АЛГОРИТМ ПОИСКА УНИКАЛЬНЫХ БИТОВЫХ КОМБИНАЦИЙ В ЦИФРОВЫХ ПОТОКАХ КОДИРОВАННЫХ РЕЧЕВЫХ СООБЩЕНИЙ С ИСПОЛЬЗОВАНИЕМ ПОЛИНОМИАЛЬНЫХ ПРИЗНАКОВ

Л.И. Двойрис, А.Н. Гомонов, Д.В. Герасин

ООО КИЦ «Техноценоз», Калининград, Российская Федерация

Описан принцип организации и функционирования алгоритма поиска уникальных битовых комбинаций, используемый в алгоритме идентификации типов вокодеров. Приведена структурно-логическая схема и блок-схема выполнения алгоритма.

Ключевые слова: алгоритм идентификации, вокодер, цифровой поток, полиномиальный признак, кодированное речевое сообщение, значение периода, декодер.

THE ALGORITHM OF SEARCHING UNIQUE BIT COMBINATIONS IN DIGITAL STREAMS OF THE CODED SPEECH MESSAGES BY USING POLYNOMIAL IDENTIFIERS

L.I. Dvoyris, A.N. Gomonov, D.V. Gerasin

Ltd KIC "Tekhnotsenoz", Kaliningrad, Russian Federation

The article proposed a concept for constructing an algorithm of searching unique bit combinations of algorithm of identification of vocoders types. The structurally logical scheme and flow-chart of the algorithm is presented.

Keywords: algorithm of identification, vocoder, digital steam, polynomial identifier, coded speech message, period value, decoder.

Важным этапом выполнения алгоритма идентификации типов вокодеров является поиск уникальных битовых комбинаций цифровых потоков (ЦП) кодированных речевых сообщений (КРС). Результат идентификации типа вокодера зависит от корректности организации и реализации указанного этапа анализа и обработки ЦП, а также от структуры ЦП (предназначен только для ЦП с явно выраженными особенностями структуры кадра).

Алгоритм поиска уникальных битовых комбинаций ЦП КРС должен удовлетворять следующим требованиям:

- точность поиска любых битовых последовательностей в различных ЦП,
- однозначность выявления заданных битовых комбинаций ЦП;
- точность выявления отличительных особенностей ЦП разных типов вокодеров и имеющих схожие битовые фрагменты;
- неизменность результата поиска от размера ЦП (от минимального до максимального);
- способность к обучению.

Для удовлетворения указанных требований разработана процедура поиска уникальных битовых комбинаций ЦП КРС, структурно-логическая схема которой представлена на рис. 1.

Все ЦП КРС предлагается условно разделить по расположению уникальных битовых комбинаций в ЦП на три категории (рис. 2):

- в столбцах кадра ЦП на всей длине (во всей структуре кадра ЦП);
- в столбцах кадра ЦП во время передачи паузы речи (в строках передачи паузы);
- в произвольных фрагментах ЦП.

Рис. 1. Структурно-логическая схема поиска уникальных битовых комбинаций ЦП КРС

Рис. 2. Структурные признаки ЦП:

а – во всей структуре кадра; *б* – во время передачи паузы; *в* – в произвольных фрагментах

Для описания битовых комбинаций предлагается использовать выражение полиномов (1) из [5]. Применительно к анализу и формату представления битовых последовательностей ЦП перепишем выражение (1) с изменением порядка следования переменных в соответствии с расположением бит в ЦП (от младшего к старшему) и вынесенным бинарным показателем b_n , получим выражение (2).

$$y = x_n^n + x_{n-1}^{n-1} + x_{n-2}^{n-2} + \dots + x_2^2 + x_1^1 + 1 \quad (1)$$

$$y = b_0 x^0 + b_1 x^1 + b_2 x^2 + \dots + b_{n-2} x^{n-2} + b_{n-1} x^{n-1} + b_n x^n \quad (2)$$

где $b_0 x^0$ – значение младшего (первого) бита цифрового потока; $b_n x^n$ – значение старшего (последнего) бита цифрового потока; b_n – бинарный показатель, принимающий два значения 0 и 1.

Сформируем критерии формирования полиномиальных признаков (ПП), используемых для поиска уникальных битовых комбинаций в ЦП КРС.

1. ПП должны иметь простую формулу (структуру), т. е. состоять из конечного числа элементов, обозначающих значение бита.

2. ПП должны выполнять точное описание и строгое соответствие каждому биту в битовой комбинации.

3. ПП должны описывать значительные (по количеству бит) битовые комбинации как в кратком, так и в полном представлении (форме записи).

4. ПП должны быть читабельными (просты к их чтению и составлению).

5. ПП должны иметь отчётливые рамки начала и конца.

1. Период поиска ПП (как правило, соответствует значению периода ЦП КРС; в отдельных случаях рекомендуется использовать двукратное значение периода; для вокодера G.726 рекомендуется использовать значения 2, 3, 4, 5 для соответствующих скоростей 16, 24, 32, 40 кбит/с).

2. Максимальное количество бит, на которое априорно может быть сдвинут ЦП (сдвиг относительно первого бита).

3. Количество уникальных битовых комбинаций, достаточное для идентификации типа вокодера (по умолчанию поиск ведётся до последнего бита ЦП).

4. Количество подряд зафиксированных уникальных битовых комбинаций, достаточных для идентификации типа вокодера (используется в паре (совместно) с предыдущим параметром для выявления структурных особенностей ЦП во время передачи паузы).

5. Допустимое количество бракованных (необнаруженных) уникальных битовых комбинаций, которое не влияет на результат идентификации (при передаче по каналу связи ЦП КРС может содержать искажённые битовые последовательности, например, ЦП вокодера G.723 ACELP во всей структуре кадра имеет первый бит – единичный, второй бит – нулевой, но встречаются исключения).

Для уточнения условий поиска уникальных битовых комбинаций предусмотрены следующие возможности для настройки параметров поиска ПП:

- указание значений параметров в относительных и абсолютных величинах, а также использование двух вариантов одновременно;

- использование наборов ПП (состоящих из нескольких ПП) для выявления фрагментов ЦП, присутствующих одновременно во всей структуре кадра и фрагментарно в определённых областях (фрагментах) ЦП (например, во время передачи паузы);

- обнаружение фрагментов ЦП с разворотом и без разворота байт;

- указание набора ПП для поиска как одного (одиночного) типа вокодера, так и одного вокодера из группы подтипов типа вокодера (например, вокодер SPEEX имеет более 30 подтипов, отличающихся скоростью передачи и способом кодирования речевой информации. Общий для всей группы подтипов вокодера SPEEX набор ПП имеет вид: 1) «B[X 217]» – для выявления уникальных битовых комбинаций, указывающих на передачу паузы; 2) «1[xxx01100] 2[xxxxxx1xx] last[1110xxxx]» – для выявления уникальных битовых комбинаций, присутствующих во всём ЦП. Отличительной особенностью каждого подтипа вокодера является значение первого байта: 17 для q2; 25 для q3; 33 для q5; 41 для q7; 49 для q9; 57 для q10 и т. д.).

Указанные параметры, условия и особенности поиска реализованы в предложенной процедуре поиска уникальных битовых комбинаций и позволяют описывать любые битовые последовательности ЦП. Например, для идентификации вокодера G.723 (ACELP) (рис. 6) необходимо использовать набор, состоящий из двух ПП:

1) «b[10]» – для поиска фрагмента битовой последовательности «[10]», расположенного в первом и втором битах каждой строки битовой структуры ЦП;

2) «B[245 58 138 6 10 4 0 0 0 0 224 254 232 143 254 232 143 170 170]» – для поиска фрагмента битовой последовательности, расположенного в месте передачи паузы.

Рис. 6. Фрагмент цифрового потока вокодера G.723 (период 160 бит) с параметрами поиска полиномиальных признаков

Таким образом, предложенные параметры поиска уникальных битовых комбинаций ЦП КРС и возможности для настройки параметров поиска ПП позволяют выполнять гибкую настройку алгоритма поиска уникальных битовых комбинаций для выявления битовых последовательностей ЦП любых типов вокодеров.

Алгоритм поиска уникальных битовых комбинаций представлен ниже (рис. 7).

Рис. 7. Алгоритм поиска уникальных битовых комбинаций в ЦП КРС

1. Выполняется загрузка ЦП КРС и идентификационных признаков (параметров поиска уникальных битовых комбинаций ЦП КРС, а также ПП (сигнатур) по указанному значению периода ЦП).

2. Выполняется поиск первой сигнатуры на указанном интервале (как правило, поиск первой сигнатуры осуществляется на интервале, соответствующем двойному значению периода, в отдельных случаях может быть задано как большее, так и меньшее значение интервала).

Позиция первого бита, на котором обнаружена сигнатура является началом (точкой) отчёта для поиска следующих битовых комбинаций по указанной сигнатуре, который выполняется через одинаковые интервалы (количество бит), равные или кратные значению периода.

Поиск продолжается до последнего бита ЦП или до указанного относительного или абсолютного значения строки в ЦП.

3. Шаг 2 повторяется для всех сигнатур, отобранных на шаге 1.

4. Выполняется анализ результатов поиска полиномиальных признаков.

Если ни одной сигнатуры не обнаружено, тип вокодера считается не идентифицированным.

Если выявлены сигнатуры, принадлежащие нескольким типам вокодеров, то тип вокодера считается идентифицированным. Значение типа вокодера выбирается в соответствии с сигнатурой, приоритет которой выше. Список приоритетности задаётся перед идентификацией. Если приоритет для оспариваемых типов вокодеров не задан, ЦП направляется на анализ с соответствующей пометкой. Далее выполняется корректировка правил приоритетности (обучается алгоритм поиска) и повторная идентификация.

Выбор ПП и составление правил поиска ПП выполняется таким образом, чтобы однозначно достиглось различие анализируемого ЦП от ЦП других типов вокодеров (ЦП других видов оборудования).

СПИСОК ЛИТЕРАТУРЫ

1. Бесекерский В.А., Попов Е.П. Теория систем автоматического управления. СПб.: Профессия, 2004. 752 с.
2. Волгин В.В., Каримов Р.Н. Оценка корреляционных функций в промышленных системах управления. М.: Энергия, 1979. 80 с.
3. Гинсберг К.С. Теория идентификации: стимулы, предпосылки и перспективы развития // Приборы и системы управления. 1996. № 12. С. 27–30.
4. Дейч А. Методы идентификации динамических объектов. М.: Энергия, 1979. 240 с.
5. Прасолов В.В. Многочлены. 2-е изд., стереотип. М.: МЦНМО, 2001. 336 с.: ил.