УДК 621.391

Ткаченко О.М., к.т.н.

ЗАДАЧА ВИЯВЛЕННЯ СИГНАЛІВ В УМОВАХ АПРІОРНОЇ НЕВИЗНАЧЕНОСТІ ТА МЕТОДИ ЇЇ РІШЕННЯ

Tkachenko O.M. The task of detecting signals in conditions of a priori uncertainty and methods of its solution. The essence and features of optimal, adaptive and nonparametric methods of statistical processing are considered. It is expedient to use optimal (classical) methods provided that the known functional type of the distribution of the sample values and all its parameters are known. Adaptive methods are used if the distribution of input data is known to the accuracy of an array of unknown parameters. Non-parametric methods are used when the functional type of input distribution is unknown and only the general differences between situations of presence and absence of a signal are specified. Studies have shown that a priori information used in the synthesis of nonparametric findings is more qualitative than quantitative. The difference between nonparametric methods from optimal and adaptive ones is that in nonparametric methods, the main emphasis is not on optimizing the characteristics of the system, but on ensuring their insensitivity to the operating conditions.

Keywords: optimal, adaptive, nonparametric, rank, distribution, signal, quality, indicator, array, sampling

Ткаченко О.М. Задача виявлення сигналів в умовах апріорної невизначеності та методи її рішення. Розглянуто сутність та особливості оптимальних, адаптивних та непараметричних методів статистичної обробки. Дослідження показали, що апріорна інформація, яка використовується під час синтезу непараметричних виявлень - має швидше якісний, ніж кількісний характер. Відмінність непараметричних методів від оптимальних та адаптивних полягає в тому, що в непараметричних методах головний акцент ставиться не на оптимізації характеристик системи, а на забезпеченні їх нечутливості до умов функціонування.

Ключові слова: оптимальний, адаптивний, непараметричний, ранг, розподіл, сигнал, якість, показник, масив, вибірка

Ткаченко О.Н. Задача обнаружения сигналов в условиях априорной неопределенности и методы ее решения. Рассмотрены сущность и особенности оптимальных, адаптивных и непараметрических методов статистической обработки. Исследования показали, что априорная информация, которая используется при синтезе непараметрических выражений - носит скорее качественный, чем количественный характер. Отличие непараметрических методов от оптимальных и адаптивных заключается в том, что в непараметрических методах главный акцент ставится не на оптимизации характеристик системы, а на обеспечении их нечувствительности к условиям функционирования.

Ключевые слова: оптимальный, адаптивный, непараметрический, ранг, распределение, сигнал, качество, показатель, массив, выборка

Вступ

Постановка задачі. Для успішного функціонування телекомунікаційних систем необхідно забезпечити, насамперед, якісний контроль параметрів мережних елементів. Отже, фактично доводиться розв'язувати задачу в умовах апріорної невизначеності.

Аналіз літературних даних. Згідно зі статистичним трактуванням задачі виявлення вхідних даних, котрі надходять на виявник, в літературі [1-3] розглядається масив вибіркових значень *X*, що має визначений розподіл імовірностей. Випадковість масиву *X* зумовлена впливом випадкових завад, а також недосконалістю технічних засобів (похибками вимірювань або перетворень фізичних величин, шумами приймально-підсилюючих пристроїв та ін.) На підставі апріорних відомостей про властивості масиву вибіркових значень *X* як за відсутності, так і за наявності сигналу, формулюються відповідно досліджувальна та альтернативна статистичні гіпотези (тобто деякі припущення про види розподілів імовірностей *X*) [4-6].

Невирішені питання. На основі аналізу літературних джерел можна зробити *наступні висновки.* У залежності від ступеня конкретизації припущень і прийнятого критерію якості вибираються тестова статистика *V(X),* тобто деякий функціонал від вхідних даних, і порогова константа *П.* Рішення приймається шляхом порівняння значення *V(X),* отриманого по даним конкретної реалізації масиву *X* з пороговою константою *П.* Якщо $V(X) \geq \Pi$, то приймається рішення на користь альтернативної гіпотези, тобто про наявність сигналу, у протилежному випадку - на користь припущеної гіпотези - про відсутність сигналу. Таким чином, виявник повинен містити обчислювач тестової статистики *V(X)* і пороговий пристрій.

Мета та задачі дослідження. Метою роботи було дослідження методів вирішення задачі виявлення сигналів в умовах апріорної невизначеності.

Для досягнення поставленої мети вирішувались наступні задачі:

- обґрунтування можливості вирішення задачі виявлення сигналу в умовах апріорної невизначеності;
- аналіз особливостей та порівняння оптимальних, адаптивних та непараметричних методів статистичної обробки.

Виклад основного матеріалу дослідження

Результат роботи визначника можна розглядати як оцінку $\vec{\nu}$ - формального параметра ситуації ν, який дорівнює 0 при відсутності сигналу, і рівного 1 - при його наявності. Через присутність випадкових завад значення параметра $\vec{\nu}$ не завжди співпадає з дійсним значенням ν, а лише з деякою ймовірністю, яка при $\vec{\nu}$ =1 і ν =1 називається ймовірністю правильного визначення *D*, а при $\vec{\nu}$ =1 і ν =0 – ймовірністю хибної тривоги *a*. Параметри *D* і *a* характеризують якість роботи визначника, яка тим вища, чим менше значення *a* і більше *D*.

Як відомо, оптимальний демодулятор за критерієм Неймана-Пірсона максимізує ймовірність правильного прийому *D*, коли задано ймовірність хибної тривоги *a*. Значення *D* та *a* залежить від виду статистики *V(x)*. Можливість оптимального вибору *V(x)* визначається повнотою і достовірністю апріорної інформації про властивості розподілів масиву вибіркових значень *X*.

Залежно від ступеня деталізації цих повідомлень використовують наступні методи статистичної обробки.

Оптимальні методи. Оптимальні (класичні) методи використовують за умови, коли відомі функціональний вид розподілу вибіркових значень і всі його параметри. У цьому випадку найвища якість визначення за відомим критерієм Неймана – Пірсона, який забезпечує фунціонал відношення правдоподібності:

$$V(x) = \alpha(x) = \frac{f(x/\nu = 1)}{f(x/\nu = 0)}, \tag{1}$$

де $f(x/\nu)$ - спільний розподіл імовірностей масиву вибіркових значень в ситуації ν.

Тоді порогова константа *П* обчислюється як розв'язок рівняння

$$P_\lambda(\Pi / \nu = 0) = 1 - a, \tag{2}$$

де $P_\lambda(\Pi / \nu)$ - інтегральна функція розподілу статистики (1).

Адаптивні методи. Адаптивні методи використовують, якщо розподіл вхідних даних відомий з точністю до масиву невідомих параметрів δ. У цьому випадку оптимізація *V(x)* за класичними критеріями неможлива. Потрібно звертатись до використання некласичних методів оптимізації, побудованих на принципах незміщеності, інваріантності подібності тощо. Використовують також некласичні критерії якості. Але незалежно від прийнятого методу загальний результат такий: у відношення правдоподібності, обчислене при відомому масиві параметрів δ, підставляємо оптимальну оцінку $\vec{\delta}$, здобуту або безпосередньо за масивом *X*, або за спеціальним масивом *Y*. Таким чином, ця група методів відрізняється від попередньої введенням до складу визначника блока адаптації, який оцінює невідомі параметри сигналу і завади.

Непараметричні методи. Непараметричні методи використовують, коли функціональний вид розподілу вхідних даних невідомий, а задано - лише загальні відмінності між ситуаціями наявності і відсутності сигналу.

Нехай масив вибіркових значень X складається з *n*-елементів $\{x_1,...,x_n\}$. У багатьох випадках спільний *n*-мірний розподіл імовірностей вибірок $f(x_1,...,x_n/\nu)$ має властивість інваріантності до перестановок аргументів, тобто:

$$f(x_1,...,x_n/\nu) = f(x_{k1},...,x_{kn}/\nu), \tag{3}$$

де $\{k_1,...,k_n\}$ - довільна перестановка цілих чисел від *1* до *n*. Умова (3) в окремих випадках виконується, якщо

$$f(x_1,...,x_n/\nu) = \prod_{i=1}^{n} f(x_i/\nu), \tag{4}$$

тобто вибірки $\{x_1,...,x_n\}$ статистично незалежні і мають однаковий одномірний розподіл $f(x/\nu)$. Якщо умови (3) і (4) виконуються для $\nu=0$ і не виконуються для $\nu=1$, то задачу визначення сигналу можна сформулювати як перевірку виконання вказаних нерівностей. Конкретний вид розподілу при цьому знати не обов’язково.

Інформативним параметром розподілу, який не залежить від його конкретного виду, може бути властивість симетрії:

$$f(x/\nu=0) = f(-x/\nu=0), \tag{5}$$

на противагу альтернативній гіпотезі про те, що розподіл несиметричний.

Існування відмінності між розподілами масиву X у ситуаціях $\nu=0$ і $\nu=1$ можна сформулювати як гіпотезу зсуву розподілу вправо:

$$F(x/\nu=1) \le F(x/\nu=0), \tag{6}$$

де F - одномірна інтегральна функція розподілу вибірок $\{x_1,...,x_n\}$. Тоді

$$F(x/\nu) = \int_{-\infty}^{x} f(y/\nu)dy. \tag{7}$$

Фактично нерівність (6) означає, що за наявності сигналу вибіркові значення фізичних величин у середньому більші, ніж за його відсутності. Можливий також зсув розподілу вліво, якому відповідає нерівність, протилежна виразу (6). У цьому разі сигнальні вибірки будуть у середньому менші за величиною, ніж завадові. Окремим випадком гіпотези зсуву є гіпотеза зсуву середнього значення:

$$f(x/\nu) = f(x-\nu\alpha), \tag{8}$$

де *a* – деяка константа.

Дослідження показують, що апріорна інформація, яка використовується під час синтезу непараметричних виявлень - має швидше якісний, ніж кількісний характер. Непараметричні методи обробки відрізняються від класичних і адаптивних. Відмінність від останніх полягає в тому, що в непараметричних методах головний акцент ставиться не на оптимізації характеристик системи, а на забезпеченні їх нечутливості до умов функціонування. Тому, непараметричними вважають системи визначення сигналів, рівень a_0 хибних тривог яких інваріантний щодо функціонального виду розподілу завади. Оскільки рівень a_0 однозначно визначається функцією розподілу тестової статистики, то звідси випливають дві вимоги до *V(x)*: по-перше, її розподіл за відсутності сигналу має бути точно відомим і незмінним, яким би не був розподіл завади на вході системи; по-друге – у разі появи сигналу інваріантність розподілу *V(x)* має порушуватись, щоб зберігалась можливість розпізнавання ситуацій $\nu=0$ і $\nu=1$. Синтез непараметричних процедур обробки сигналів проводиться переважно евристичними методами.

Деякі закономірності цих методів: практично всі непараметричні визначники містять як складовий елемент пристрої, що виконують деяке інваріантне перетворення *S* масиву вибіркових значень *X*. У результаті цього перетворення утворюється новий масив *Z=SX*, розподіл елементів якого при відсутності сигналу ($\nu = 0$) точно відомий.

Перетворення *S*, яке вибирається евристично, дає змогу звести задачу визначення сигналу на фоні завад з невідомим розподілом до задачі перевірки простої гіпотези відносно розподілу масиву *Z*. Відповідно і синтез непараметричних визначників виконують в два етапи: на першому вибирають вид інваріантного перетворення *S*, на другому – спосіб обробки перетворюваних даних.

Одним з найпростіших прикладів перетворення є жорстке обмеження:

$$Z_i = \operatorname{sgn}(x_i) = \begin{cases} 1, x_i \geq 0 \\ 0, x_i < 0 \end{cases} . \tag{9}$$

Якщо гіпотеза симетрії (5) правильна, то неважко переконатись, що поява одиниць і нулів на виході обмежувача (9) рівноймовірна.

$$P(z_i = 1 / \nu = 0) = P(z_i = 0 / \nu = 0) = 1/2 \quad . \tag{10}$$

Отже, (9) переводить масив з довільним симетричним, відносно нульового рівня розподілом, до нового масиву. З появою сигналу симетрія розподілу порушується (що особливо важливо). Тоді задачу виявлення можна звести до перевірки простої гіпотези (10) щодо альтернативної гіпотези $P(z_i = 1 / \nu = 1) > 1/2$.

Рішенням такої задачі є критерій знаків, який передбачає додавання знаків елементів вибірки

$$V(x) = \sum_{i=1}^{n} \operatorname{sgn}(x_i) \tag{11}$$

і аналіз накопиченого значення з порогом. Статистика (11) - це кількість переваг в серії *n* незалежних досліджень, тобто підпорядковується біноміальному розподілу, причому n за відсутності сигналу ймовірність переваги дорівнює 1/2, а за його наявності перевищує 1/2. Отже, (11) задовольняє дві вище наведені умови: при $\nu = 0$ її розподіл точно відомий і незмінний, яким би не був вихідний симетричний розподіл $f(x / \nu = 0)$, а при $\nu = 1$ розподіл статистики (11) чутливий до наявності сигналу.

Окрім жорсткого обмеження, відомі також наступні види перетворень.

Перестановка елементів вибірки. Якщо справедлива гіпотеза випадковості (3), то всі перестановки елементів вибірки рівноймовірні незалежно від виду їх розподілу.

Порядкові статистики. Вихідна вибірка $\{x_1, ..., x_n\}$ упорядковується за величиною і організується в так званий варіаційний ряд, в якому $x^{(1)}$ - найменший елемент вибірки, $x^{(2)}$ - другий за абсолютною величиною елемент; $x^{(R)}$ - *R*-й за абсолютною величиною елемент; $x^{(n)}$ - максимальний елемент, тобто:

$$x^{(1)} < x^{(2)} ... < x^{(R)} ... < x^{(n)} . \tag{12}$$

Випадкова величина $x^{(R)}$ називається *R*-ою порядковою статистикою. Її інваріантна властивість полягає в тому, що зі збільшенням обсягу вибірки значення $x^{(R)}$ збігається за ймовірністю з квантованим рівнем *R/(n+1)* незалежно від виду розподілу вихідної вибірки (якщо лише для нього виконується рівність (4)).

Ранги. Рангом *i*-го елементу x_i масиву вибіркових значень X вважається порядковий номер R_i цього елементу в варіаційному ряді, тобто:

$$x_i = x^{(R_i)} . \tag{13}$$

Формально процедуру обчислення рангу можна представити у вигляді:

$$R_i = \sum_{k=1}^{n} \mathrm{sgn}(x_i - x_k). \tag{14}$$

Сукупність рангів $\{R_1,...,R_n\}$ усіх елементів вибірки $\{x_1,...,x_n\}$ утворює деяку перестановку чисел від 1 до *n*. Згідно гіпотези випадковості (3) усі такі перестановки рівноймовірні. Отже, незалежно від конкретного закону розподілу вихідної вибірки $\{x_1,...,x_n\}$ спільний розподіл рангів $\{R_1,...,R_n\}$ є рівномірним:

$$P(R_1,...,R_n) = \frac{1}{n!}. \tag{15}$$

Висновки

Апріорна інформація, яка використовується під час синтезу непараметричних виявлень - має швидше якісний, ніж кількісний характер. Непараметричні методи обробки відрізняються від класичних і адаптивних. Відмінність від останніх полягає в тому, що в непараметричних методах головний акцент ставиться не на оптимізації характеристик системи, а на забезпеченні їх нечутливості до умов функціонування.

Аналіз методів виявлення сигналу показує, що наведений перелік інваріантних перетворень не є, безумовно, вичерпним, але він підтверджує неоднозначність вибору перетворень масиву *S*. Критерієм вибору є не тільки дослідження заданих інваріантних властивостей рівня неправдоподібних тривог до виду розподілу, але і максимально можливе зберігання інформації про сигнал, що дає змогу його виявити.

Список використаної літератури

1. Толубко В.Б. Методи оптимізації / В.Б. Толубко, Л.Н. Беркман. – Київ: ДУТ, 2016. – 442 с.

2. Соломенчук В. Д. Оптические транспортные сети / В. Д. Соломенчук, В. А. Мищенко, К. Н. Гура. – Киев: Центр последипломного образования ПАО «Укртелеком», 2014. – 294 с.

3. Стеклов В.К. Оптимізація та моделювання пристроїв та систем зв'язку / В.К. Стеклов, Л.Н. Беркман, Є.В. Кільчицький. – Київ: Техніка, 2004. – 576 с.

4. Стеклов В. К. Телекоммуникационные сети / В. К. Стеклов, Л. Н. Беркман. – Київ: Техніка, 2000. – 392 с.

5. Лихтциндер Б.Я. Интеллектуальные сети связи. / Б. Я. Лихтциндер, М. А. Кузякин, А. В. Росляков, С. М. Фомичев. – Москва: Эко-Трендз, 2000. – 205 с.

6. Бертсекас Д. Сети передачи данных / Д. Бертсекас, Р. Галлагер. – Москва: Мир, 1989. – 544 с.

Автори статті

Ткаченко Ольга Миколаївна – кандидат технічних наук, доцент кафедри телекомунікаційних систем та мереж, Державний університет телекомунікацій, Київ, Україна.

Authors of the article

Tkachenko Olha Mykolaivna – candidate of science (technic), associate professor, associate professor of Department of the Telecommunication systems and networks, State University of Telecommunications, Kyiv, Ukraine.

Дата надходження в редакцію 26.03.2018 р. Рецензент: д.т.н., доцент В.Ф. Заїка