

УДК: 621.371.25, 550.388.2

Об одной модели смеси распределений вероятностей в радиотехнических измерениях

А. А. Колчев^а, А. Е. Недопекин^б

Марийский государственный университет,
физико-математический факультет,
Россия, 424001, г. Йошкар-Ола, пл. Ленина, д. 1

E-mail: ^аkolchevaa@mail.ru, ^бagasfer911@yandex.ru

Получено 20 июня 2012 г.,
после доработки 24 июля 2012 г.

В данной работе представлена модель смеси распределений вероятностей сигнала и шума. Как правило, при анализе данных в условиях неопределенности приходится использовать непараметрические критерии. Однако при анализе нестационарных данных при наличии неопределенности по виду закона распределения и его параметрам они могут оказаться малоэффективными. Рассматриваемая модель подразумевает реализацию случая априорной непараметрической неопределенности при обработке сигнала в условиях, когда возможно разделение сигнала и шума как компонентов, относящихся к разным генеральным совокупностям.

Ключевые слова: статистическая неопределенность, модель, выделение сигнала

On one particular model of a mixture of the probability distributions in the radio measurements

A. A. Kolchev, A. E. Nedopekin

Mari State University, Faculty of Physics and Mathematics, MarSU, 1 Lenin sq., Yoshkar-Ola, 424001, Russia

Abstract. – This paper presents a model mixture of probability distributions of signal and noise. Typically, when analyzing the data under conditions of uncertainty it is necessary to use nonparametric tests. However, such an analysis of nonstationary data in the presence of uncertainty on the mean of the distribution and its parameters may be ineffective. The model involves the implementation of a case of a priori non-parametric uncertainty in the processing of the signal at a time when the separation of signal and noise are related to different general population, is feasible.

Keywords: statistical uncertainty, model, selection of signal

Citation: *Computer Research and Modeling*, 2012, vol. 4, no. 3, pp. 563–568 (Russian).

Введение

В радиотехнических задачах существует потребность выделения сигнала на фоне шума или помех. Классические способы выделения предусматривают использование критерия максимального правдоподобия при известных априорных предположениях о виде распределения сигнала и фона, препятствующего выделению [Ширман, 1970]. В случае отсутствия знаний о виде законов распределения, когда реализуется ситуация непараметрической неопределенности, рекомендуется использовать непараметрические методы (знаковые, ранговые, знаково-ранговые) [Левин, 1976].

В ряде случаев, когда возникает задача обнаружения и выделения сигналов, целевые отсчеты в рассматриваемой группе данных могут иметь определяющее значение для структуры (топологии) анализируемых выборок. В задачах радиофизических исследований нестационарных сред, например ионосферы, могут реализовываться ситуации априорной непараметрической неопределенности, когда не известны ни параметры, ни вид распределения смеси отсчетов, соответствующих интересуемому данным и фоновых отсчетов [Колчев, Недопекин, 2011]. При этом на первый план выходит именно выделение заранее неизвестных структур, обусловленных поведением изучаемой среды. В этом случае особое внимание следует уделять структуре рассматриваемых выборок. Использование же непараметрических обнаружителей, опирающихся на сравнение соседних по времени выборок смеси сигнал-шум, может испытывать трудности при рассмотрении нестационарных данных в условиях статистической неопределенности, когда закон распределения не просто неизвестен, но испытывает постоянные изменения.

Цель – изучение модели смеси распределений при выделении сигнала в условиях непараметрической неопределенности.

Модель смеси

В работе [Соколов, 2002] изложены вопросы, касающиеся топологии выборок в связи с фрактальными закономерностями. Однако далеко не всегда в экспериментальных данных существуют подобные закономерности. Рассмотрим ситуацию, когда анализируемые данные получены в условиях непараметрической неопределенности из исследования нестационарной среды, имеющей нефрактальный случайный характер. Одна из подобных ситуаций, при разрешении которой неэффективны непараметрические статистики, описана в [Колчев, Недопекин, 2011]. Была рассмотрена обработка широкополосного КВ линейно-частотно модулированного (ЛЧМ) сигнала ионосферного распространения после когерентного приема.

При обработке принятого сигнала ЛЧМ-ионозонда происходит его сжатие в частотной области путем умножения на сигнал, комплексно-сопряженный излученному сигналу. Затем производится выделение низкочастотной составляющей — так называемого разностного сигнала [Иванов, 2003]. При этом полезный сигнал преобразуется к квазигармоническому виду, а помехи, вызванные посторонними радиотехническими устройствами, преобразуются в импульсы малой длительности – так называемые импульсные и сосредоточенные помехи. Наличие сосредоточенных помех способно усложнить выделение ЛЧМ сигнала, особенно в автоматическом режиме, главным образом за счет уменьшения величины отношения сигнал-шум S/N .

Смесь сигнал-шум не имеет преобладающего закона распределения, более того, законы распределения подвержены постоянной смене. Эффективное выделение отсчетов сигнала возможно при соблюдении следующих условий: целевые выделяемые отсчеты составляют малую долю от общего объема выборки, целевые отсчеты отличаются по значению (или параметру) от остальных отсчетов.

Теперь задача выделения сводится к разделению вариационного ряда анализируемой выборки на части, соответствующие интересующим отсчетам и всем остальным отсчетам. Структурно такая выборка представляет собой выборку из генеральной совокупности, чья плотность распределения вероятности имеет так называемый «утяжеленный хвост», в котором сконцен-

трированы целевые отсчеты. И хотя в общем случае сумма двух распределений вероятностей представляет собой их свертку [Новицкий, Зограф, 1991], при соблюдении вышеуказанных условий сумма распределений целевых отсчетов и фона может быть представлена как сумма плотностей вероятностей:

$$f(x) = h_1 f_1(x) + h_2 f_2(x), \quad (1)$$

где $f_1(x)$ – плотность распределения шума (фона), а $f_2(x)$ – сигнала. Весовые коэффициенты (доля в общей генеральной совокупности) значительно отличаются: $h_2 \ll h_1$. При непараметрической неопределенности вид функций $f_1(x)$ и $f_2(x)$ неизвестен, поэтому будем проводить качественное исследование данной структурной модели в обобщенном случае, когда указанные функции представлены простейшими соотношениями. Пусть $x \in [0, 1]$, z – отношение сигнал-шум в линейном масштабе, которое в логарифмическом масштабе можно оценить как $S/N = 20 \cdot \lg(z)$. Функция распределения задается выражениями вида

$$f(x) = \begin{cases} (1-h_2) \cdot z, & x \in [0, 1/z], \\ z \cdot h_2 / (z-1), & x \in [1/z, 1], \end{cases} \quad \int_0^1 f(x) = 1, \quad (2)$$

то есть представляет собой два прямоугольника, первый из которых соответствует шуму, а второй – сигналу, при этом сохраняется нормировка функции плотности вероятностей (см. рис. 1, а). Величина $1/z$ на рис. 1, а) будет соответствовать длине второго прямоугольника, выполняющего роль $f_2(x)$. Для функции на рисунке $z = 6$, $h_2 = 0,15$. На рис. 1, б) показан пример типичной выборки смеси сигнал-шум в частотной области для оцифрованного сигнала, полученного с выхода системы сжатия сигнала ЛЧМ-иозонда при наклонном зондировании ионосферы; на рис. 1, в) – оценка функции плотности, соответствующая данной выборке. Частотные составляющие в спектре разностного ЛЧМ сигнала в окрестности сигнала ~ 600 и 850 Гц соответствуют различным модам принятого сигнала.

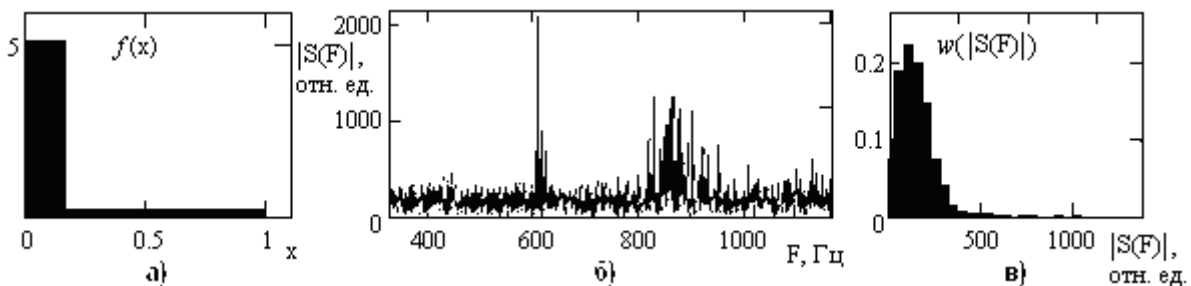


Рис. 1

Статистические параметры для данной функции были рассмотрены при изменяющихся параметрах z и h_2 (рис. 2 а) асимметрия A , б) эксцесс E , в) нормированный момент 6-го порядка μ_6/σ^6 , σ – среднеквадратическое отклонение). Также рассматривались отношения μ_n/μ_n^0 нормированных центральных моментов при наличии сигнала к значениям одноименных центральных моментов при их отсутствии. Пример такой величины для эксцесса при различных значениях S/N представлен на рис. 2, г) – проекция поверхности E на плоскость (h_2, E) . Из рисунка видно, что с ростом отношения S/N при малой доле h_2 наблюдается рост значений рассматриваемых параметров. Для асимметрии и эксцесса рост значений начинается с 4–5 дБ, для центрального момента 6-го порядка рост значений начинается с 10 дБ. Рассмотренные центральные моменты достигают максимума при h_2 , равном 0,01. Эксцесс распределения дает чувствительность, большую, чем асимметрия, но меньшую, чем μ_6/σ^6 , зато начиная с меньшей величины порога S/N .

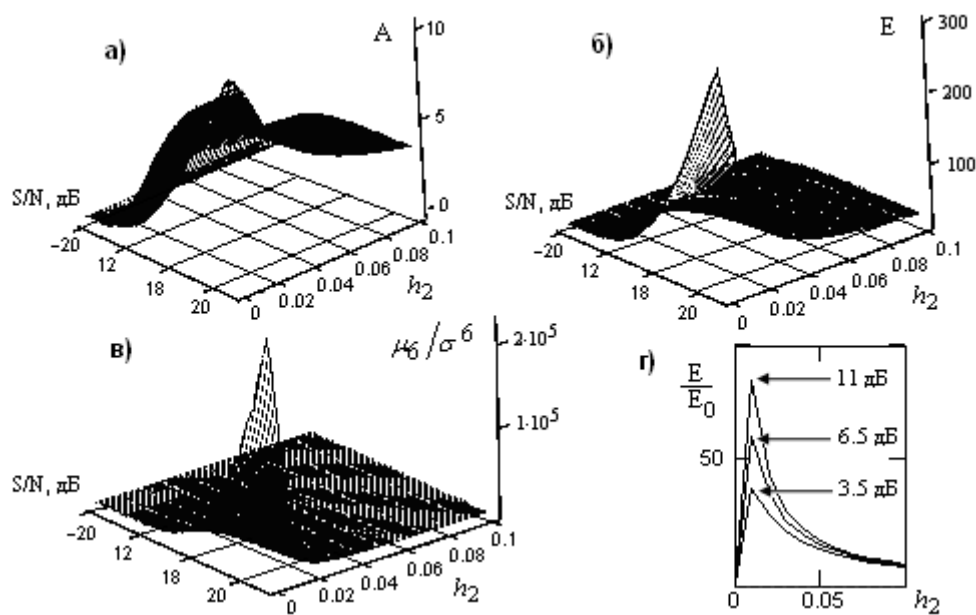


Рис. 2

Возможности использования модели

В рассмотренной смеси (1) сигнальные отсчеты, составляющие малую долю от общего объема выборки и сконцентрированные в «утяжеленном хвосте», могут рассматриваться как аномальные отсчеты на фоне отсчетов шума. При анализе даже столь простой модели видно, что наличие аномальных отсчетов может значительно менять поведение параметров суммарного распределения. Особый интерес представляет поведение параметров при малых значениях весового коэффициента h_2 от 0,01 до 0,05, что соответствует случаю спектрального анализа сигналов радиозондирования. Чем больше z , тем выше отношение сигнал–шум и тем больше изменяются рассмотренные параметры. При этом с ростом порядка центральный момент реагирует на наличие аномальных отсчетов сильнее. Дальнейшее убывание кривых после значения h_2 , равного 0,01–0,03, вызвано тем, что в общей массе отсчетов выборки сигнальные отсчеты перестают быть аномальными.

Как видно из анализа модели, при неизвестных законах распределения и сделанных выше предположениях о смеси распределений признаком наличия сигнала может служить некоторый статистический параметр. Обнаружение сигнала в рамках данной модели будет сводиться к классификации анализируемых выборок амплитудных спектров по двум классам: «шум» и «сигнал». Для возможности обнаружения сигнала в частотной выборке необходимо знать тот параметр, который в наибольшей степени реагирует на наличие аномальных отсчетов, наилучшим образом коррелирует с физическими параметрами сигнала (например, его мощностью) и при некоторых критических значениях обеспечивает наименьшие ошибки пропуска сигнала (ошибка первого рода) и ложного обнаружения (ошибка второго рода).

Выделение сигнала в рамках данной модели будет сводиться к расщеплению суммы распределений (1), а именно кластеризации анализируемой выборки на две части по некоторому решающему правилу – критерию выделения. Для этого могут быть использованы критерии тестирования статистических выбросов, применяемые для обработки результатов экспериментов [Колчев и др., 2006].

Пример использования

Для случая сигнала ЛЧМ-ионозонда в частотной области, для которого известно об априорной непараметрической неопределенности, возможны оба следствия из описанной модели: обнаружение сигнала по статистическому параметру и его отсчетов в частотной области при помощи критерия цензурирования.

На рис. 3, б) представлено поведение эксцесса распределения для выборок в частотной области, содержащих сигнал, в натурном опыте, а на рис. 3, а) – в представленной формулой (2) модели. Видна большая схожесть поведения параметров, из рисунка видно, что наиболее часто встречающееся значение h_2 для рассмотренного в эксперименте сигнала заключено в интервале 1–3 %.

Использование статистического параметра в качестве статистики обнаружения для рассмотренного экспериментального случая дало следующие результаты. Из числа статистических параметров наилучшую ранговую корреляцию ($-0,943$) с S/N показал контрэксцесс χ , получаемый из эксцесса E как $1/\sqrt{E}$. Равновероятностная ошибка обнаружения сигнала по контрэксцессу составила около 6 %, а по знаковой статистике – около 26 %.

Применение методики ослабления сосредоточенных помех, применяемой во временной области [Колчев, Щирый, 2006], позволяет достигнуть выигрыша по S/N в среднем на 1,5–2 дБ. При этом результат для равновероятностной ошибки обнаружения с использованием контрэксцесса улучшается до величины $\sim 4,7$ – $4,9$ %.

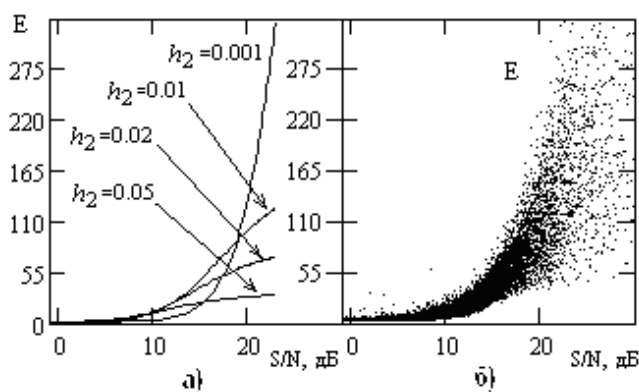


Рис. 3

Использование контрэксцесса как статистики-обнаружителя было опробовано при определении максимальной наблюдаемой частоты (МНЧ) трассы ионосферного зондирования, определяемой в частотной области как последний спектр, содержащий сигнал. Средняя относительная ошибка при сравнении с результатами оператора составила 1,95 % при среднем квадратическом отклонении в 4,22 % относительной ошибки. Этот результат, полученный в автоматическом режиме без каких-либо априорных предположений, можно признать хорошим, так как он сопоставим с результатом, получаемым при определении МНЧ с использованием модели ионосферы IRI-2007 [Грозов и др., 2012].

Выводы

В условиях непараметрической неопределенности при затрудненном использовании непараметрических критериев предлагается рассматривать смесь сигнал-шум как выборку, содержащую аномальные значения сигнала. Структура выборки представима как сумма двух распределений, соответствующих сигналу и шуму. Установлено, что в качестве индикаторов наличия сигнала в выборке могут служить ее статистические параметры. При этом для обнаружения и выделения сигнала не используются априорные предположения о виде закона распределения.

Список литературы

- Вопросы перспективной радиолокации / Коллективная монография под ред. А. В. Соколова. – М.: Радиотехника, 2002. – 512 с.
- Грозов В. П., Котович Г. В., Пономарчук С. Н.* Комплекс автоматической интерпретации данных зондирования ионосферы // XXIII Всероссийская научная конференция «Распространение радиоволн», 23–26 мая 2011 г. Сборник докладов. Йошкар-Ола: Марийский государственный технический университет, 2011. – Т. 1. – С. 192–195.
- Иванов В. А., Куркин В. И., Носов В. Е., Урядов В. П., Шумаев В. В.* ЛЧМ-ионозонд и его применение в ионосферных исследованиях // Известия высших учебных заведений. Радиофизика. – 2003. – Т. 46, № 11. – С. 919–952.
- Колчев А. А., Недопекин А. Е.* Экспериментальное подтверждение априорной непараметрической неопределенности для сигнала ЛЧМ-ионозонда в спектральной области // Вестник Национального технического университета «Харьковский политехнический институт». Сборник научных статей. Тематический выпуск: Радиофизика и ионосфера. Харьков: НТУ ХПИ, 2011. – № 44. – С. 91–100.
- Колчев А. А., Недопекин А. Е.* Обнаружение сигнала ЛЧМ-ионозонда в условиях априорной непараметрической неопределенности // Труды XVII Международной научно-технической конференции «Радиолокация, навигация, связь» (RLNC'2011). Воронеж: НПФ «САКВО-ЕЕ» ООО, 2011. – Т. 3. – С. 2469–2477.
- Колчев А. А., Недопекин А. Е., Фомин Д. А.* Развитие алгоритмов обнаружения и различения сигналов в радиофизических измерениях // Обозрение прикладной и промышленной математики. – 2006. – Т. 13, вып. 5. – С. 866–867.
- Колчев А. А., Щирый А. О.* Режекция сосредоточенных по спектру помех при ЛЧМ-зондировании ионосферы // Известия вузов. Радиофизика. – 2006. – Т. 49, № 9. – С. 751–759.
- Левин Б. Р.* Теоретические основы радиотехники. Кн. 3. – М.: Советское радио, 1976. – 288 с.
- Новицкий П. В., Зограф И. А.* Оценка погрешностей результатов измерений. – Л.: Энергоатомиздат. Ленингр. отделение, 1991. – 302 с.
- Теоретические основы радиолокации / Под ред. Я. Д. Ширмана. – М.: Советское радио, 1970. – 560 с.