

ПОВЫШЕНИЕ ДОСТОВЕРНОСТИ ПЕРВИЧНОЙ ОБРАБОТКИ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

Марчук В.И.

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

Целью обработки сигналов является решение двух задач: определение истинных параметров сигнала, искаженного действием помех или среды, через которую передается сигнал, и оценка влияния характеристик этой среды на параметры сигнала. Характерной особенностью методов измерения является многоэтапность, причем, каждый этап измерений вносит свои систематические составляющие погрешности. Выявление и оценка этих составляющих погрешности становится основной задачей обработки данных. В результате, обработка данных превращается в сложную комплексную задачу, требующую для своего решения привлечения разнообразных методов и средств.

Под первичной обработкой данных (data cleansing and scrubbing) понимается процедура "очистки" исходных данных, заключающаяся в устранении избыточности, противоречивости и в очищении от шума перед помещением в склад данных. Более сложная обработка может включать восстановление пропущенных значений в исходных данных.

Поток измерительной информации представляет собою, в общем виде, совокупность полезного сигнала (измеряемой функции), систематических погрешностей результатов измерений (приборные и модельные), сбойных (аномальных или грубых погрешностей) значений результатов измерений, случайных погрешностей и ряда других. При исследовании новых изделий (уникальность, изменение условий испытаний и др.), априорная информация о системе исследования отсутствует или очень бедна. Методы решения таких задач начали разрабатываться только в последние годы, результаты здесь скромные, что объясняется чрезвычайной ее трудностью. Сложность математического описания исследуемых систем, нестационарность и нелинейность процессов, протекающих в них, случайный характер воздействующих возмущений обуславливают большую сложность задачи построения их математической модели.

В данной работе речь идет в основном об обработке результатов измерений представленных единственной реализацией нестационарного случайного процесса при ограниченной объеме априорной информации, которые достаточно часто встречаются на практике, математическая модель при этом еще более усложняется. В связи с этим предлагается в качестве исходной модели потока измерительной информации использовать упрощенную математическую модель, которая может быть записана в следующем виде

$$Y(t) = S(t) + \varepsilon_{ан}(t) + \varepsilon_{н}(t),$$

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

Для обработки результатов измерений, которые могут быть представлены данной математической моделью, предлагается воспользоваться структурной схемой, представленной на рисунке 1. На каждом из этапов обработки, в основе, используется новый метод, основанный на размножении оценок, исследуемого нестационарного процесса (результатов измерений) [1].

Для получения оценок полезного сигнала, в предлагаемом методе производится разбиение исходного временного отрезка $[t_1, t_n]$ (длительности реализации результатов измерений) на случайные промежутки (покрытия) точками, равномерно распределенными на данном отрезке, с последующей проверкой того, чтобы длина покрытий была не менее некоторого значения l . На каждом покрытии определяются оценки коэффициентов аппроксимирующего полинома $a + bt + ct^2$ методом наименьших квадратов. Следовательно, оценка полезного сигнала представляет собою последовательность полиномов второй степени полученных для каждого из покрытий. Таким образом получается первая оценка функции полезного сигнала. Разбивая исходную реализацию другими случайными числами, распределенных по равномерному закону, и производя аппроксимацию на каждом из покрытий полиномом второй степени методом наименьших квадратов, получаем вторую оценку полезного сигнала. Следовательно, повторяя вышеперечисленные действия, можно получить любое количество оценок полезного сигнала по единственной реализации результатов измерений. При этом следует отметить, что вероятность того, что границы покрытий при размножении оценок совпадут практически равна нулю [2]. При обнаружении аномальных (сбойных) измерений

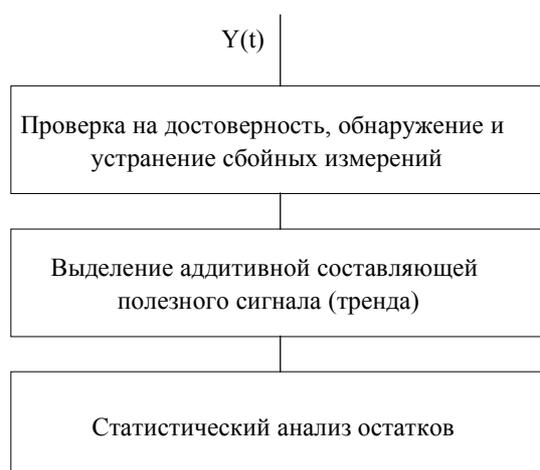


Рисунок 1.

на каждом покрытии относительно полученной оценки полезного сигнала оценивается значение дисперсии шумовой составляющей, определяются пороги, относительно которых штрафуются те измерения, значения которых их превышают. Данная процедура производится для каждой из полученных оценок полезного сигнала. Таким образом, получают массив штрафов для каждого из измерений. Те измерения, которые получили наибольшее число штрафов, считаются аномальными (сбойными) и исключаются из дальнейшего анализа [3].

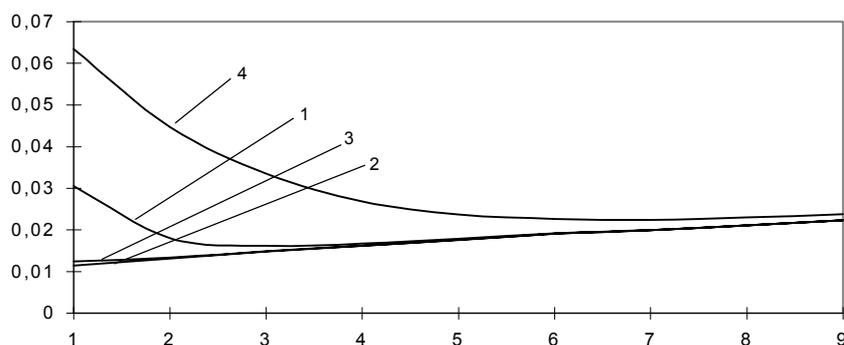


Рисунок 2

Очищенный от аномальных (сбойных) измерений исходная реализация обрабатывается методом размножения оценок, с целью очищения от шумовой составляющей [4]. Проведенные исследования показали, что метод размножения оценок обладает свойством инвариантности к закону распределения шумовой составляющей и его можно рассматривать как метод свободный от модели полезного сигнала. Данное предположение подтверждается результатами имитационного моделирования, результаты которого приведены на рисунке 2, где приведена зависимость среднеквадратической ошибки выделения полезного сигнала при наличии шумовой составляющей, имеющей нормальный закон распределения, для различных моделей полезного сигнала (1- синусоидальный, 2 - параболический, 3 - экспоненциальный, 4- сложный (модель работы двигательной установки)). Как видно из рисунка 2, с ростом дисперсии шумовой составляющей, среднеквадратическая погрешность стремится к одному значению. Полученный результат представляет значительный интерес в связи с тем, что погрешность выделения для различных моделей сигналов одна и та же и можно высказать предположение о том, что метод является свободным от модели полезного сигнала. На основании этого факта можно высказать гипотезу о том, что погрешность выделения полезного сигнала определяется погрешностью вносимой шумовой составляющей. Иначе говоря, погрешность выделения полезного сигнала, при наличии аддитивной шумовой составляющей, становится значительно меньшей, чем погрешность вносимая структурой самой аддитивной шумовой составляющей. Это предположение подтверждается результатами, представленными на рисунке 3, где приведены зависимости \overline{m}_{CKO} (среднее значение среднеквадратической ошибки – кривая 1), \overline{m}_{CAO} (среднее значение средней абсолютной ошибки –

кривая 2), $\bar{\sigma}_{CAO}$ (кривая 3), $\bar{\sigma}_{CKO}$ (кривая 4) при нормальном законе распределения шумовой составляющей (дисперсия изменяется от 0,02 до 0,14). Как видно из результатов, представленных на рисунке 3, все кривые имеют линейную зависимость от дисперсии шумовой составляющей.

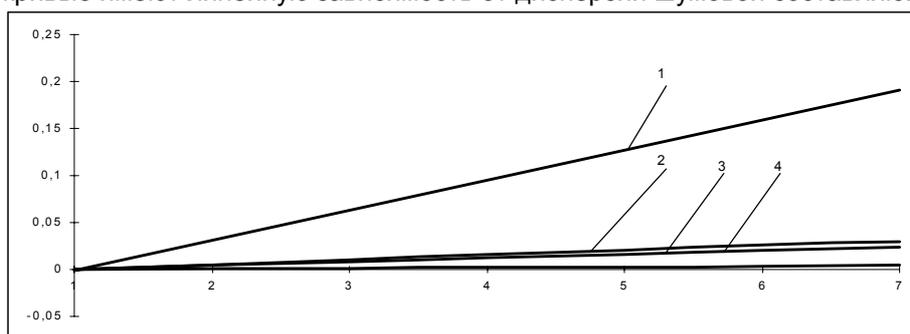


Рисунок 3.

На основании вышеизложенного можно сделать следующие выводы:

- Приведенная структура первичной обработки результатов измерений может быть рекомендована для практического применения при ограниченном объеме априорной информации о функции измеряемого процесса и статистических характеристиках аддитивной шумовой составляющей и аномальных (сбойных) измерений.
- Предложенный метод размножения является инвариантным к статистическим характеристикам аддитивной шумовой составляющей и его можно рассматривать, как метод свободный от модели измеряемого процесса.

Литература

1. 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.)
2. Марчук В.И. Новый метод обработки результатов измерений при ограниченном объеме априорной информации. В кн.: Сборник материалов международного симпозиума "Аэрокосмические приборные технологии" (АПТ 2002, 17-20 сентября 2002 г.), Санкт-Петербург, ГУАП, 2002. -160с. (97-98 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
4. Марчук В.И., Саакян О.В., Саакян Г.Р. Пространственно-временная аппроксимация при выделении сигнала для случая априорной недостаточности// Изв. вузов. Сев.- Кавк. регион. Техн. науки. 1999. №3. С. 67 – 70.

Работа выполнена в рамках НИР по заданию Министерства образования РФ.



INCREASE OF RELIABILITY OF PRIMARY PROCESSING OF RESULTS OF MEASUREMENTS

Marchuk V.

South - Russian state university of economy and service
346500, Schachty, Rostov region, Shevchenko street 147

The aim of processing of signals may be represented as a decision of two problems: the definition of true parameters of a false or a deformed signal under the influence of environmental sounds where the signal is transmitted and estimation of the parameters of a signal influence of environmental characteristics of a signal parameters is made. A specific feature of methods of measurement is its multistaged. Each stage of measurement has systematic errors. Revealing and estimation of error components is the main task of data processing. As a results of it data processing is a complex process. It is necessary to pay special attention to various research methods

By doing data cleansing and scrubbing we understand it as the procedure of clearing of initial data. This process is consisted of the process of elimination of redundancy, discrepancy and clarification from noise before placing it into the data bank. More complex processing may include the process of restoration of missed meanings in initial data.

A lot of measuring information may be represented as a kind set of a useful signal of a measured function. Any regular errors in the results of measurements of abnormal meanings and casual errors also may be considered as a useful signal. In research of new products we have less aprioristic information or this information is ratlur poor so we need new methods. These methods to solve such problems have started to be developed last years so the results leave much to be desired. This fact may be explained from the point of complexity of mathematical description of researched systems. Due to non-stationary and non linear processes occurred in them and casual character of influence, so it is very difficult to do a mathematical model.

The work of processing of results of measurements submitted by a unique realization of non-stationary casual process in a case of limited aprioristic information. Mathematical model is very complicated so as it is offered to use a simple mathematical model which can be represented as $Y(t) = S(t) + \varepsilon_{an}(t) + \varepsilon_n(t)$, where $Y(t)$ - the results of measurements of an output of a researched system as initial model of a current measuring information; $S(t)$ - function of a useful signal; $\varepsilon_n(t)$ - additive sound a component; $\varepsilon_{an}(t)$ - abnormal measurements.

On each of processing stage as the basis of a new method duplication of estimations for non-stationary process (results of measurements) [1] is used.

For estimating of useful signal by this method a splitting of a initial time $[t_1, t_n]$ duration into intervals by regular points is made. It must be stated that the following control of the length of intervals was not less that l . On each interval estimations of factors of an approximating polynom $a + bt + ct^2$ will be defined by a method of the least squares. Hence, the estimation of a useful signal is represented as sequence of polynoms of the degree obtained for each interval. Thus the first estimation of function of a useful signal is obtained. Initial realization of other random numbers, equally distributed by making approximation for each interval of polynom of the second degree by least squares method we may have the second estimation of a useful signal. So repeating all stages we may get any quantity of estimations of a useful signal in realization of results of measurements. Thus it is necessary to note, that probability of the borders of intervals will coincide. It is equal to zero [2]. At detection of abnormal measurements for each interval according to this estimation of a useful signal the meaning of a sound dispersion is estimated by a component. Thresholds for measuring of these meanings may be found. The exceed meanings of thresholds are fined.

This procedure is made for each of received estimations of a useful signal. Thus, a file of penalties for each of measurements is obtained. Those measurements which have received the greatest number of penalties, are considered to be abnormal and are excluded from the further analysis [3].

Having cleared from abnormal measurements initial realization is processed by a method of duplication of estimations, with the purpose of clarification from a sound component [4]. The results of researches have shown, that the method of duplication of estimations has a property of invariancy to the law of distribution of a sound component and it can be considered as a method free from model of a useful signal. The given assumption proves to be real results of imitating modelling. The growth of a dispersion of a sound component, the mean error aspires to one meaning. This result is of significant interest as an error of allocation for various models of signals is the same and it is possible to conclude that the method is free from model of a useful signal. It is possible to state that the error of allocation of a useful signal is defined by an error brought by a sound component. In other words, an error of allocation of a useful signal (in additive

sound a component) becomes significant smaller, than an error brought by a structure of the most additive sound component.

On this basis it is possible to make conclusions:

- The resulted structure of primary processing of results of measurements can be recommended for practical application at the limited volume of the aprioristic information about function of measured process and statistical characteristics of additive sound component and abnormal measurements.
- Method of duplication is an invariant to statistical characteristics of additive sound a component and it can be considered, as a method free from model of measured process.

The literature

1. Marchuk V.I. New the approach to the analysis of results of measurements In кн.: New technologies of management of movement of technical objects: Materials of 3-rd World Scientific Conference /Rostov-on-Don., 2000, 147 with. (129-132 with.)
2. Marchuk V.I. New a method of processing of results of measurements at the limited volume of the aprioristic information. In кн.: the Collection of materials of the international symposium " Space instrument technologies " (АПТ 2002, on September, 17-20, 2002), Saint Petersburg, 2002.-160с. (97-98 with.)
3. Марчук V.I. Уланов 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
4. Marchuk V.I., Saakjan O.V., Саакян G.R.existential approximation at allocation of a signal for a case of aprioristic insufficiency // Изв. High schools. Сев.-Кавк. Region. Техн. Sciences. 1999.№3. With. 67 – 70.