

ОБНАРУЖЕНИЕ АНОМАЛЬНЫХ ИЗМЕРЕНИЙ ПРИ АНАЛИЗЕ НЕСТАЦИОНАРНЫХ ПРОЦЕССОВ

Марчук В.И., Шерстобитов А.И.

Южно-российский государственный университет экономики и сервиса
346500, Шахты, Ростовской обл., ул. Шевченко, 147, каф. «Радиотехника информационные системы»

Поток измерительной информации представляет собой, в общем виде, реализацию нестационарного случайного процесса: совокупность функции измеряемого процесса, аддитивной шумовой составляющей, аномальных (сбойных) результатов измерений. При проведении исследования, как правило, априорная информация о характере измеряемого процесса отсутствует или крайне мала. Сложность математического описания такого рода исследуемых процессов, наличие случайных составляющих и неопределенность, как появления, так и интенсивности аномальных измерений, обуславливают большую сложность решения поставленной задачи.

Данная работа посвящена обработке результатов измерений, представленных единственной реализацией нестационарного случайного процесса при ограниченной объеме априорной информации, которые достаточно часто встречаются на практике. Упрощенная математическая модель потока измерительной информации может быть представлена в следующем виде [1]:

$$Y(t) = S(t) + \varepsilon_{ан}(t) + \varepsilon_n(t), \quad (1)$$

где $Y(t)$ - результаты измерений на выходе исследуемой системы; $S(t)$ - функция полезного сигнала (функция тренда); $\varepsilon_n(t)$ - аддитивная шумовая составляющая; $\varepsilon_{ан}(t)$ - аномальные (сбойные) измерения.

В работе [1] рассматриваются методы обнаружения аномальных измерений при анализе стационарных процессов. В случае нестационарного случайного процесса (1), в литературе отсутствуют эффективные методики обнаружения и устранения аномальных измерений.

В данной работе за основу принят метод размножения оценок, описанный в [2], который обладает высокой эффективностью при обработке единственной реализации нестационарного случайного процесса. В отличие от метода, рассматриваемого в работе [3], на каждом интервале (покрытии) исходной реализации проводится аппроксимация по методу наименьших квадратов полиномом первой степени, находится разность между оценкой и исходными данными на каждом из полученном покрытии. Для разностного процесса (остатков) на каждом покрытии определяется максимальное значение и устанавливается некоторый порог $\varepsilon = A \cdot Y_{\max}(t)$, где $0 < A \leq 1$. Превышение значений разностного процесса, на каждом из покрытий, установленного порога, штрафуются, т.е. если

$$Y(t_i) > \varepsilon, \quad (2)$$

то данное значение получает одно штрафное значение, равное 1. В соответствии с методом размножения оценок [2] вышеизложенная процедура, определения штрафов, повторяется для каждой из размноженных оценок исходной реализации. При этом происходит накопление значений штрафов для элементов исходной реализации при выполнении условия (2). По окончании обработки, для всех оштрафованных значений исходной реализации строится плотность распределения. Для заданного уровня значимости, по плотности распределения штрафов определяются, какие из измерений исходной реализации являются аномальными.

Эффективность предложенного алгоритма определялась на основе проведения имитационного моделирования. Для модели исходной реализации (1) рассматривалась шумовая составляющая, имеющая нормальный закон распределения с нулевым математическим ожиданием и дисперсией, изменяющейся по величине и характеру (стационарная по реализации и нестационарная). В качестве модели измеряемого процесса были использованы следующие нормированные функции: синусоидальная; прямоугольный униполярный импульс; экспоненциальная и сложная (модель функционирования двигательной установки). В качестве модели аномальных измерений использовались одиночные аномалии, которые располагались: равномерно по выборке, в начале, в середине и конце выборки.

На основе проведения имитационного моделирования были получены зависимости вероятности ложной тревоги $\alpha = f(R, P)$ и вероятности правильного обнаружения $\beta = f(R, P)$ аномальных измерений, где R - количество покрытий на длину реализаций; P - количество размноженных оценок исходной реализации. Анализ полученных результатов позволил сделать вывод о том, что характер зависимостей α и β , при различных видах модели измеряемого процесса, величины и характера дисперсии (стационарная и нестационарная), существенно не

изменяется. Результаты проведенных исследований (рисунок 1) позволяют сделать вывод о том, что величина α незначительно зависит от числа разложений оценок исходной реализации, а при увеличении числа разбиений на покрытия в пределах от 3 до 12 изменяется от 0,02 до 0,1.

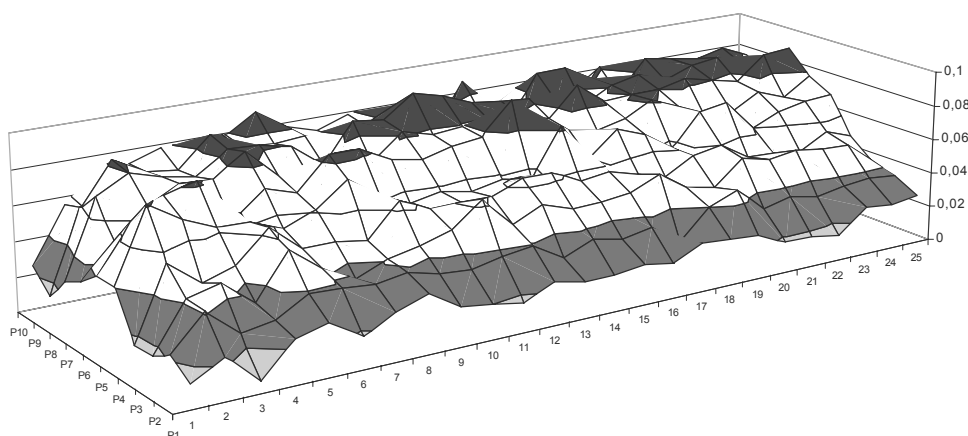


Рисунок 1

Вероятность правильного обнаружения аномальных измерений, в основном, не зависит от количества разбиения исходной реализации на покрытия, однако, наблюдается зависимость значения β от числа разложений оценок исходной реализации. На рисунке 2 представлены результаты имитационного моделирования $\beta = f(R, P)$ при изменении числа разложения оценок исходной реализации от 1 до 50 с шагом 2. Анализ полученных результатов показывает, что значение β при $P \geq 10$ практически не зависит от числа разложенных оценок исходной реализации и приблизительно равна 1.

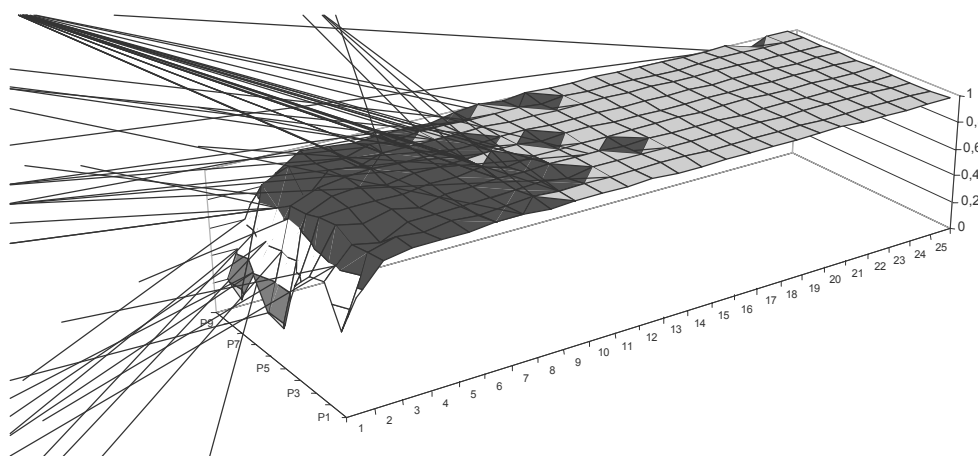


Рисунок 2

Таким образом, на основании проведенных исследований можно сделать следующие выводы:

- предложенный метод обнаружения аномальных измерений при обработке нестационарного случайного процесса, позволяет обнаружить одиночные аномальные измерения с вероятностью, практически, равной 1 независимо от их места положения в исходной реализации.

Литература

1. Отбраковка аномальных результатов измерений / А.Ф.Фомин, О.Н. Новоселов, А.В. Плющев. - М.: Энергоатомиздат, 1985. -200 с.
2. Марчук В.И. Новый подход к анализу результатов измерений В кн.: Новые технологии управления движением технических объектов: Материалы 3-й Междунар. науч.-технич.конф./Ростов-на-Дону. Изд. СКНЦ ВШ, 2000, 147 с. (129-132 с.)
3. Марчук В.И. Уланов А.П. Методы обнаружения и отбраковки аномальных результатов измерений В кн.: Изв. вузов. Сев.-Кавк. регион. Техн.науки. 2001.№2. с.7-8

Работа выполнена в рамках НИР по заданию Министерства образования РФ.

DETECTION OF ABNORMAL MEASUREMENTS IN THE ANALYSIS OF NON-STATIONARY PROCESSES

Marchuk V., Sherstobitov A.

The South - Russian state University of economies and service
346500, Schachty, Rostov region, Shevchenko street 147 the faculty of Radio engineering informational systems

A lot of measuring information may be represented as a realization of non-stationary casual process, i.e. function of a measured process of additive sound compound, abnormal results of measurements. It may be stated that carrying out a research, same times the aprioristic information on character of measured process is absent or is smallest. Complexity of mathematical description of some researched processes, in case of casual compounds and uncertainty, occurrences, and intensity of abnormal measurements, may being difficulty to decision of a task in view.

This work is devoted to appearance processing of measurement results represented a unique realization of non-stationary casual process under limited of information volume which is enough often used in practice. The simplified mathematical model of the measuring information can be submitted as [1]: $Y(t) = S(t) + \varepsilon_{an}(t) + \varepsilon_n(t)$, where $Y(t)$ - output results of measurements of the researched system; $S(t)$ - function of a useful signal (function of a trend); $\varepsilon_n(t)$ - additive sound compound; $\varepsilon_{an}(t)$ - abnormal measurements.

In work [1] methods of detection of abnormal measurements are considered to be in the analysis of stationary processes. In case of non-stationary casual process $Y(t) = S(t) + \varepsilon_{an}(t) + \varepsilon_n(t)$, we see that there are no effective techniques of detection and elimination of abnormal measurements in literature.

In this work an a basis the method of duplication estimations, described [2] which has high efficiency at processing unique realization of non-stationary casual process is accepted. Unlike method we considered in work [3], in each interval of initial realization approximation on a method of the least squares is carried out by a polynom of the first degree as there is a difference between an estimation difference and the initial data on each of the obtained covering. For different sided process on each covering maximum value may be defined and some threshold $\varepsilon = A \cdot Y_{\max}(t)$, where $0 < A \leq 1$ may be established. Excess of values different sided process, on each of the coverings, the established threshold, is fined, i.e. if $Y(t_i) > \varepsilon$ the given value receives one penalty value that is equal to 1. According to a penalty method of duplication of estimations [2] the mentioned procedure of definitions of penalties can be repeated for each of multiplied estimations of initial realization. Thus there is an accumulation of values of penalties for elements of initial realization under condition $Y(t_i) > \varepsilon$. In the end of processing, for all fined values of initial realization the density of distribution may be calculated. For low level of significance a density of distribution of penalties is determined, from measurements of initial realization are considered to be abnormal.

The efficient algorithm was defined on the basis of carrying out imitating modelling. As a model of initial realization is $Y(t) = S(t) + \varepsilon_{an}(t) + \varepsilon_n(t)$, a sound compound due to normal law of distribution with a zero population mean and a dispersion, changing on size and character (stationary on realization and non-stationary) was considered. As model of measured process the following normalized functions have been used: of sine wave; as a rectangular unipolar pulse; that is exponent and difficult (model of functioning of impellent installation). As model of abnormal measurements single anomalies which have bun settled down were used: in regular intervals of sample, in the beginning, in the middle and the end of sample.

As a result of carrying out imitating modelling dependences of probability of a false alarm, we have $\alpha = f(R, P)$ and probability of correct detection $\beta = f(R, P)$ of abnormal measurements, where R - quantity(amount) of coverings for length of realizations have been defined; P - amount of the multiplied estimations of initial realization. The analysis of the result let as conclude that character of dependences α and β , at various kinds of model in a measured process, sizes and character of a dispersion (stationary and non-stationary), are essentially unchangeable. As the results of the researches we may say that the size α insignificantly depends on the number of duplication estimations of initial realization, and the increase of number of splittings at coverings in limits from 3 up to 12 changes from 0,02 up to 0,1.

The probability of correct detection in abnormal measurements, basically, does not depend on amount of splitting of initial realization on coverings dependence on value β on number of duplication estimations of initial realization, however, it is observed at $P \leq 10$. The analysis of the results show, that

value β at $P \geq 10$ practically does not depend on number of the multiplied estimations of initial realization and is approximately equal to 1.

Thus, on the basis of the research it is possible to say that:

- the suggested method of detection of abnormal measurements at processing non-stationary casual process, allows to find out single abnormal measurements with probability, that is , equal to 1 inspite of irrespective of their of position in initial realization.

The literature

1. Rejection of abnormal results of measurements / A.F.Fomin, O.H. New settlers, A.V.Pljushchev.
2. Marchuk V.I. New the approach to the analysis of results of measurements. New technologies of management of movement of technical objects: Materials of 3-rd World Scientific Conference /Rostov-on-Don.
3. Marchuk V.I. Ulanov A.P.method of detection and rejection of abnormal results of measurements In кн.: Изв. High schools. Сев.-Кавк. Region. Техн.науки. 2001. № 2. with 7-8