АВТОМАТИЧЕСКАЯ КЛАССИФИКАЦИЯ СИГНАЛЬНЫХ СОЗВЕЗДИЙ ЦИФРОВОЙ МОДУЛЯЦИИ ПО МИНИМУМУ РАССТОЯНИЯ КУЛЬБАКА-ЛЕЙБЛЕРА

Горячкин О.В.

Поволжская государственная академия информатики, радиотехники и связи,
443010, г. Самара, ул. Л. Толстого 23, тел. (8462) 391143,
E-mail: gor@mail.radiant.ru

Задача разработки алгоритмов автоматической классификации вида модуляции, актуальная проблема, возникающая при разработке универсальных демодуляторов и средств радиоконтроля, в настоящие время насчитывает значительное число подходов к своему решению.

Основные методы классификации вида цифровой модуляции можно разделить на две основные группы, различающиеся по большому счету способом построения пространства признаков классификатора: классификация вида цифровой модуляции по сигнальным созвездиям; классификация вида цифровой модуляции по сигнальным реализациям.

В докладе мы рассматриваем алгоритм классификации вида модуляции по сигнальным созвездиям основанный на минимизации расстояния Кульбака-Лейблера, эквивалентный алгоритму максимального правдоподобия. Алгоритм позволяет классифицировать различные виды цифровой линейной модуляции без памяти [1].

Линейный квадратурный прием обеспечивает регистрацию сигнального созвездия в виде последовательности комплексных отсчетов сигнала $\{z_1, z_2, ..., z_N\}$. В отсутствии ошибок синхронизации и замираний в канале каждый такой отсчет представляет собой сумму двух независимых случайных комплексных величин:

$$z_k = c_k + n_k,\ k = 1,...,N, \qquad (1)$$

где: $\{c_k\}$ - последовательность независимых случайных величин принимающих дискретные значения равновероятно из множества $C_j = \{c^1, c^2, ..., c^{M_j}\}$, которое мы будем называть созвездием j-го способа модуляции, $j = 1,...,K$; $\{n_k\}$ - последовательность отсчетов белого гауссовского шума с нулевым математическим ожиданием и дисперсией σ^2.

Созвездие PSK4 при различных отношениях сигнал/шум показаны на Рис.1.

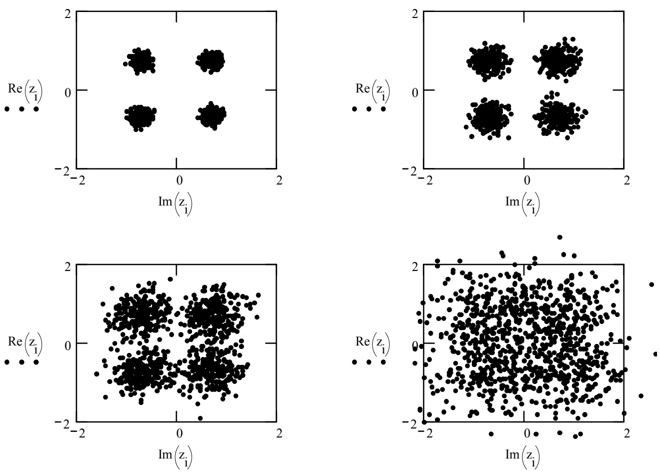

Рис.1. Созвездие PSK4 при отношении сигнал/шум = 5, 10, 15, 20 Дб соответственно.

Диаграмму, показанную на Рис.1 можно интерпретировать как случайное поле, полученное суммой N независимых случайных функций:

$$\xi(r) = \sum_{k=1}^{N} \delta\left(r - \left(c_k + n_k\right)\right) \qquad (2)$$

$$\delta(r) = \begin{cases} 1, r = 0 \\ 0, r \neq 0 \end{cases}$$

где: r - вектор на комплексной плоскости.

В общем виде алгоритм максимального правдоподобия сводится к следующей процедуре:

$$\hat{C} = \arg\left(\max_j \left(p\left(z_1, z_2, ..., z_N \middle| C_j \right)\right)\right) \quad (3)$$

В случае (1) мы можем записать:

$$\frac{1}{N}\log\left(p\left(z_1, z_2, ..., z_N \middle| C_j\right)\right) = \frac{1}{N}\sum_{i=1}^{N}\log\left(p\left(z_i \middle| C_j\right)\right) \quad (4)$$

Рассматривая предел функции правдоподобия при $N \to \infty$ и используя свойства асимптотической несмещенности оценки максимального правдоподобия можно записать следующее выражение, являющиеся неким аналогом функции неопределенности сигнальных созвездий:

$$F\left(C_i, C_j\right) = \mathbf{M}\left\{\ln\left(p\left(z \middle| C_j\right)\right)\right\} = \int p\left(z \middle| C_j\right)\log\left(p\left(z \middle| C_i\right)\right)dz \quad (5)$$

Выражение (5) означает также, что для больших N алгоритм максимального правдоподобия сводится к оценке условной плотности $\hat{p}\left(z \middle| C_i\right)$ в виде:

$$\hat{p}\left(z \middle| C_i\right) = \frac{1}{N}\sum_{k=1}^{N}\delta\left(z - z_k\right) \quad (6)$$

Алгоритм классификации соответственно к поиску распределения $p\left(z \middle| C_j\right)$ наиболее близкого к $\hat{p}\left(z \middle| C_i\right)$ с точки зрения расстояния Кульбака-Лейблера:

$$\hat{C} = \arg\left(\min_j \left(\int \hat{p}\left(z \middle| C_i\right) \cdot \log\left(\frac{\hat{p}\left(z \middle| C_i\right)}{p\left(z \middle| C_j\right)}\right)dz \right)\right) \quad (7)$$

Особенности геометрии пространства признаков классификатора, построенного по алгоритму (7) характеризует матрица расстояний Кульбака-Лейблера $\{\rho_{ij}\}$, которая в свою очередь зависит от параметров распределения шума.

$$\rho_{ij} = \int p\left(z \middle| C_i\right) \cdot \log\left(\frac{p\left(z \middle| C_i\right)}{p\left(z \middle| C_j\right)}\right)dz \quad (8)$$

В таблице показаны значения расстояний Кульбака-Лейблера для случая некоторых наиболее распространенных видов цифровой модуляции.

Условная плотность вероятности принятая в расчетах в соответствии с (1) имеет вид:

$$p\left(z \middle| C_j\right) = \left(\frac{1}{M}\right) \cdot \sum_{k=1}^{M_j}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}\left|z_i - c_j^k\right|^2\right) \quad (9)$$

Матрица расстояний Кульбака-Лейблера несимметрична и зависит от порядка перебора созвездий. По строкам таблицы 1 записана модуляция принимаемого сигнала, по столбцам перебираемые при классификации виды модуляции.

Значения матрицы расстояний Кульбака-Лейблера имеют глубокий смысл, устанавливаемый известной леммой Стейна [2], которая в наших обозначениях означает, что при решении задачи двухальтернативной классификации *i*-го и *j*-го сигнальных созвездий по критерию Неймана-Пирсона справедливо следующее асимптотическое свойство:

$$\lim_{q_{ij} \to 0}\lim_{N \to \infty}\frac{1}{N}\log\left(p_{ij}\right) = -\rho_{ij} \quad (10)$$

где: q_{ij} - вероятность ошибки первого рода, p_{ij} - вероятность ошибки второго рода *i*-го и *j*-го сигнального созвездия соответственно. Справедливо и симметричное свойство:

$$\lim_{p_{ij} \to 0}\lim_{N \to \infty}\frac{1}{N}\log\left(q_{ij}\right) = -\rho_{ij} \quad (11)$$

Таблица 1
Коэффициенты матрицы расстояний Кульбака-Лейблера для SNR = 5Дб.

	PSK2	PSK4	PSK8	PSK16	PAM4	PAM8	PAM16	QAM16	QAM64
PSK2	0	0.547	0.567	0.639	0.152	0.315	0.415	0.771	0.714
PSK4	0.944	0	0.021	0.043	0.789	0.464	0.768	0.174	0.117
PSK8	0.881	0.057	0	2.6E-5	0.712	0.696	0.708	0.062	0.029
PSK16	0.881	0.057	3.4E-5	0	0.712	0.696	0.708	0.062	0.029
PAM4	0.162	0.439	0.456	0.456	0	0.029	0.063	0.618	0.428
PAM8	0.326	0.761	0.471	0.471	0.028	0	6.4E-3	0.647	0.399
PAM16	0.421	0.495	0.497	0.497	0.061	6.5E-3	0	0.679	0.406
QAM16	1.113	0.223	0.074	0.074	0.951	0.98	1.015	0	0.056
QAM64	0.928	0.12	0.029	0.029	0.63	0.602	0.608	0.057	0

Это означает, что при проверке двух альтернатив, асимптотически вероятность ошибки классификации тесно связана с расстоянием Кульбака-Лейблера. Качественно, используя лемму Стейна можно оценить число отсчетов сигнала для достижения требуемой вероятности ошибки классификации $N \approx \frac{1}{\rho}\log\left(\frac{1}{p_{ош}}\right)$. В качестве примера, анализируя данные таблицы 1, заметим, что для классификации с вероятностью ошибки меньшей 10^{-3} для различения PAM16 и PAM8 требуется порядка 1600 отсчетов сигнала, для различения QAM16 и PSK2 достаточно 10 отсчетов.

Литература

1. Прокис Дж. Цифровая связь. Пер с англ. / под ред. Д.Д. Кловского. – М. Радио и связь. 2000. – 800с.
2. H. Chernoff. Large-sample theory: Parametric case. Ann. Math. Stat., 27:1-22,1956

———♦———

AUTOMATIC CLASSIFICATION OF MODULATION CONSTELLATIONS BASED ON A MINIMUM OF KULLBACK-LEIBLER DISTANCE

Goriachkin O.

Volga State Academy of Telecommunication and Informatic (PGATI),
23 L. Tolstoy str., Samara 443010, Russia,
Tel.: (8462) 391143, E-mail: gor@mail.radiant.ru

The task of development of algorithms of the automatic classification of modulation types is actual problem originating at development of universal demodulators and radiocontrol systems, in the present time includes significant number of the approaches to the solution.

The basic discriminatory analysis of types of digital modulation can be divided into two basic groups distinguishing in a mode of a construction of indications space of the qualifier: classification of types of digital modulation on signal constellations; a classification of types of digital modulation on signal realizations.

In the paper we consider algorithm of a classification of modulation types on signal constellations based on minimization of Kullback-Leibler distance, which is equivalent to algorithm of a maximum likelihood. The algorithm allows classifying various types of digital linear modulation without memory. It are PSK, QAM, PAM.

In the paper values of Kullback-Leibler distance matrix are analyzed for various types of digital modulation using Stein’s lemma.