
**ТЕОРИЯ И МЕТОДЫ
ОБРАБОТКИ СИГНАЛОВ**

УДК 621.396

**БАЙЕСОВСКОЕ ОЦЕНИВАНИЕ С ПОСЛЕДОВАТЕЛЬНОМ ОТКАЗОМ
И УЧЕТОМ АПРИОРНЫХ ЗНАНИЙ**

© 2020 г. М. Г. Бакулин^{а, *}, В. Б. Крейнделин^а, В. А. Григорьев^б,
В. О. Аксенов^{б, **}, А. С. Щесняк^б

^аМосковский технический университет связи и информатики,
ул. Авиамоторная, 8а, Москва 111024, Российская Федерация

^бУниверситет ИТМО, Кронверкский просп., 49, Санкт-Петербург, 197101 Российская Федерация

*E-mail: m.g.bakulin@gmail.com

**E-mail: voaksenov@ifmo.ru

Поступила в редакцию 09.08.2018 г.

После доработки 18.03.2019 г.

Принята к публикации 27.03.2019 г.

Рассмотрен метод синтеза алгоритмов оценивания, использующий байесовский подход при последовательном учете априорных знаний. Показано, что отсутствие знаний о виде априорного распределения или отказ от этих знаний из-за невозможности вывода байесовского алгоритма оценивания может быть имитирован заменой истинного априорного распределения другим распределением с большей энтропией, при этом известные алгоритмы небайесовского оценивания могут быть получены путем классического байесовского правила. Предложен подход последовательного отказа от части априорных знаний и его последующего учета для синтеза алгоритма оценивания. Приведены примеры практического использования предложенного подхода для синтеза алгоритма демодуляции сигналов в системах связи с ММО-каналом.

DOI: 10.31857/S0033849420030031

1. ПОСТАНОВКА ЗАДАЧИ

Основным разделом статистического синтеза и анализа является теория статистических решений и оценок. В связи с широким развитием техники передачи цифровой информации сфера использования теории оценивания существенно расширилась, и она стала широко применяться в области демодуляции цифровых сигналов, т.е. сигналов с конечным дискретным множеством значений.

От объема учитываемых априорных знаний зависит не только эффективность полученных оценок, но и сложность реализации алгоритма оценивания. Например, отказываясь от каких-либо априорных сведений, можно, вместо оценивания дискретного параметра использовать более простой алгоритм оценивания непрерывного параметра, а на этапе принятия окончательного решения преобразовать эти оценки, задействуя априорную информацию о дискретности параметра.

Важно не дублировать одну и ту же априорную информацию.

В статье показано, как использовать байесовский подход с последовательным отказом и последующим учетом априорных знаний об оцениваемых параметрах.

Справедливо уравнение наблюдения:

$$Y = HX(\Theta) + \eta, \quad (1)$$

где $\eta \sim \mathcal{CN}(0, R_\eta)$ – N -мерный комплексный вектор отсчетов гауссовского шума наблюдения, H – $(N \times M)$ -мерная известная матрица; Y – N -мерный комплексный вектор наблюдения; $X(\Theta) = [x_1(\theta_1) \ x_2(\theta_2) \ \dots \ x_M(\theta_M)]^T$ – M -мерный вектор комплексных символов, каждый из которых является нелинейной комплекснозначной функцией $x_i(\theta_i)$ параметра θ_i , $i = \overline{1, M}$; $\Theta = [\theta_1 \ \dots \ \theta_M]^T$ – M -мерный вектор случайных величин. Функциональные зависимости $x_i(\cdot)$ могут быть как одинаковыми для всех i , так и разными.

Задача состоит в том, чтобы по полученным наблюдениям Y оценить вектор символов Θ , причем здесь рассматривается случай, когда число оцениваемых параметров (размерность вектора Θ) равно числу комплексных символов (размерности вектора X). Это условие необязательное, и в общем случае переменная θ_i может быть вектором.

В данной постановке задача хорошо известна и имеет множество приложений: оценивание случайных процессов [1, 2], многопользовательское детектирование [3, 4], детектирование в условиях межсимвольных или межканальных помех [5], детектирование сигналов в ММО системах [6–8] и т.п.

Полная информация об оцениваемом векторе Θ содержится в апостериорном распределении, которое вычисляется по формуле Байеса [1, 2]

$$p_{ps}(\Theta|Y) = \frac{\Lambda(Y|\Theta) p_{pr}(\Theta)}{\int_{\Theta} \Lambda(Y|\Theta) p_{pr}(\Theta) d\Theta},$$

где $p_{pr}(\Theta)$ – априорное распределение вектора Θ ; $\Lambda(Y|\Theta)$ – функция правдоподобия, определяемая уравнением наблюдения, –

$$\Lambda(Y|\Theta) = \frac{1}{(\pi)^N \det(R_\eta)} \times \exp\left(- (Y - HX(\Theta))' R_\eta^{-1} (Y - HX(\Theta))\right).$$

Зная апостериорное распределение, можно вычислить точечные или интервальные оценки вектора Θ .

Процедура вычисления апостериорного распределения может быть трудоемкой из-за нелинейности уравнения наблюдения относительно оцениваемого параметра, многомерности и зависимости компонентов оцениваемого вектора.

Для решения такой задачи часто используют приближенные методы, например, “линейные” методы оценивания (детектирования), суть которых состоит в следующем.

Введем обозначение для вектора комплексных случайных величин $X = [x_1 \ x_2 \ \dots \ x_M]^T$, где $x_i \triangleq x_i(\theta_i)$. Тогда уравнение наблюдения относительно данного вектора будет линейным

$$Y = HX + \eta, \quad (2)$$

и, следовательно, для оценивания вектора X можно применять известные методы, использующие частичную априорную информацию. К последним, например, относятся метод минимума среднеквадратического отклонения (МСКО)¹ [1], предусматривающий использование информации об априорном математическом ожидании \bar{X}_{pr} и корреляционной матрице V_{pr} и описываемый следующими выражениями:

$$\begin{aligned} \hat{X}_{МСКО} &= \bar{X}_{pr} + K_{МСКО}(Y - H\bar{X}_{pr}), \\ V_{МСКО} &= (H' R_\eta^{-1} H + V_{pr}^{-1})^{-1} = \\ &= V_{pr} - K_{МСКО}(H V_{pr}), \\ K_{МСКО} &= V_{pr} H' (H V_{pr} H' + R_\eta)^{-1}, \end{aligned} \quad (3)$$

или метод наименьших квадратов (МНК) [1], также известный в теории демодуляции как декорре-

лятор (в англоязычных публикациях этот метод известен под названием “Zero-Forcing”)

$$\begin{aligned} \hat{X}_{МНК} &= (H' R_\eta^{-1} H)^{-1} H' R_\eta^{-1} Y, \\ V_{МНК} &= (H' R_\eta^{-1} H)^{-1}. \end{aligned} \quad (4)$$

Далее, как правило, используется следующий подход. Распределение ошибок оценивания предполагается гауссовским $p(X - \hat{X}) \sim \mathbb{CN}(\hat{X}_{ps}, V_{ps})$ с параметрами, определяемыми методом оценивания. После этого каждая оценка $\hat{x}_{i,ps}$ рассматривается независимо от оценок других символов как новое уравнение наблюдения [9]

$$\hat{x}_{i,ps} = x_i(\theta_i) + \mu_i, \quad i = \overline{1, M},$$

с гауссовским шумом $\mu_i \sim \mathbb{CN}(0, v_{ps}^{(ii)})$, где $v_{ps}^{(ii)}$ – i -й диагональный элемент корреляционной матрицы V_{ps} .

Затем по формуле Байеса вычисляется апостериорное распределение случайного параметра θ_i

$$p_{ps}(\theta_i | \hat{x}_{i,ps}) = \frac{\Lambda(\hat{x}_{i,ps} | \theta_i) p_{pr}(\theta_i)}{\int_{\theta_i} \Lambda(\hat{x}_{i,ps} | \theta_i) p_{pr}(\theta_i) d\theta_i}, \quad (5)$$

и его оценка

$$\hat{\theta}_i = \mathbb{F}(p_{ps}(\theta_i | \hat{x}_{i,ps})), \quad (6)$$

где $\mathbb{F}(p_{ps}(\theta_i | \hat{x}_{i,ps}))$ – некоторый нелинейный функционал от апостериорного распределения, определяемый видом выбранного критерия оптимизации (максимума апостериорной плотности или вероятности, минимума СКО и т.п.).

Традиционное определение “линейные” к алгоритмам детектирования (3), (4), (5), (6) применяется не совсем корректно, так как в итоге окончательные оценки символов $\hat{\theta}_i$ вычисляются путем нелинейного преобразования исходного наблюдения Y . Оно относится лишь к первой, промежуточной части алгоритма – оценивания символов x_i . Сама возможность использования линейных алгоритмов основана на том факте, что символы x_i являются функциями разных случайных параметров $x_i(\theta_i)$, т.е. они функционально независимы. Это позволяет после получения оценок перейти от многомерной задачи оценивания к M одномерным задачам и использовать раздельную обработку и оценивание искомым параметров θ_i . Вся процедура использования линейных алгоритмов основана на нескольких полуэмпирических подходах, не связанных общей идеей. Поэтому в некоторых случаях, как будет показано далее, это приводит к потерям в качестве оценивания.

¹ В англоязычной литературе он известен под названием Minimum Mean Square Error (MMSE).

Далее предлагается подход, соединяющий воедино все этапы оценивания и позволяющий более точно вычислить искомое апостериорное распределение и получить оценки с более высоким качеством. Предложенный метод проверен на примере демодуляции и декодирования в системах ММО.

2. БАЙЕСОВСКИЙ ПОДХОД ВЫЧИСЛЕНИЯ АПОСТЕРИОРНОГО РАСПРЕДЕЛЕНИЯ С НЕПОЛНЫМ УЧЕТОМ АПРИОРНОЙ ИНФОРМАЦИИ

Рассмотрим задачу оценивания комплексного вектора X при наблюдении (4). Вся информация об оцениваемом векторе содержится в апостериорном распределении

$$p_{ps}(X|Y) = \frac{\Lambda(Y|X)p_{pr}(X)}{\int_X \Lambda(Y|X)p_{pr}(X)dX},$$

при вычислении которого используется априорное распределение $p_{pr}(X)$ и функция правдоподобия

$$\Lambda(Y|X) = \frac{1}{(\pi)^N \det(R_\eta)} \times \exp\left(-(Y - HX)' R_\eta^{-1} (Y - HX)\right).$$

При негауссовом многомерном априорном распределении задача вычисления апостериорного распределения является достаточно трудоемкой и не всегда может быть решена. Поэтому существуют определенные небайесовские критерии, позволяющие вычислять оценки вектора X более просто.

Среди этих критериев можно выделить такие, которые используют не полную априорную информацию, а лишь ее часть, или вообще ее не используют. Среди них можно выделить, например, критерий минимума среднеквадратического отклонения [1], в котором в качестве априорной информации используется только априорное математическое ожидание и корреляционная матрица, или критерий наименьших квадратов, который вообще не использует никакой априорной информации.

Рассмотрим подход вычисления оптимальных оценок, основанный на вычислении и использовании апостериорного распределения при неполном учете априорной информации.

Пусть в исходном случае имеется полная априорная информация, которая заключена в априорном распределении $p_{pr}(X)$. Объем этой априорной информации характеризуется энтропией источника сообщений

$$H_{pr} = -\int_X p_{pr}(X) \ln(p_{pr}(X)) dX.$$

Отказ от какой-либо части априорной информации приводит к увеличению энтропии источника

сообщения. Следовательно, процедура отказа может быть имитирована заменой исходного априорного распределения $p_{ps}(X)$ другим распределением $\tilde{p}_{ps}(X)$, характеризующим источник сообщения с большей энтропией. Существуют различные варианты перехода от распределений с меньшей энтропией к распределению с более высокой энтропией.

Известно [10, 11], что среди распределений с фиксированной средней мощностью наибольшей энтропией обладает гауссовское распределение, причем наилучшей аппроксимацией негауссовской плотности гауссовской по критерию минимума расстояния Кульбака является распределение, у которого математическое ожидание и дисперсия равны истинным значениям [12]. То же самое имеет место и для распределения векторной случайной величины, т.е. максимум энтропии

$$H_{\max} = \max_{\tilde{p}_{pr}(X)} \left(-\int_X \tilde{p}_{pr}(X) \ln(\tilde{p}_{pr}(X)) dX \right),$$

при ограничениях на функцию плотности вероятности

$$\begin{aligned} \int_X \tilde{p}_{pr}(X) dX &= 1, \\ \int_X X \tilde{p}_{pr}(X) dX &= \bar{X}_{pr}, \\ \int_X (X - \bar{X}_{pr})(X - \bar{X}_{pr})' \tilde{p}_{pr}(X) dX &= V_{pr}, \end{aligned}$$

будет иметь место при следующем априорном распределении:

$$\tilde{p}_{pr}(X) \sim \mathbb{CN}(\bar{X}_{pr}, V_{pr}).$$

Следующий путь увеличения энтропии состоит в отказе от зависимости априорного распределения [12], т.е. в данном случае имеется в виду использование вместо истинной корреляционной матрицы V_{pr} диагональной априорной корреляционной матрицы

$$D_{pr} = \text{diag}(d_{pr}^{(i)}, i = \overline{1, M}),$$

где $d_{pr}^{(i)} = v_{pr}^{(ii)}$, $v_{pr}^{(ii)}$ – диагональные элементы корреляционной матрицы V_{pr} .

И, наконец, последним шагом в увеличении энтропии априорной информации является увеличение априорных дисперсий вплоть до бесконечности, т.е. использование следующего априорного распределения²

$$\bar{p}_{ps}(X) \sim \mathbb{CN}(O_M, dI_M), \text{ при } d \rightarrow \infty.$$

² Иногда для полного исключения априорной информации применяют распределение, равномерное в бесконечных пределах. Результат при этом получается одинаковым, но использование гауссовского распределения позволяет делать это в рамках одного общего подхода, основанного на постепенном увеличении энтропии сообщения.

Таким образом, предлагаемый подход постепенного отказа от априорной информации, сопровождающегося использованием априорных распределений с большим значением энтропии, позволяет создавать алгоритмы оценивания разной степени приближения и разной сложности, оставаясь при этом в рамках единого байесовского подхода. В табл. 1 приведены последовательные шаги исключения априорной информации и получаемые при этом алгоритмы вычисления апостериорного распределения (его параметров), а также связь между этими параметрами и известными оценками.

Сравнивая выражения для линейных оценок МСКО (3) и МНК (4), с выражениями для параметров апостериорных распределений из второй и четвертой строк этой таблицы, соответственно, можно заметить, что алгоритм вычисления линейной оценки МСКО является результатом синтеза с использованием байесовского подхода без учета знаний о негауссовости априорного распределения, а линейный алгоритм МНК также может быть получен с использованием байесовского подхода, при полном отсутствии априорной информации, что имитируется гауссовским априорным распределением с бесконечной дисперсией (бесконечной энтропией).

3. ПОСЛЕДОВАТЕЛЬНЫЙ УЧЕТ АПРИОРНОЙ ИНФОРМАЦИИ ПРИ БАЙЕСОВСКОМ ОЦЕНИВАНИИ

Отказ от информации о негауссовости априорного распределения для модели (2), позволяет получить простые алгоритмы оценивания, однако приводит к потерям. Но поскольку эта информация известна, то возникает вопрос, нельзя ли на следующем этапе использовать, хотя бы частично, информацию о негауссовости априорного распределения без существенного усложнения алгоритма.

Рассмотрим следующий подход. Отказ от негауссовости привел к тому, что вычисленное апостериорное распределение для модели (2) является гауссовским $\tilde{p}_{ps}(X|Y) \sim \mathbb{CN}(\bar{X}_{ps}, V_{ps})$ для всех вариантов отказа от априорных сведений (строки 2–4 табл. 1). После вычисления параметров апостериорного распределения откажемся от апостериорной зависимости между элементами вектора X , тогда можно записать

$$\tilde{p}_{ps}(X|Y) = \prod_{i=1}^M \tilde{p}_{ps}(x_i|Y), \quad (7)$$

где $\tilde{p}_{ps}(x_i|Y) \sim \mathbb{CN}(\bar{x}_{i,ps}, d_{ps}^{(i)})$, $d_{ps}^{(i)} \triangleq v_{ps}^{(ii)}$ – i -й диагональный элемент корреляционной матрицы V_{ps} .

В апостериорном распределении $\tilde{p}_{ps}(x_i|Y)$ со-держится часть априорной информации, исполь-

зуемой в виде гауссовского распределения $\tilde{p}_{pr}(x_i)$. Представим $\tilde{p}_{ps}(x_i)$ как результат, полученный вычислением по формуле Байеса

$$\begin{aligned} \tilde{p}_{ps}(x_i) &= \frac{\Lambda_{eq}(z_i|x_i) \tilde{p}_{pr}(x_i)}{\int_{x_i} \Lambda_{eq}(z_i|x_i) \tilde{p}_{pr}(x_i) dx_i} = \\ &= C \Lambda_{eq}(z_i|x_i) \tilde{p}_{pr}(x_i), \end{aligned}$$

где $\Lambda_{eq}(z_i|x_i)$ – некоторая эквивалентная функция правдоподобия с эквивалентным наблюдением z_i , которая может быть найдена следующим образом:

$$\Lambda_{eq}(z_i|x_i) = \frac{\tilde{p}_{ps}(x_i)}{\tilde{p}_{pr}(x_i)}.$$

Так как плотности $\tilde{p}_{pr}(x_i)$ и $\tilde{p}_{ps}(x_i)$ являются нормальными, то и эквивалентная функция правдоподобия также будет описываться гауссовской плотностью. Опуская промежуточные преобразования, получим

$$\Lambda_{eq}(z_i|x_i) = C \exp\left(-\frac{1}{\sigma_v^2} |z_i - x_i|^2\right), \quad (8)$$

где

$$\begin{aligned} z_i &= -\frac{d_{ps}^{(i)}}{d_{pr}^{(i)} - d_{ps}^{(i)}} \bar{x}_{i,pr} + \frac{d_{pr}^{(i)}}{d_{pr}^{(i)} - d_{ps}^{(i)}} \hat{x}_{i,ps} = \\ &= \bar{x}_{i,pr} + k_i (\hat{x}_{i,ps} - \bar{x}_{i,pr}), \quad (9) \\ \sigma_{v,i}^2 &= \frac{d_{pr}^{(i)} d_{ps}^{(i)}}{d_{pr}^{(i)} - d_{ps}^{(i)}} = k_i d_{ps}^{(i)}, \quad k_i = \frac{d_{pr}^{(i)}}{d_{pr}^{(i)} - d_{ps}^{(i)}}. \end{aligned}$$

Функция правдоподобия $\Lambda_{eq}(z_i|x_i)$ может рассматриваться как результат использования в качестве модели следующего эквивалентного уравнения наблюдения:

$$z_i = x_i + v_i,$$

где $v_i \sim \mathbb{CN}(0, \sigma_{v,i}^2)$ – отсчет эквивалентного комплексного гауссовского шума.

Применяя эквивалентную функцию правдоподобия, можно найти уточненное апостериорное распределение, использующее истинное априорное распределение для каждого элемента вектора X

$$p_{ps}(x_i|z_i) = C \Lambda_{eq}(z_i|x_i) p_{pr}(x_i),$$

и из него найти необходимые оценки, например, оценку, оптимальную по критерию МСКО

$$\hat{x}_i = \int x_i p_{ps}(x_i|z_i) dx_i.$$

Полученные выражения можно использовать для синтеза нелинейного алгоритма оценивания

Таблица 1. Сопоставление априорной информации с априорным и апостериорным распределениями, с оценками

Априорная информация	Априорное распределение	Апостериорное распределение	Оценки
Полная априорная информация	$p_{pr}(X)$	$p_{ps}(X Y) = C \Lambda(Y X) p_{pr}(X)$	$\hat{X}_{МАП} = \arg \max_X (p_{ps}(X Y))$ $\hat{X}_{МСКО} = \int \hat{X} p_{ps}(X Y) dX$
Отказ от информации о негауссовости априорного распределения	$\tilde{p}_{pr}(X) \sim \text{CN}(\bar{X}_{pr}, V_{pr})$	$\tilde{p}_{ps}(X Y) \sim \text{CN}(\bar{X}_{ps}, V_{ps})$ $\bar{X}_{ps} = \bar{X}_{pr} + K(Y - H\bar{X}_{pr})$ $V_{ps} = (H' R_{\eta}^{-1} H + V_{pr}^{-1})^{-1} = V_{pr} - K H V_{pr}$ <p style="text-align: center;">где</p> $K = (H' R_{\eta}^{-1} H + V_{pr}^{-1})^{-1} H' R_{\eta}^{-1} =$ $= V_{pr} H' (H V_{pr} H' + R_{\eta})^{-1}$	$\tilde{X}_{МАП,2} = \tilde{X}_{МСКО,2} = \bar{X}_{ps}$
Отказ от информации о зависимости (коррелированности) априорного распределения	$\tilde{p}_{pr}(X) = \prod_{i=1}^M \tilde{p}_{pr}(x_i^{(i)})$ $\tilde{p}_{pr}(x_i) \sim \text{CN}(\bar{x}_{pr}^{(i)}, d_{pr}^{(i)})$	$\tilde{p}_{ps}(X Y) \sim \text{CN}(\bar{X}_{ps}, V_{ps})$ $\bar{X}_{ps} = \bar{X}_{pr} + K(Y - H\bar{X}_{pr})$ $V_{ps} = (H' R_{\eta}^{-1} H + D_{pr}^{-1})^{-1} = D_{pr} - K H D_{pr}$ <p style="text-align: center;">где</p> $K = (H' R_{\eta}^{-1} H + D_{pr}^{-1})^{-1} H' R_{\eta}^{-1} =$ $= D_{pr} H' (H D_{pr} H' + R_{\eta})^{-1}$ $D_{pr} = \text{diag}(d_{pr}^{(i)}, i = \overline{1, M})$	$\tilde{X}_{МАП,3} = \tilde{X}_{МСКО,3} = \bar{X}_{ps}$
Полный отказ от априорной информации	$\tilde{p}_{pr}(X) \sim \text{CN}(O_M, dI_M),$ $d \rightarrow \infty$	$\tilde{p}_{ps}(X Y) \sim \text{CN}(\bar{X}_{ps}, V_{ps})$ $\bar{X}_{ps} = (H' R_{\eta}^{-1} H)^{-1} H' R_{\eta}^{-1} Y$ $V_{ps} = (H' R_{\eta}^{-1} H)^{-1}$	$\tilde{X}_{МАП,4} = \tilde{X}_{МСКО,4} = \bar{X}_{ps}$

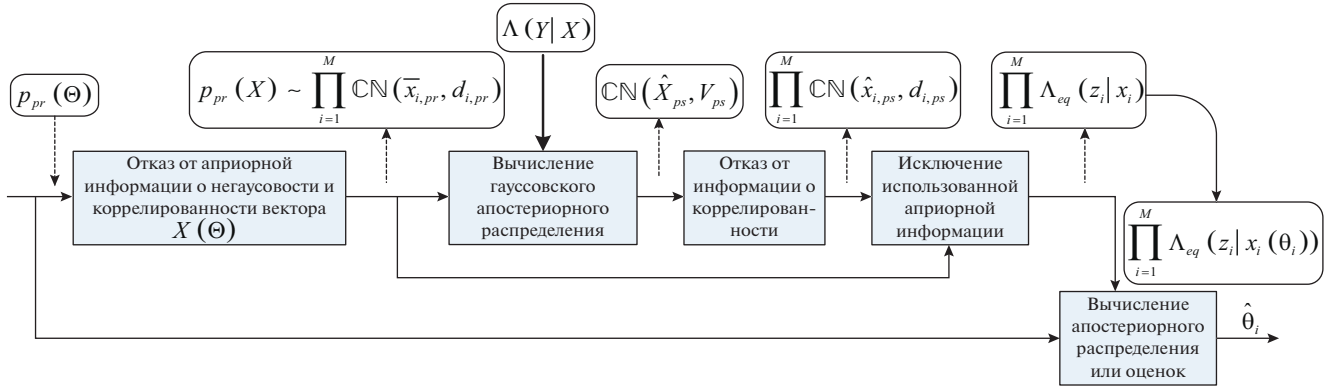


Рис. 1. Блок-схема вычисления апостериорного распределения с последовательным учетом априорной информации.

вектора Θ в модели (1). Пусть априорное распределение вектора Θ является независимым, т.е. представимо в виде

$$p_{pr}(\Theta) = \prod_{i=1}^M p_{pr}(\theta_i).$$

Это может быть следствием отказа от априорной информации о зависимости случайных величин и следствием отсутствия этой информации.

Следующим шагом будет переход от априорного распределения $p_{pr}(\Theta)$ к априорному распределению вектора $X \triangleq X(\Theta)$

$$\tilde{p}_{pr}(X) \sim \mathcal{CN}(\bar{X}_{pr}, V_{pr}) = \prod_{i=1}^M \mathcal{CN}(\bar{x}_{i,pr}, d_{i,pr}),$$

где параметры этого распределения определяют исходным априорным распределением

$$\bar{x}_{i,pr} = E\{x_i(\theta_i)\}, \quad d_{i,pr}^{(i)} = E\{\|x_i(\theta_i) - \bar{x}_{i,pr}\|^2\}.$$

Очевидно, что истинное априорное распределение $p_{pr}(X)$ не является гауссовским, так как, во-первых, функция $x_i(\theta_i)$ нелинейная с ограниченной областью значений, а, во-вторых, сам параметр θ_i , как правило, имеет негауссовское априорное распределение, однако на данном шаге мы не будем использовать информацию о негауссовости априорного распределения случайной величины $x_i \triangleq x_i(\theta_i)$.

В соответствии со строкой 3 табл. 1 апостериорное распределение при данном объеме учитываемой априорной информации будет гауссовским $p_{ps}(X) \sim \mathcal{CN}(\hat{X}_{ps}, V_{ps})$ с параметрами

$$\begin{aligned} \hat{X}_{ps} &= \bar{X}_{pr} + K(Y - H\bar{X}_{pr}), \\ V_{ps} &= V_{pr} - K(HD_{pr}), \end{aligned}$$

где $K = D_{pr}H'(HD_{pr}H' + R_{\eta})^{-1}$.

Следующим шагом является переход от апостериорного распределения коррелированных случайных величин к независимому распределению (7) и вычисление параметров эквивалентной функции правдоподобия $\Lambda_{eq}(z_i|x_i)$ в соответствии с (9).

Учитывая, что $x_i \triangleq x_i(\theta_i)$, можем записать апостериорное распределение для оцениваемой случайной величины θ_i

$$p_{ps}(\theta_i|z_i) = C\Lambda_{eq}(z_i|x_i(\theta_i))p_{pr}(\theta_i).$$

Далее апостериорное распределение может использоваться по назначению в зависимости от выбранного критерия.

На рис. 1 приведена блок-схема описанного подхода вычисления апостериорного распределения (или оценок) вектора параметров для модели (1). Отметим, что традиционное использование линейных алгоритмов оценивания для этой модели, например, алгоритма МСКО, отличается отсутствием блока исключения использованной априорной информации, что в итоге может приводить к ухудшению качества вычисленных оценок. Использованная дополнительная процедура вычисления параметров эквивалентной функции правдоподобия незначительно увеличивает сложность. В общем случае добавляется M операций действительного деления, $3M$ операции действительного умножения и $5M$ операций сложения.

4. ДЕМОДУЛЯЦИЯ В СИСТЕМАХ ММО С ПОСЛЕДОВАТЕЛЬНЫМ ОТКАЗОМ И УЧЕТОМ АПРИОРНОЙ ИНФОРМАЦИИ

Пусть дана модель уравнения наблюдения для системы связи с ММО-каналом

$$Y = HS(\Theta) + \eta, \tag{10}$$

где $\eta \sim N(0, R_{\eta})$ – комплексный гауссовский шум, $\Theta = [\theta_1 \dots \theta_m]^T$ – m -мерный вектор двоич-

ных символов (битов) $\theta_i \in \{-1; 1\}$, $i = \overline{1, m}$, которые нужно оценить; Y – N -мерный комплексный вектор наблюдения; $S(\Theta)$ – M -мерный вектор комплексных квадратурно-амплитудно модулированных (КАМ) символов. Для классической системы с пространственным мультиплексированием имеем

$$S(\Theta) = S(\theta_1, \dots, \theta_m) = \left[s_1 \left(\vartheta_1^{(1)}, \dots, \vartheta_k^{(1)} \right) \dots s_M \left(\vartheta_1^{(M)}, \dots, \vartheta_k^{(M)} \right) \right]^T,$$

где $\vartheta_k^{(l)}$ – k -й из совокупности K двоичных символов, входящих в состав l -го КАМ-символа, число точек созвездия которого равно 2^K . Между элементами вектора Θ и битами КАМ-символов существует следующее соответствие:

$$\theta_i = \vartheta_k^{(l)}, \text{ при } i = k + (l - 1)K, \\ k = \overline{1, K}, \quad l = \overline{1, M},$$

при этом $m = KM$. Общее число комбинаций вектора КАМ-символов $S(\Theta)$ равно 2^{KM} .

Случайные величины θ_i являются дискретными и в данном случае могут принимать всего два значения, 1 и -1 . Поэтому для них оценки могут быть двух видов – “жесткие” и “мягкие”. Оценки, принимающие значения согласно области значений оцениваемой случайной величины, называются “жесткими” оценками. Они, как правило, соответствуют критериям максимума апостериорной вероятности или максимального правдоподобия

$$\hat{\theta}_{i, \text{МАВ}} = \arg \max_{\theta \in \{-1; 1\}} (p_{ps}(\theta_i | Y)), \\ \hat{\theta}_{i, \text{МП}} = \arg \max_{\theta \in \{-1; 1\}} (\Lambda(Y | \theta_i)).$$

“Мягкие” оценки или “мягкие” решения могут принимать значения, отличающие от области значений оцениваемой случайной величины. Для одномерной двоичной случайной величины $\theta = \pm 1$ распределение характеризуется вероятностями $P\{\theta = 1\} = p_0$ и $P\{\theta = -1\} = p_1 = 1 - p_0$. Аналитически это распределение можно представить двумя способами [6, 13]:

$$P(\theta) = \frac{1}{2}(1 + \theta\mu) = \frac{\exp(\theta\lambda)}{\exp(\lambda) + \exp(-\lambda)},$$

где $\mu = E\{\theta\} = (p_0 - p_1)$ – математическое ожидание; $\lambda = \frac{1}{2} \ln \frac{p_0}{p_1}$.

Между параметрами μ и λ существует однозначная связь

$$\mu = \text{th}(\lambda).$$

Таким образом, в параметре μ или в связанном с ним параметре $\lambda = \text{arcth}(\mu)$ содержится полная информация о распределении двоичной случайной величины. Нетрудно показать, что параметр μ является оптимальной оценкой бита по критерию минимума среднеквадратического отклонения, т.е.

$$\tilde{\theta}_{\text{МСКО}} = \arg \min_{\theta_{\text{МСКО}}} \left(E(\theta - \tilde{\theta}_{\text{МСКО}})^2 \right) = \\ = \arg \min_{\theta_{\text{МСКО}}} \left(\sum_{\theta \in \{-1; 1\}} (\theta - \tilde{\theta}_{\text{МСКО}})^2 P(\theta) \right) = \mu.$$

Оба параметра μ и λ могут использоваться в качестве “мягкой” оценки бита. Однако, как показывает практика, с точки зрения точности вычислений предпочтительней оказываются логарифмические параметры, т.е. параметр λ . Поэтому в дальнейшем в качестве “мягкой” оценки бита будем использовать именно его, т.е.

$$\tilde{\theta} \triangleq \lambda = \frac{1}{2} \ln \frac{p_{ps}(\theta = 1 | Y)}{p_{ps}(\theta = -1 | Y)}. \quad (11)$$

Следует отметить, что из “мягких” оценок можно легко вычислить “жесткую” оценку

$$\hat{\theta}_{\text{МАВ}} = \text{sign}(\mu) = \text{sign}(\lambda).$$

Рассмотрим вычисление “мягких” оценок для демодулятора сигналов в ММО-канале. Для модели наблюдения (10) с учетом равной вероятности комбинаций и представления сигналов в виде функции битов $S(\Theta)$ апостериорное распределение будет определяться выражением

$$p_{ps}(\Theta | Y) = C p_{pr}(\Theta) \times \exp\left(- (Y - HS(\Theta))' R_n^{-1} (Y - HS(\Theta))\right).$$

Для вычисления апостериорного распределения i -го бита необходимо усреднить многомерное апостериорное распределение $p_{ps}(\Theta | Y)$ по всем битам кроме i -го. Для дискретного распределения операция усреднения выполняется путем суммирования

$$p_{ps}(\theta_i | \tilde{y}) = \sum_{\substack{\theta_k \in \{-1; 1\}, k=1, m \\ k \neq i}} p_{ps}(\Theta | Y).$$

С учетом этого и формулы (10) “мягкая” оценка i -го бита будет определяться выражением

$$\tilde{\theta}_i \triangleq \lambda_i = \frac{1}{2} \ln \frac{\sum_{\substack{\theta_k \in \{-1; 1\}, k=1, m \\ k \neq i}} p_{ps}(\theta_1, \dots, \theta_m | Y)|_{\theta_i=1}}{\sum_{\substack{\theta_k \in \{-1; 1\}, k=1, m \\ k \neq i}} p_{ps}(\theta_1, \dots, \theta_m | Y)|_{\theta_i=-1}}.$$

Таким образом, оптимальный алгоритм вычисления “мягких” оценок битов для ММО-канала в общем случае описывается следующим выражением:

$$\tilde{\theta}_i \triangleq \lambda_i = \frac{1}{2} \ln \frac{\sum_{\substack{\theta_k \in \{-1;1\}, k=1,m \\ k \neq i}} p_{pr}(\theta_1, \dots, \theta_m) \exp\left(-(Y - HS(\theta_1, \dots, \theta_m))' R_{\eta}^{-1} (Y - HS(\theta_1, \dots, \theta_m))\right) \Big|_{\theta_i=1}}{\sum_{\substack{\theta_k \in \{-1;1\}, k=1,m \\ k \neq i}} p_{pr}(\theta_1, \dots, \theta_m) \exp\left(-(Y - HS(\theta_1, \dots, \theta_m))' R_{\eta}^{-1} (Y - HS(\theta_1, \dots, \theta_m))\right) \Big|_{\theta_i=-1}}. \quad (12)$$

Алгоритм вычисления мягких оценок (12) имеет очень высокую сложность $\sim KM2^{KM}$ даже при независимых и равновероятных комбинациях битов. Данное выражение допускает некоторое упрощение, но порядок сложности все равно остается $\sim 2^{KM}$. Поэтому на практике используют различные приближенные методы, которые позволяют уменьшить число перебираемых комбинаций. Среди наиболее простых методов можно выделить так называемые “линейные” алгоритмы оценивания.

Суть этих методов состоит в том, что сначала путем линейной обработки находятся оценки КАМ-символов, а затем эти оценки используются в качестве новых независимых наблюдений для мягкой демодуляции битов. При этом сложность демодуляции одного символа будет пропорциональна 2^K .

Используем байесовский подход, основанный на последовательном учете априорной информации, описанный в предыдущих разделах.

Сначала вычислим параметры априорного распределения вектора $S = [s_1, \dots, s_M] \triangleq S(\theta_1, \dots, \theta_m)$. Для независимых и равновероятных битов имеем

$$\begin{aligned} \bar{s}_{i,pr} &= E\left\{s_i \left(\vartheta_1^{(i)}, \dots, \vartheta_K^{(i)}\right)\right\} = 0, \\ d_{pr}^{(i)} &= E\left\{\left|s_i \left(\vartheta_1^{(i)}, \dots, \vartheta_K^{(i)}\right)\right|^2\right\} = 1. \end{aligned}$$

Очевидно, что априорные распределения случайных величин s_i будут негауссовскими, но согласно описанному ранее подходу на данном шаге мы отказываемся от этой информации и используем распределение, наихудшее с точки зрения максимума энтропии сообщения, т.е. гауссовское априорное распределение $p_{pr}(S) = \mathbb{CN}(0, I_M)$, где I_M – единичная матрица размером $(M \times M)$.

С учетом этого (согласно третьей строке табл. 1) параметры апостериорного гауссовского распределения $p_{ps}(S) = \mathbb{CN}(\hat{S}_{ps}, V_{ps})$ будут

$$\begin{aligned} \hat{S}_{ps} &= (H' R_{\eta}^{-1} H + I)^{-1} H' R_{\eta}^{-1} Y, \\ V_{ps} &= (H' R_{\eta}^{-1} H + I)^{-1}. \end{aligned}$$

Как уже отмечалось, эти оценки полностью совпадают с линейными оценками МСКО.

Для дальнейшего синтеза алгоритма демодуляции отказываемся от апостериорной зависимости компонентов вектора S и используем апостериорное распределение

$$p_{ps}(S) = \prod_{i=1}^M \mathbb{CN}(\hat{s}_{i,ps}, d_{ps}^{(i)}).$$

Для исключения априорной информации, которую мы использовали при вычислении априорных математических ожиданий и дисперсий перейдем к эквивалентным функциям правдоподобия $\Lambda_{eq}(z_i | x_i)$ (8) с параметрами

$$z_i = \frac{\hat{s}_{i,ps}}{1 - d_{ps}^{(i)}}, \quad \sigma_{v,i}^2 = \frac{d_{ps}^{(i)}}{1 - d_{ps}^{(i)}}, \quad (13)$$

где при выводе этих выражений учитывалось, что $\bar{s}_{i,pr} = 0$, $d_{pr}^{(i)} = 1$.

Следует отметить, что отличие полученного алгоритма от известного алгоритма демодуляции с использованием линейных оценок МСКО в данном случае заключается только в операциях вычисления параметров эквивалентной функции правдоподобия (13), что усложняет алгоритм всего на M операций действительного деления, M операций действительного сложения и на $3M$ операций действительного умножения.

Учитывая независимость и равновероятность комбинаций битов, получим

$$\tilde{\vartheta}_k^{(i)} \triangleq \lambda_k^{(i)} = \frac{1}{2} \ln \frac{\sum_{\substack{\vartheta_j^{(i)} \in \{-1;1\}, j=1,K \\ j \neq k}} \exp\left(-\frac{1}{2\sigma_{v,i}^2} \left|z_i - s_i(\vartheta_1^{(i)}, \dots, \vartheta_K^{(i)})\right|^2\right) \Big|_{\vartheta_k^{(i)}=1}}{\sum_{\substack{\vartheta_j^{(i)} \in \{-1;1\}, j=1,K \\ j \neq k}} \exp\left(-\frac{1}{2\sigma_{v,i}^2} \left|z_i - s_i(\vartheta_1^{(i)}, \dots, \vartheta_K^{(i)})\right|^2\right) \Big|_{\vartheta_k^{(i)}=-1}}, \quad i = \overline{1, M}, \quad k = \overline{1, K}. \quad (14)$$

Нетрудно видеть, что сложность вычисления “мягких” оценок битов пропорциональна $MK2^K$ плюс сложность МСКО-алгоритма ($\sim M^3$) и плюс сложность вычисления параметров эквивалентных функций правдоподобия ($\sim M$). Если учесть, что в большинстве систем связи используются квадратные КАМ-созвездия с кодированием Грея, которое допускает отдельную обработку квадратурных составляющих, то сложность вычисления по формуле (14) будет составлять $2MK2^{K/2}$.

Недостатком (14) является большой динамический диапазон суммируемых чисел в числителе и знаменателе дроби, что ухудшает точность вычисления при ограниченной разрядности вычислений. Чтобы избежать этого, можно использовать следующий прием. Введем обозначение для метрик аргументов экспоненциальной функции

$$\Delta_{k,l_k^{(+)}}^{(i)} = -\frac{1}{\sigma_{v,i}^2} \left| z_i - s_i(\vartheta_1^{(i)}, \dots, \vartheta_K^{(i)}) \right|^2 \Bigg|_{\substack{\vartheta_j^{(i)} \in \{-1; 1\}, j=1, \overline{K}, j \neq k \\ \vartheta_k^{(i)} = 1}},$$

$$\Delta_{k,l_k^{(-)}}^{(i)} = -\frac{1}{\sigma_{v,i}^2} \left| z_i - s_i(\vartheta_1^{(i)}, \dots, \vartheta_K^{(i)}) \right|^2 \Bigg|_{\substack{\vartheta_j^{(i)} \in \{-1; 1\}, j=1, \overline{K}, j \neq k \\ \vartheta_k^{(i)} = -1}},$$

где $l_k^{(+)} = \overline{1, 2^{K/2}}$ – номера комбинаций двоичных битов $\vartheta_1^{(i)}, \dots, \vartheta_K^{(i)}$ для всех возможных значений $\vartheta_j^{(i)} \in \{-1; 1\}$, и фиксированном значении k -го бита $\vartheta_k^{(i)} = 1$, аналогично $l_k^{(-)} = \overline{1, 2^{K/2}}$ – номера комбинаций при $\vartheta_k^{(i)} = -1$. С учетом этого выражение (14) можно преобразовать к виду

$$\tilde{\vartheta}_k^{(i)} \triangleq \lambda_k^{(i)} = \frac{1}{2} \left(\max \Delta_{k,l_k^{(+)}}^{(i)} - \max \Delta_{k,l_k^{(-)}}^{(i)} \right) + \frac{1}{2} \ln \frac{\sum_{l_k^{(+)}=1}^{2^{K/2}} \exp(\Delta_{k,l_k^{(+)}}^{(i)} - \max \Delta_{k,l_k^{(+)}}^{(i)})}{\sum_{l_k^{(-)}=1}^{2^{K/2}} \exp(\Delta_{k,l_k^{(-)}}^{(i)} - \max \Delta_{k,l_k^{(-)}}^{(i)})}.$$

Преимущества этой формулы состоят в том, что динамический диапазон аргументов экспоненциальной функции, самих сумм и их отношений существенно меньше, чем в исходном выражении. Из этой формулы вытекает приближенный алгоритм вычисления “мягкой” оценки, если ограничиться первым слагаемым

$$\tilde{\vartheta}_k^{(i)} \triangleq \lambda_k^{(i)} = \frac{1}{2} \left(\max \Delta_{k,l_k^{(+)}}^{(i)} - \max \Delta_{k,l_k^{(-)}}^{(i)} \right).$$

5. МОДЕЛИРОВАНИЕ

Для проверки предложенного подхода последовательного учета априорной информации и срав-

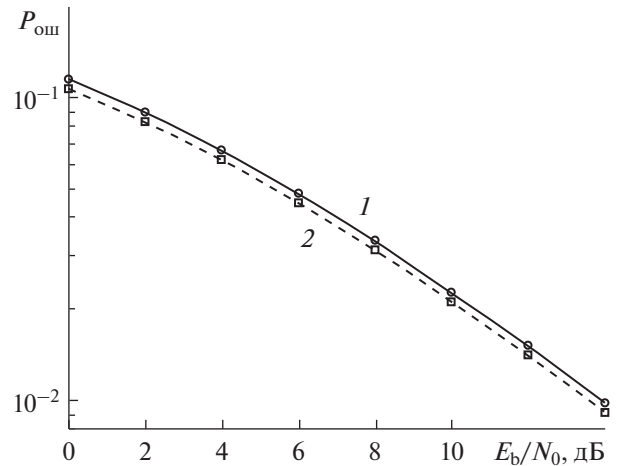


Рис. 2. Зависимости вероятности ошибки на бит от отношения сигнал/шум на бит в системе с ММО-каналом без кодирования для МСКО-детектора (1) и детектора, использующего предложенный подход с последовательным отказом и учетом априорной информации (2).

нения его с аналогичными алгоритмами было проведено моделирование системы с ММО-каналом со следующими условиями моделирования: число передающих антенн $M = 8$, число приемных антенн $N = 8$, модуляция 16 КАМ, ММО-канал с независимыми релейскими замираниями.

Были исследованы зависимости вероятности ошибки на бит для системы без кодирования и вероятности ошибки на бит и на кадр для систем с кодированием от отношения сигнал/шум на бит. В качестве канального кода использовался турбокод со скоростью кодирования 1/2 и длиной кадра 573 бита.

Детектор, синтезированный с использованием предложенного подхода (рис. 2), имеет характеристики помехоустойчивости на 0.3 дБ лучше, чем аналогичный по сложности МСКО-детектор. Это улучшение незначительное, так как в системе без кодирования используются жесткие оценки битов, которые соответствуют критерию максимума апостериорной вероятности и используют не всю информацию.

Для системы с кодированием явно видно преимущество предложенного подхода (рис. 3а и 3б). Помехоустойчивость полученного алгоритма лучше на 1...1.5 дБ. Это объясняется более корректным использованием априорной информации и, как следствие, более точным вычислением “мягких” оценок битов.

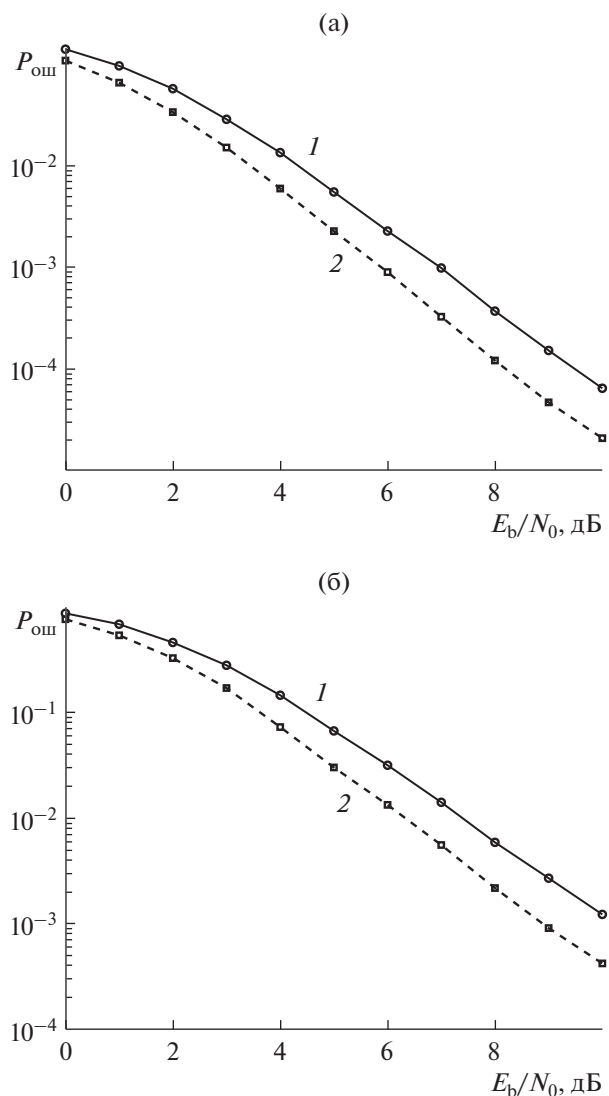


Рис. 3. Зависимости вероятности ошибки на бит (а) и на кадр (б) от отношения сигнал/шум на бит в системе с ММО-каналом с кодированием для МСКО-детектора (1) и детектора, использующего предложенный подход с последовательным отказом и учетом априорной информации (2).

ЗАКЛЮЧЕНИЕ

Таким образом, предложен подход к решению задачи нелинейного оценивания, основанный на последовательном учете априорной информации. Показано, что путем отказа от части априорной информации можно упрощать алгоритмы оценивания. Причем отказ от априорной информации может быть имитирован заменой исходного априорного распределения другим с большей энтропией, например, вместо негуассовского — гауссовское,

вместо многомерного зависимого — многомерное независимое, вместо распределения с фиксированной дисперсией — распределение с бесконечной дисперсией, и т.п.

При этом синтез алгоритмов оценивания остается в рамках единого байесовского подхода и оперирует понятиями априорного распределения, функции правдоподобия и апостериорного распределения. Это позволяет учесть априорную информацию, используемую на промежуточных шагах синтеза, и более точно использовать неучтенную информацию на последующих шагах.

Приведенный пример использования данного подхода для синтеза алгоритма оценивания двоичных битов при передаче информации в системах связи с ММО показывает преимущество при разработке демодуляторов с линейными алгоритмами оценивания символов за счет более точного учета априорной информации, вследствие чего формируются оценки битов с лучшим качеством оценивания, чем известные линейные алгоритмы.

СПИСОК ЛИТЕРАТУРЫ

1. Седж Э., Мелс Дж. Теория оценивания и ее применение в связи и управлении. М.: Связь, 1976.
2. Тихонов В.И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем. М.: Радио и связь, 1991.
3. Verd'u S. Multiuser Detection. Cambridge: Univ. Press, 1998.
4. *Advances in Multiuser Detection* / Ed. M.L. Honig. Hoboken: John Wiley & Sons, Inc., 2009.
5. Li Y., Vucetic B., Sato Y. // IEEE Trans. 1995. V. IT-41. № 3. P. 704.
6. Шлома А.М., Бакулин М.Г., Крейнделин В.Б., Шумов А.П. Новые алгоритмы формирования и обработки сигналов в системах подвижной связи / Под ред. А.М. Шломы. М.: Горячая линия-Телеком, 2008.
7. Бакулин М.Г., Варукина Л.А., Крейнделин В.Б. Технология ММО: принципы и алгоритмы. М.: Горячая линия-Телеком, 2014.
8. Bai L., Choi J. Low Complexity MIMO Detection. N.Y.: Springer, 2012.
9. Tüchler M., Singer A.C., Koetter R. // IEEE Trans. 2002. V. SP-50. № 3. P. 673.
10. Ту Дж., Гонсалес Р. Принципы распознавания образов. М.: Мир, 1978.
11. Куликовский Л.Ф., Мотов В.В. Теоретические основы информационных процессов. М.: Высш. школа, 1987.
12. Шеннон К. Работы по теории информации и кибернетике. М.: Изд-во иностр. лит., 1963.
13. Бакулин М.Г. // Электросвязь. 2000. № 1. С. 11.