad (18)$$

где  $\tilde{m}_{\dots}$  - математическое ожидание оцениваемой величины  $D_g^{\text{расч}} (D_b^{\text{расч}})$ .

Указанные оценки определялись по результатам имитационного моделирования предложенного алгоритма (рисунок 2) в десяти сериях испытаний для различных параметров модели Гилберта. В каждой серии объем испытаний определялся числом состояний канала ( $Q = 50, 100, 500, 1000, 2000$ ), в зависимости от чего вычислялся объем двоичных элементов  $Q_{\text{дв}}$ . Анализ результатов испытаний позволил сделать следующие промежуточные выводы:

1) при  $Q > 500$  погрешность определения  $D_g^{\text{расч}}$  и  $D_b^{\text{расч}}$  относительно истинных значений, при которых генерировался поток ошибок, не превышает 18,9 %;

2) при  $Q > 1000$  указанная погрешность не превышает 13,2 %;

3) ширина доверительного интервала для  $D_g^{\text{расч}}$  не превышает 15,3 % при  $Q > 500$ , 11,2 % при  $Q > 1000$  и 8,4 % при  $Q > 2000$ ;

4) ширина доверительного интервала для  $D_b^{\text{расч}}$  не превышает 13,6 % при  $Q > 500$ , 10,5 % при  $Q > 1000$  и 7,9 % при  $Q > 2000$ .

### Выводы

Полученные количественные оценки позволяют сделать вывод о том, что предложенный алгоритм идентификации канала связи обеспечивает приемлемую для инженерных расчетов точность оценок параметров ДКС, описываемого моделью Гилберта при объеме испытаний  $Q > 1000$  смен состояний канала.

Результаты испытаний алгоритмов, представленного на рисунке 2 и предложенного в [4], показывают, что при больших объемах выборки

( $Q = 1000$ ,  $Q_{\text{дв}} \approx 50000$ ) точность обоих алгоритмов приблизительно одинакова. При уменьшении объема испытаний ( $Q = 500$ ) предлагаемый алгоритм несколько уступает в точности определения параметров модели Гилберта (на 1,7 % для  $D_g^{\text{расч}}$  и 2,7 % для  $D_b^{\text{расч}}$ ), однако, как и в первом случае, не влияет на пропускную способность системы передачи.

Использование предложенного алгоритма идентификации в АСК РС позволит изменять их структуру (алгоритм функционирования, параметры) с целью повышения эффективности функционирования.

### Библиографический список

1. Басов О.О., Рыболовлев А.А., Никитин В.В. Модель дискретного канала связи с группированием ошибок для адаптивных систем кодирования речевого сигнала / Труды Всероссийской научной конференции «Проектирование научных и инженерных приложений в среде MATLAB». – СПб.: Изд-во С.-Петербург. ун-та, 2007. — 1603 с.: ил. – С.1338-1353.
2. Басов О.О., Рыболовлев А.А. Методика проектирования адаптивных систем кодирования речевого сигнала. Информатизация и информационная безопасность правоохранительных органов. – М.: Академия управления МВД России, 2007. – С.407-412.
3. Головин О.В. Декаметровая радиосвязь. – М.: Радио и связь, 1990. – 240 с.: ил.
4. Мелентьев О.Г. Теоретические аспекты передачи данных по каналам с группирующимися ошибками / Под ред. проф. В.П. Шувалова. – М.: Горячая линия – Телеком, 2007. – 232 с.: ил.
5. Калинин Ю.К. Разборчивость речи в цифровых вокодерах. – М.: Радио и связь, 1991. – 220 с.: ил.
6. Михайлов В.Г., Златоустова Л.В. Измерение параметров речи / Под ред. М.А. Сапожкова. – М.: Радио и связь, 1987. – 168 с.: ил.
7. Сапожков М.А., Михайлов В.Г. Вокодерная связь. – М.: Радио и связь, 1983. – 248 с.: ил.
8. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи / Под ред. О.И. Шелухина. – М.: Радио и связь, 2000. – 456 с.: ил.
9. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006. – 816 с.
10. Венцель Е.С. Теория вероятностей. – М.: Наука, 1969. – 576 с.