ENSURING THE AUTHENTICITY OF DATA PROCESSING ON THE BASIS OF IDENTIFICATION OF NON-STATIONARY OBJECTS IN CONDITIONS OF UNCERTAINTY OF REGULAR ERROR FACTORS

G. I. Zaripova
Samarkand State University,
140104, Samarkand, Uzbekistan

Probability structure of information authenticity is the main reason for complexity of solving the problems of identification of non-stationary objects; improve the performance of automated control systems to technological processes, and also ensuring of efficiency of data processing. Thus, significant factors that reduce authenticity of data transfer and processing become non-stationarity of processes, insufficiency of priory knowledge and large parametrical uncertainty in models to describing objects.

In this context, development of methodical bases to constructing methods and software-algorithmic complexes for improve the authenticity of data processing, taking into account kind, properties, distribution laws, regular error represent actual scientific and technical problem.

In traditional approaches to development of methods to improve the authenticity of information the solutions of tasks are gotten on the basis of statistical and dynamic modeling, algorithmic implementation and experimental studies with extensive a priory data. Thus the most typical statistical characteristics of factors used for estimation of the regular error are mathematical expectation, mean-squared deviation, distribution laws, and auto correlation functions, coefficients of pair and mutual correlation connections, other dynamic characteristics of random time series (RTS), describing non-stationary objects.

Feature of the present issue is the development of methodical bases for multivariate analysis of regular error function in methods of RTS' identification and approximation on the basis of mechanisms to revealing and use statistical, dynamic characteristics of information and probabilities of errors' distribution with limited retrospective data of information process.

The offered approach to ensure the authenticity of data assumes search of extremum of influences function, creating in future an opportunity to evaluate the minimal regular error of RTS identification. Thus the regular error of identification on model of non-stationary objects is represented by the sum of additional factor errors caused at each stage of information transformation.

The technique of multivariate analysis is developed on the basis of identification of non-stationary objects in view of estimates of influence degree on a regular error on all stages of information processes. The structural components of the influencing factors are considered. The models and algorithms of RTS identification are developed on the basis of account of a regular error according to given technique.

Functions in the form of regression dependences are used as model of identification and approximations of non-stationary object and in them estimates of error are defined by parameters of mathematical expectation and mean-squared deviation with the normalized level of influencing factor coefficient's sign.

The perspective and effective approach is offered to improve the authenticity of data processing by overlapping opportunities of statistical and dynamic models of identification with methods to estimate influence factors on a regular error. Questions of synthesis of statistical and dynamic models for non-stationary objects identification are investigated as the basic fundamentals of methods to improve the authenticity of non-stationary objects' data.

With the purpose of a regular error reduction during RTS identification the technique is offered to use the method checking observance the balance ratio entered into structure of dynamic model of non-stationary object. To optimize the analysis and processing of data in structure of dynamic model for non-stationary object identification the additional balance ratio are entered, and they are set based on normative requirements revealed on a long enough time interval. The mathematical model is formalized for identification of non-stationary object with procedures of check of observance of balance ratio. The check of observance of balance ratio in dynamic models of identification of non-stationary object is carried out by method of target shift of RTS sequence values inside a range of probability distribution with account iteratively, conditions of non-stationarity and parametrical uncertainty.

The general solution of task is gotten in the form of continuous and differentiable by all variable equations. The algorithm is developed for identification of non-stationary objects by correcting balance ratio with linear dynamic equations.

The method to dynamic identification is investigated under various distribution laws for a regular error and properties of RTS non-stationarity. The composition of specific characteristics of input variable, models and algorithms to control of a regular error, adjustment and correction of parameters of model are determined and also obtained estimates of minimization of RTS dispersion and target parameters are investigated. It is proved, that realization of methods to synthesis models of multivariate influences of a regular error with dynamic model of non-stationary object, methods to improve information authenticity by check and correction of balance ratio contribute to achieving improve the stability of identification, and also efficiency of data analysis and processing.

Key words: non-stationary object, hybrid identification, authenticity, regular error, factors of influence, multifactor analysis, dynamic model, balance control.

References

1. Zaripova G. I. Adaptivniy kontrol dostovernosti texnologicheskix parametrov na osnove modeley nechetkogo vivoda // Materiali VIII mejdunarodnoy nauchno-prakticheskoy konferenciyi „Nauchnaya diskussiya: voprosi texnicheskix nauk“, Moskva: „Mejdunarodniy centr nauki i obrazovaniya“, 2013. P. 31–37.

2. Miroshnik I. V., Nikiforov V. O., Fradkov A. L. Nelineynoe i adaptivnoe upravleniye slojnimi dinamicheskimi sistemami. S-Pb.: Nauka, 2000.

3. Mif N. P. Modeli i osenki pogreshnosti texnixheskix izmereniy. M.: Standarti, 1976.

4. Igamberdiyev X. Z., Sevinov J. U., Zaripov O. O. Regulyarniye metodi i algoritmi sinteza adaptivnix system upravleniya s nastraivayemimi modelyami. T.: TashGTU, 2014.

5. Karabutov N. N. Adaptivnaya identifikasiya system. Informasionniy sintez. M.: Kom Kniga, 2006.

6. Zaripova G. I., Akhatov A. R. Methods and Algorithms to Control Information Authenticity during Transfer and Handling of Data of Continuous Objects on the basis of a Neuro-Fuzzy Network // 2013 International Conference in Central Asia on Internet (ICI), Tashkent, 8–10 October 2013, Section 7, IEEE. Tashkent, 2013. P. 12–18.

7. Jumanov I. I., Abdullayev A. N. Kontrol tochnosti peredachi informasii v sistemax avtomatizasii izmereniya I obrabotki dannix nestasionarnoy prirodi // „Intellektualniye sistemi dlya industrialnoy avtamatizasii“ WCIS – 2006, TGTU, Tashkent. P. 213–218.

8. Yarushkina N. G. Osnovi nechotkix i gibridnix sistem. Uchebnoye posobiye. M.: Finansi i statistika. 2004.

ОБЕСПЕЧЕНИЕ ДОСТОВЕРНОСТИ ОБРАБОТКИ ДАННЫХ НА ОСНОВЕ ИДЕНТИФИКАЦИИ НЕСТАЦИОНАРНЫХ ОБЪЕКТОВ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ ФАКТОРОВ СИСТЕМАТИЧЕСКОЙ ПОГРЕШНОСТИ

Г. И. Зарипова

Самаркандский государственный университет,
140104, г. Самарканд, Узбекистан

УДК 658.512.011

Сформулирована задача и разработаны методы повышения достоверности информации на основе синтеза методов статистической, динамической, нечеткой идентификации, порогового контроля, контроля по приращениям и с предсказанием, оценки влияния факторов на систематическую погрешность и механизмов настройки параметров моделей для оптимизации обработки данных нестационарных объектов. Предложены методы многофакторного анализа для повышения эффективности идентификации и аппроксимации объектов, представляющие данные в виде случайных временных рядов. Разработанные методы рекомендованы для реализации в виде программно-алгоритмических комплексов обеспечения достоверности обработки данных на основе гибридной идентификации с учетом нелинейности влияющих факторов и неопределенности параметров в динамических моделях.

Ключевые слова: нестационарный объект, гибридная идентификация, достоверность, систематическая погрешность, факторы влияния, многофакторный анализ, динамическая модель, балансовый контроль.

Актуальность темы. Качество функционирования автоматизированных систем управления технологическими процессами, в том числе эффективность методов обработки данных, в большой степени зависит от достоверности сообщений, формируемых на этапах измерения, ввода, передачи информации. Вероятностная структура достоверности информации является основной причиной сложности решения задач идентификации и аппроксимации нестационарных объектов при недостаточных априорных сведениях и большой параметрической неопределенности моделей, обусловленных влиянием факторов в систематической погрешности на всех этапах переработки информации [1, 2].

Необходимо отметить, что в существующих системах управления технологическими процессами разработка методов повышения достоверности информации нестационарных объектов на основе идентификации переходных процессов, оценки влияния факторов на систематическую погрешность и получение на этой основе механизмов оптимизации обработки данных является перспективным направлением исследований [2, 3]. В этой связи разработка методических основ построения программно-алгоритмических комплексов повышения точности обработки данных для анализа и прогнозирования случайных временных рядов (СВР) с учетом вида, свойств, законов распределения, параметрической неопределенности в моделях идентификации и систематической погрешности представляют актуальную научно-техническую проблему.

В традиционных подходах к разработке методов повышения достоверности информации получены решения задач проектирования адекватных моделей идентификации СВР, повышения точности информационных процессов на основе статистического и динамического моделирования, алгоритмической реализации и проведения экспериментальных исследований при обширных априорных данных. При этом наиболее типичными статистическими характеристиками факторов, используемых для оценки систематической погрешности, а, следовательно, и эффективности алгоритмов контроля достоверности информации, являются математическое ожидание, среднеквадратическое отклонение, законы распределения, автокорреляционные функции, коэффициенты парных, взаимных корреляционных связей и другие динамические характеристики [3].

Особенностью проводимого в настоящей работе исследования является разработка методических основ многофакторного анализа нелинейной функции систематической погрешности, методов повышения эффективности идентификации и аппроксимации нестационарных объектов, представляемых в виде СВР на основе усовершенствованных механизмов выявления и использования статистических, динамических характеристик информации и вероятностей распределения погрешностей при ограниченных ретроспективных данных информационного процесса.

Методика анализа влияний структурных факторов на систематическую погрешность. Предлагается подход, по которому предполагается поиск экстремума функции влияний, создающий в последующем возможность оценить минимальную систематическую погрешность идентификации в системах анализа и обработки данных. Систематическая погрешность идентификации по модели нестационарных объектов представляется суммой дополнительных факторных погрешностей, вызываемых на каждом этапе преобразования информации [4, 5].

В [6] разработаны методики детерминированного и многофакторного анализов для оценки функций каждого из влияющих факторов погрешности. Важными результатами работы являются выражения оценки коэффициентов степеней влияния A_i и коэффициентов эластичности факторов $Э_i$, полученные для широкого спектра аналитических зависимостей.

Предлагаемая в настоящей работе методика многофакторного анализа основывается на идентификации нестационарных объектов с учетом оценок степени влияния на систематическую погрешность на всех этапах информационных процессов. В связи с этим рассмотрим следующие структурные составляющие влияющих факторов:

– погрешность данных, связанная с искажениями в информации на этапах измерений, ввода и передачи (фактор A_1);

– погрешность, обусловленная нестационарностью СВР (фактор A_2);

– погрешность из-за неадекватности моделей идентификации и аппроксимации, обусловленная неопределенностью параметров (фактор A_3).

Оценку систематической составляющей погрешности модели идентификации нестационарного объекта с учетом факторов A_1, A_2, A_3 представим как

$$\Delta = \overline{\Delta}_c = \Delta(A_1) + \Delta(A_2) + \Delta(A_3). \tag{1}$$

Отражение взаимозависимости факторов A_1, A_2, A_3 определим через назначение трех уровней влияющих факторов, среди которых первый принимается как основной постоянный, а два остальных являются дополнительными переменными. Для получения оценок влияний формируются матрицы комбинаций ортогональных пересечений с учетом неизменяемости первого основного фактора.

Возможные комбинации задаются девятью значениями погрешностей, где верхний индекс указывает уровень влияющего фактора, а нижний — номер вида влияющих факторов:

$$\left.\begin{matrix} \Delta(A_1',A_2',A_3'); & \Delta(A_1',A_2',A_3''); & \Delta(A_1',A_2',A_3''') \\ \Delta(A_1',A_2'',A_3'); & \Delta(A_1',A_2'',A_3''); & \Delta(A_1',A_2'',A_3''') \\ \Delta(A_1',A_2''',A_3'); & \Delta(A_1',A_2''',A_3''); & \Delta(A_1',A_2''',A_3''') \end{matrix}\right\}. \tag{2}$$

Усредненное значение погрешности для учета влияния первого фактора первого уровня, исходя из (1), вычисляется как

$$\overline{\Delta}(A_1') = \frac{1}{3}[\Delta(A_1',A_2',A_3') + \Delta(A_1',A_2',A_3'') + \Delta(A_1',A_2',A_3''')]. \tag{3}$$

Аналогично вычисляются средние значения погрешностей для второго и третьего уровней основного влияющего фактора влияния

$$\begin{matrix} \overline{\Delta}(A_1'') = \frac{1}{3}[\Delta(A_1',A_2'',A_3') + \Delta(A_1',A_2'',A_3'') + \Delta(A_1',A_2'',A_3''')] \\ \overline{\Delta}(A_1''') = \frac{1}{3}[\Delta(A_1',A_2''',A_3') + \Delta(A_1',A_2''',A_3'') + \Delta(A_1',A_2''',A_3''')] \end{matrix} \tag{3'}$$

Для редукции избыточных связей при многофакторном анализе целесообразно применение абсолютной разности дополнительных и основного влияющих факторов, которая оценивается как

$$\left[\overline{\Delta}(A_i''') - \overline{\Delta}(A_i')\right] - \left[\overline{\Delta}(A_i'') - \overline{\Delta}(A_i')\right] = \overline{\Delta}(A_i''') - 2\overline{\Delta}(A_i''). \tag{4}$$

Если разность (4) при изменении величин от A_i' до A_i''' несущественна, то функция систематической погрешности с учетом влияния i-го фактора представляется линейной стохастической зависимостью первого приближения

$$P(A_i) = K(A_i - A_{=i}), \tag{5}$$

где $A_{=i}$ — нормальное значение i-ой влияющей величины;

K — коэффициент коррекции линейной оценки i-ой влияющей величины, которая задается в виде

$$\tilde{K} = \sum_m \frac{\overline{\Delta}(A_i^m)}{A_i^m}, m = \{','','''\}. \tag{6}$$

В случае нелинейности функции влияния принимается, что

$$P(A_1',A_1'',A_1''') = \Delta(A_1',A_1'',A_1''') - \overline{\Delta_c}, \tag{7}$$

и результаты нормирования каждого A_i при многоуровневом влиянии представляются в виде схемы

$$\left.\begin{matrix} \lg P(A_1',A_2',A_3'); & \lg P(A_1',A_2',A_3''); & \lg P(A_1',A_2',A_3''') \\ \lg P(A_1',A_2'',A_3'); & \lg P(A_1',A_2'',A_3''); & \lg P(A_1',A_2'',A_3''') \\ \lg P(A_1',A_2''',A_3'); & \lg P(A_1',A_2''',A_3''); & \lg P(A_1',A_2''',A_3''') \end{matrix}\right\}. \tag{8}$$

Тогда функции оценки влияния для трех факторов влияния A_i на примере A_1 представляются в виде

$$
\begin{aligned}
P_1(A_1^{'}) &= \frac{1}{3}[\lg P(A_1^{'},A_2^{'},A_3^{'''}) + \lg P(A_1^{'},A_2^{''},A_3^{'}) + \lg P(A_1^{'},A_2^{'''},A_3^{''})];\\
P_1(A_1^{''}) &= \frac{1}{3}[\lg P(A_1^{'},A_2^{''},A_3^{'}) + \lg P(A_1^{'},A_2^{''},A_3^{''}) + \lg P(A_1^{'},A_2^{''},A_3^{'''})]; \qquad (9)\\
P_1(A_1^{'''}) &= \frac{1}{3}[\lg P(A_1^{'},A_2^{'''},A_3^{'}) + \lg P(A_1^{'},A_2^{'''},A_3^{''}) + \lg P(A_1^{'},A_2^{'''},A_3^{'''})].
\end{aligned}
$$

Очевидно, что учет разности нормирования для комбинаций (2) позволит получить более точную оценку систематической погрешности трехуровневых влияний.

В связи с этим зададим коэффициент $\bar{C}$, который определяется аналогично выражениям (3), (4), исходя из схемы

$$
\left.\begin{matrix}
C(A_1^{'},A_2^{'},A_3^{'}); & C(A_1^{'},A_2^{'},A_3^{''}); & C(A_1^{'},A_2^{'},A_3^{'''})\\
C(A_1^{'},A_2^{''},A_3^{'}); & C(A_1^{'},A_2^{''},A_3^{''}); & C(A_1^{'},A_2^{''},A_3^{'''})\\
C(A_1^{'},A_2^{'''},A_3^{'}); & C(A_1^{'},A_2^{'''},A_3^{''}); & C(A_1^{'},A_2^{'''},A_3^{'''})
\end{matrix}\right\}, \qquad (10)
$$

где

$$
C(A_i^m) = \lg P(A_i^m) - \Delta_i^m, i = (1,2,3), m = \{',\,'',\,'''\}.
$$

В результате стохастическая функция совместного влияния всех факторов на систематическую погрешность определяется как

$$
P(A) = \bar{C} P_1(A_1^m) P_2(A_2^m) P_3(A_3^m). \qquad (11)
$$

Отметим, что в дополнение к (11) в качестве функции статистической зависимости трехуровневого влияния факторов $P_1(A_1^{'})$, $P_2(A_2^{''})$, $P_3(A_3^{'''})$ могут применяться различные аналитические зависимости, которые способствуют повышению точности оценки идентификации СВР.

Предложим оценку коэффициента статистической зависимости основных факторов от дополнительных вычислить как

$$
\overline{P}(A_1^i,A_2^j,A_3^l) = \frac{P(A_1^i,A_2^j,A_3^l)}{P(A_1^i,A_2^l,A_3^j)\cdot P(A_1^l,A_2^j,A_3^i)}. \qquad (12)
$$

Тогда результативная функция оценки степени влияния трехуровневых факторов получается путем усреднения значений составляющих и представляется для оптимизации анализа и обработки данных в виде

$$
\tilde{P}(A) = \bar{C} \bar{P}_1(A_1^m) \bar{P}_2(A_2^m) \bar{P}_3(A_3^m). \qquad (13)
$$

Показатель точности анализа и обработки данных оценивается по критерию минимального среднеквадратического отклонения, который является функцией, зависящей от усредненного коэффициента $\bar{C}$, степени влияния $\tilde{P}(A_i)$ при факторах $\bar{P}_1(A_1^m)$, $\bar{P}_2(A_2^m)$, $\bar{P}_3(A_3^m)$.

Теперь изложим результаты разработки моделей и алгоритмов идентификации СВР на основе учета систематической погрешности по приведенной методике.

Анализ и оценка систематической погрешности на основе статистических моделей идентификации СВР. В качестве модели идентификации и аппроксимации нестационарного объекта использованы функции в виде регрессионных зависимостей, в которых оценки погрешности определяются по следующим параметрам:

— математическое ожидание

Таблица

Оценка влияний факторов на систематическую погрешность идентификации

№ п/п	Наименование фактора	Натуральные уровни		Нормализованный уровень
		нижний	верхний	
1	A_1 — погрешность измерения СВР	24,0	20,0	-1
2	A_2 — погрешность передачи информации	17,5	12,5	-1
3	A_3 — погрешность идентификации СВР	20,4	26,4	$+1$

$$C_{MO} = b_0 + \gamma \sum b_i A_i + \gamma \sum_{1 \le i < j \le 3} b_{ij} A_i A_j, \tag{14}$$

— среднеквадратическое отклонение

$$C_\sigma = \sqrt{b_0 + \gamma \sum_{i \le 3} b_i A_i + \gamma \sum_{1 \le i < j \le 3} b_{ij} A_i A_j}, \tag{15}$$

где γ — нормализованный уровень знака коэффициента влияющего фактора.

Для тестирования методики идентификации нестационарных объектов представлены данные швейного производства для прогнозирования почасовой производительности технологических оборудований в зависимости от материальных, энергетических и финансовых факторов. При оценке систематической погрешности модели определены также коэффициенты степени влияния и эластичности факторов.

В табл. приведены результаты формализации многофакторного анализа погрешности с учетом исследуемых условных факторов и их нормализованные уровни.

Для исследуемой предметной области оценки систематической погрешности при статистической модели идентификации и аппроксимации СВР представляются по следующим параметрам:

— математического ожидания

$$C_{MO} = 0{,}64 - 0{,}1A_1 - 0{,}13A_2 + 0{,}49A_3; \tag{16}$$

— среднеквадратического отклонения

$$C_k = t_c\sqrt{0{,}96 + 0{,}41A_3 - 0{,}006A_1A_3 - 0{,}008A_2A_3}. \tag{17}$$

Точность идентификации, аппроксимации, анализа, обработки данных определяется по упрощенным оценкам параметров:

- математического ожидания

$$C_{MO} = 0{,}64 + 0{,}49A_3; \tag{18}$$

- среднеквадратического отклонения

$$C_\sigma = \sqrt{0{,}96 + 0{,}41A_3}. \tag{19}$$

Следует отметить, что вышеизложенные методы многофакторного анализа структуры систематической погрешности идентификации и аппроксимации СВР образуют базовую основу методологии дальнейшего совершенствования и развития методов, моделей и алгоритмов анализа и обработки данных нестационарных объектов. Перспективным и эффективным направлением использования результатов проведенных исследований является решение задач повышения достоверности обработки данных путем совмещения возможностей статистических и динамических моделей идентификации с методами оценки влияющих факторов на систематическую погрешность. В связи с этим, в качестве базовой основы разработки методов повышения достоверности данных нестационарных объектов исследуем вопросы синтеза статистических и динамических моделей идентификации нестационарных объектов.

Повышение точности идентификации на основе синтеза статистических и динамических моделей. В [7] разработаны методы повышения точности идентификации нестационарных объектов на основе порогового контроля, контроля по приращениям и путем прямого и адаптивного предсказаний СВР. В предложенных методах контроля погрешности учитываются вероятности ошибок, возникающие при измерении, вводе, передаче, хранении, анализе и обработке информации, а также особенности и свойства статистических моделей идентификации. Реализованные алгоритмы уменьшают систематическую погрешность путем отождествления искаженной информации со значениями пороговых границ контроля, средними значениями, предыдущим отсчетом и предсказанным значением СВР. Для полученных оптимальных границ контроля по выражениям оценки минимальной среднеквадратической погрешности обработки данных доказана эффективность алгоритмов по критериям вероятности необнаруженных ошибок, минимальной среднеквадратической погрешности, трудоемкости и стоимости контроля. Определено, что в результате применения методов точность (достоверность) обработки данных повышается на два порядка по сравнению с известными статистическими методами точечного и интервального контроля, значительно снижается дисперсия исходного идентифицируемого процесса, устраняются отрицательно влияющие нестационарные составляющие и случайные всплески в динамике СВР.

В настоящем исследовании разработаны методы повышения точности идентификации нестационарных объектов на основе синтеза динамических моделей с отмеченными алгоритмами контроля достоверности информации, что позволит повысить достоверность информации за счет дополнительных знаний об условиях нестационарности процесса и параметрической неопределенности моделей идентификации.

Одним из подходов к уменьшению систематической погрешности идентификации СВР является использование методов проверки соблюдения балансовых соотношений, введенных в структуру динамической модели нестационарного объекта. Для оптимизации анализа и обработки данных в структуру динамической модели идентификации нестационарного объекта вводятся дополнительные балансовые соотношения, которые устанавливаются исходя из нормативных требований, выявленных за достаточно длительный интервал времени.

Определено, что многоитеративным процессам обработки данных свойственны потеря устойчивости динамической модели нестационарного объекта и, как следствие, нарушение условия балансовых соотношений влияющих факторов и результативных.

Для проверки соблюдения балансовых соотношений в динамических моделях идентификации нестационарного объекта предлагается метод целевого сдвига значений последовательности СВР внутри диапазона вероятностного распределения с учетом итеративности, условий нестационарности и параметрической неопределенности.

Опишем математическую модель идентификации нестационарного объекта с процедурами проверки соблюдения балансовых соотношений.

Пусть задано условие

$$F_j(X) = F_j(x_1,...,x_i,...,x_n) = 0, \tag{20}$$

где $X = (x_1,...,x_i,...,x_n)$ — вектор размерности n, состоящий из значений факторов A_i, влияющих на систематическую погрешность СВР, $j \in \overline{1,m}$, $i \in \overline{1,n}$, m — число уравнений балансов.

Считается, что балансовое соотношение в уравнении (21) соблюдается при подстановке в него оценок значений СВР в виде

$$x_i^c = x_i - |\Delta x_i|\,, \tag{21}$$

где x_i^c — значение i-го идентифицированного фактора; Δx_i — систематическая погрешность идентификации.

Система уравнений (20) с учетом (21) запишется в виде

$$F_j(\overline{\mathbf{X}^c}) = F_j(x_1^c,...,x_i^c,...,x_n^c) = \Delta F_j, j \in \overline{1,m}, i \in \overline{1,n}. \tag{22}$$

где $\overline{\mathbf{X}^c}\,\{x_1^c,...,x_i^c,...,x_n^c\}$ — вектор значений i-го идентифицированного фактора; ΔF_j — небалансы, эквивалентные систематической погрешности.

Для нахождения общего решения задачи будем считать, что уравнения (20) являются непрерывными и дифференцируемыми по всем переменным. Тогда их можно разложить в ряд Тейлора по степеням приращений Δx_i

$$F_j(\overline{X}) = F_j(\overline{X^c}) + \sum_{i=1}^{n} \frac{\partial F_j(\overline{X^c})}{\partial x_i^c} \cdot \Delta x_i + ... = 0. \tag{23}$$

Принимается допущение, что значение параметра Δx_i по сравнению с истинными значениями обработанных величин являются очень малыми.

В связи с этим в (23) не принимаются во внимание нелинейные члены ряда Тейлора, содержащие в качестве сомножителей величины высших порядков $(\Delta x_i)^k$, где $k = 2,3,...$.

С учетом выражений (22) и (23) получим упрощенное выражение для проверки балансовых соотношений:

$$\Delta F_j + \sum_{i=1}^{n} a_{j,i} \cdot \Delta x_i = 0, \tag{24}$$

где $a_{j,i} = \frac{\partial F_j(\overline{\mathbf{X}^c})}{\partial x_i^c}$ — частные производные системы уравнений (22).

Отметим, что задача относится к классу задач оптимального распределения при наличии ограничений и может быть решена методом неопределенных множителей Лагранжа.

Решение задачи методом неопределенных множителей Лагранжа. Пусть значение параметра Δx_i для проверки и корректировки балансовых соотношений находится по методике оптимального распределения корректирующих поправок.

В систему уравнений балансовых соотношений предъявляются ограничения в виде равенств (23) размерности m.

В соответствии с этим критерий оптимизации Лагранжа запишется как

$$L = \sum_{i=1}^{n} p_i \cdot \left(\frac{\Delta x_i}{x_i^{\max}} \right)^2 \to \min, \tag{25}$$

где $x_i^{\max}$ — верхний предел значений диапазона вероятностного распределения i-го идентифицированного фактора;

p_i — весовые коэффициенты влияния i-го идентифицированного фактора x_i на систематическую погрешность.

Весовые коэффициенты p_i рассчитываются по следующему соотношению:

$$p_i = \frac{1}{\delta_{x_i}^2} / \sum_{i=1}^{n} \frac{1}{\delta_{x_i}^2}, \tag{26}$$

где δ_{x_i} — допускаемый предел к границам влияния погрешности структурных факторов.

С учетом (26) функция оптимизации Лагранжа запишется в виде

$$L = \sum_{i=1}^{n} p_i \cdot (\frac{\Delta x_i}{x_i^c})^2 + \sum_{j=1}^{m} (\sum_{i=1}^{n} a_{j,i} \cdot \Delta x_i + \Delta F_j) \cdot \lambda_j \to \min, \tag{27}$$

где λ_j — множитель Лагранжа, который необходимо определить.

Для получения значения параметра Δx_i продифференцируем функцию Лагранжа по переменным Δx_i и λ_j и получим $(n+m)$ систему нелинейных уравнений:

$$2 \cdot \frac{p_i}{(x_i^c)^2} \cdot \Delta x_i + \sum_{j=1}^{m} a_{j,i} \cdot \lambda_j = 0, \sum_{i=1}^{n} a_{j,i} \cdot \Delta x_i + \Delta F_j = 0, j \in \overline{1,m}, i \in \overline{1,n}. \tag{28}$$

Величина Δx_i получается из первых n уравнений

$$\Delta x_i = -\frac{(x_i^c)^2}{2 \cdot p_i} \cdot \sum_{j=1}^{m} a_{j,i} \cdot \lambda_j. \tag{29}$$

А из оставшихся m уравнений получаются значения параметра λ_j

$$b_{j,l} \cdot \lambda_l = -\Delta F_j, \tag{30}$$

где $b_{j,l} = \sum_{i=1}^{n} a_{j,i} \cdot a_{l,i} \cdot \frac{(x_i^c)^2}{2 \cdot p_i}$, $j \in \overline{1,m}$, $l \in \overline{1,m}$.

Задачей алгоритма является сведение к нулю невязок ΔF_j, которые обусловлены нелинейностью уравнений. Для этого предложен многошаговый алгоритм проверки соблюдения балансовых соотношений с последовательным уточнением на каждом шаге итерации частных производных $a_{j,i}$.

Многошаговый алгоритм контроля систематической погрешности. Вычисляются величины $a_{j,i}^k$, Δx_i^k для k-го шага итерации, и затем корректируются их значения для следующего шага алгоритма.

Итерации алгоритма продолжаются до выполнения условия

$$\left|\Delta F_j^k\right| \le \varepsilon, \tag{31}$$

где ε — точность идентификации нестационарного объекта, задаваемая балансовыми уравнениями.

На каждом шаге итерации алгоритма определяется разность скорректированного и истинного значений переменной факторного влияния Δx_i в систематическую погрешность. При этом наличие большого значения Δx_i с постоянным знаком означает значимость влияния фактора.

Если значение систематической погрешности соответствует условию $|\Delta x_i| \le \Delta x_i^*$, то значения Δx_i следует корректировать в предположении, что Δx_i^* — допустимое значение систематической погрешности.

Систематическая погрешность, образованная влияниями всех структурных факторов, рассчитывается как

$$\Delta x_i^{\text{А}}(N) = \frac{1}{N} \cdot \sum_{j=1}^{N} \Delta x_i(j \cdot t), \tag{32}$$

где t — период выполнения системы анализа и обработки данных;

$\Delta x_i^{\text{А}}(N)$ — усредненное значение влияния факторов при N числе итераций алгоритма.

Среднеквадратическое отклонение ошибок алгоритма контроля погрешности идентификации оценивается по выражению

$$\sigma_i(N) = \sqrt{\frac{\sum_{j=1}^{N} \left(\Delta x_i(j \cdot t) - \Delta x_i^{\text{А}}(N)\right)^2}{N-1}}. \tag{33}$$

Отметим, что эффективность методов проверки балансовых соотношений по критерию (33) зависит от характера модели идентификации.

Ниже предлагается алгоритм идентификации нестационарных объектов путем корректировки балансовых соотношений с линейными динамическими уравнениями.

Алгоритм проверки балансовых соотношений путем корректировки параметров линейных динамических моделей. Для алгоритма формируются массивы следующих параметров: x_i^c — усредненное значение влияния факторов; ΔF_j^* — допустимая погрешность идентификации по нелинейной динамической модели нестационарного объекта; Δx_i^* — допустимое значение систематической погрешности системы; p_i — весовые коэффициенты i-го фактора x_i влияния на систематическую погрешность.

Алгоритм включает следующие шаги.

Шаг 1. Ввод и инициализация входных переменных, $k = \overline{0,m}$.

Шаг 2. Проверка небалансов ΔF_j в системе уравнений (22).

Шаг 3. Проверка условия

$$|\Delta F_j| \le \Delta F_j^* \tag{34}$$

и сравнение расчетных и допустимых значений ΔF_j.

Шаг 3.1. Если выполняется условие (34), то результаты расчета признаются достоверными. Остановка алгоритма и перевод инициализированной переменной на первоначальное состояние $k = 0$.

Шаг 3.2. Если наблюдается нарушение условия (34) хотя бы для одного уравнения, то запускаются процедуры настройки и соответствующих корректировок параметров. Инициализированной переменной k присваивается значение 1.

Шаг 3.3. Цикл проверки небалансов ΔF_j и расчет оценок параметра Δx_i.

Шаг 4. Вычисление коэффициентов $a_{j,i}$ для линеаризованных уравнений.

Шаг 5. Формирование матрицы коэффициентов $a_{j,i}$ системы линейных уравнений (22) и матрицы-столбца небалансов ΔF_j.

Шаг 5.1. Если расчет выполняется для случая $n = m$, то алгоритм прекращает работу.

Шаг 5.2. Если $n > m$, то в матрицу включаются дополнительные строки.

Шаг 6. Задается алгоритм численного решения системы уравнений (24) или (28).

Шаг 6.1. Получение оценки систематической погрешности для всех уровней факторных влияний.

Шаг 7. Коррекция значения $\left|\Delta F_j^k\right|$ по формуле (31).

Шаг 8. Вывод результатов расчета и прекращение работы алгоритма.

Доказано, что реализация методов синтеза моделей многофакторных влияний систематической погрешности с динамической моделью нестационарного объекта методами повышения достоверности информации путем порогового контроля, контроля по приращениям, адаптивными моделями предсказаний, моделями и алгоритмами проверки и коррекции балансовых соотношений способствуют достижению повышения устойчивости выполнения программно-алгоритмических комплексов идентификации, аппроксимации, анализа и обработки данных. Кроме того, изложенная методика линейной динамической модели идентификации исследована при различных законах распределения систематической погрешности и свойствах нестационарности СВР. Определены состав специфических характеристик входных переменных, моделей и алгоритмов контроля систематической погрешности, настройки и коррекции параметров модели, а также исследованы полученные оценки минимизации дисперсии СВР и выходных параметров.

Ниже возможности предложенного подхода расширяются введением блока нелинейности, в котором эффективность механизма идентификации нестационарных объектов в условиях параметрической неопределенности моделей достигается за счет синтеза нечетких уравнений с алгоритмами логических выводов.

Оптимизация идентификации нестационарного объекта на основе синтеза нечетких уравнений с алгоритмами логических выводов. В качестве удобного и эффективного инструмента оптимизации достоверности обработки данных могут выступать гибридные интеллектуальные модели идентификации нестационарных объектов, в которых синтезируется математический аппарат нечетких уравнений, логических выводов, когнитивных карт, нечетких когнитивных карт [8]. При этом функция погрешности идентификации с учетом структурного влияния факторов представляется в виде матричного уравнения

$$\tilde{P}(A_i) = P_i(A_i) \cdot w_{ij}. \tag{35}$$

Требуется, чтобы

$$\tilde{P}(A_i) = \min(P_i(A_i) \cdot w_{ij}),$$

где A_i — факторы влияния структурных составляющих систематической погрешности; w_{ij} — весовой коэффициент степени влияния.

Синтезированные модели, методы и алгоритмы реализованы при допущении, что заданная матрица состоит из набора строк, каждая из которых характеризует активность факторов в t, $t+1$, $t+2$ моментах времени. Разность значений факторов на t-м и $(t+1)$-м шагах итерации алгоритма принимается в качестве составляющей систематической погрешности идентификации. Процедуры проверки, настройки и корректировки параметров модели идентификации выполняются аналогично работе предыдущего алгоритма с линейными уравнениями.

Алгоритм генерирует вектор приращений $\Delta \mathrm{x}_i$ для получения оценок $A_i(t)$ – значения активности влияния i-го фактора погрешности в момент времени t, а также $A_i(t+1)$, $A_i(t+2)$ — значения активности влияний i-го фактора в t, $t+1$, $t+2$ моменты времени.

Исходный вектор приращений в момент времени t задается как

$$\Delta \mathrm{x}_i = \frac{A_i(t+1) - A_i(t)}{A_i(t)}.$$

Результирующий вектор приращений в $t+1$ моменты времени определяется как

$$\mathrm{y}_i = \frac{A_i(t+2) - A_i(t)}{A_i(t)}.$$

Эффективность модели идентификации оценивается по величине

$$E(W) = \frac{1}{2} \sum_{t=1}^{K} \sum_{i=1}^{N} (o(i) - y(i))^2 \to \min, \tag{36}$$

где $o_i(t)$ — приращение i-го фактора, полученное на обучающем векторе x(t).

Качество расчетов на $(i+2)$ шаге итерации алгоритма подтверждается расчетами по формуле (35), значения которых должны совпадать со значением, полученным в предыдущем шаге итерации. В отличие от предыдущих подходов, реализованная гибридная модель идентификации отражает картину влияния факторов на систематическую погрешность на всех уровнях информационных процессов, а корректировка параметров модели производится путем внесения изменений в когнитивные карты, формируемые из исходного вектора приращений. Работа процедур проверки и корректировки параметров аналогична работе процедур алгоритма с линейными уравнениями.

Заключение. Разработаны методические основы совершенствования и развития существующих методов повышения достоверности информации на основе синтеза методов статистической, динамической и нечеткой идентификации с алгоритмами порогового контроля погрешности СВР, контроля по приращениям и с предсказанием, оценки влияния факторов на систематическую погрешность, настройки параметров моделей для оптимизации обработки данных нестационарных объектов. Предложены методы многофакторного анализа для определения степени влияния составляющих структурной погрешности, определяемых на различных уровнях информационного процесса для обеспечения эффективности идентификации и аппроксимации СВР.

Полученные методики на основе гибридной идентификации рекомендованы для реализации в виде программно-алгоритмических комплексов, в которых учитываются вид и свойства законов распределений, нелинейность влияющих факторов, неопределенность в параметрах моделей и эффективно проверяется соблюдение балансовых соотношений в динамических моделях для повышения достоверности информации.

Полученные результаты исследований подтверждают перспективность дальнейшего совершенствования и развития гибридных моделей идентификации нестационарного объекта с целью уменьшения систематической погрешности и повышения точности анализа и обработки данных на основе математического аппарата мягких вычислений, а именно синтеза моделей нечеткой логики, алгоритмов обучения нейро-нечетких сетей, методов поиска по матрице нечеткой когнитивной карты и генетических алгоритмов.

Список литературы

1. Зарипова Г. И. Адаптивный контроль достоверности технологических параметров на основе моделей нечеткого вывода // Материалы VIII международной научно-практической конференции „Научная дискуссия: вопросы технических наук“, 4 апреля 2013 г. Москва: „Международный центр науки и образования“, 2013. С. 31–37.

2. Мирошник И. В., Никифоров В. О., Фрадков А. Л. Нелинейное и адаптивное управление сложными динамическими системами. СПб.: Наука, 2000.

3. Миф Н. П. Модели и оценки погрешности технических измерений. М.: Стандарты, 1976.

4. Игамбердиев Х. З., Севинов Ж. У., Зарипов О. О. Регулярные методы и алгоритмы синтеза адаптивных систем управления с настраиваемыми моделями. Т.: ТашГТУ, 2014.

5. Карабутов Н. Н. Адаптивная идентификация систем. Информационный синтез. М.: Ком Книга, 2006.

6. Zaripova G. I., Akhatov A. R. Methods and Algorithms to Control Information Authenticity during Transfer and Handling of Data of Continuous Objects on the basis of a Neuro-Fuzzy Network // 2013 International Conference in Central Asia on Internet (ICI), Tashkent, 8–10 October 2013, Section 7, IEEE. Tashkent, 2013. P. 12–18.

7. Жуманов И. И., Абдуллаев А. Н. Контроль точности передачи информации в системах автоматизации измерения и обработки данных нестационарной природы // „Интеллектуальные системы для индустриальной автоматизации“ WCIS — 2006, ТГТУ, Ташкент. С. 213–218.

8. Ярушкина Н. Г. Основы нечетких и гибридных систем: Учебное пособие. М.: Финансы и статистика. 2004.

Зарипова Гульсара Исраиловна — соискатель кафедры информационных технологий Самаркандского государственного университета тел.: +998662293558; e-mail: zaripovagi@mail.ru.

Гульсара Исраиловна Зарипова окончила механико-математический факультет Самаркандского государственного университета по специальности „Математика“ в 1996 году. С 1996 года работала преподавателем математики и информатики в средней школе № 21 Пастдаргомского района Самаркандской области. В 2004–2006 гг. работала заведующей кафедрой математики Педагогического колледжа № 1 г. Самарканда. С 2006 года является заместителем директора этого колледжа. С 2008 года ведет научную работу по подготовке диссертации на кафедре информационных технологий Самаркандского государственного университета в качестве научного исследователя-соискателя. Г. И. Зариповой опубликовано 57 научных и 4 научно-методических работы в области информационных технологий, интеллектуального анализа данных, повышения достоверности передачи и обработки информации. Ее текущие исследовательские интересы включают технологии обработки информации на базе математического аппарата мягких вычислений, в частности, нейронных сетей, нечетких множеств и нечеткой логики.

Gulsara Israilovna Zaripova received her M. S. degree in Mathematics from the Samarkand State University (1997). From 1996 she worked the teacher of mathematics and computer science in secondary school № 21 of Pastdargom area in Samarkand region. In 2004–2006 she worked the manager by faculty of mathematics at the Pedagogical college № 1 in Samarkand. Since 2006 she is the deputy director of this college. In 2008 she starts scientific issue as the scientific researcher on Department of Information Technologies of the Samarkand State University. G. I. Zaripova published over 57 papers in the areas of information technologies, Data Mining, improving reliability of data transfer and processing. Her current research interests include technologies of data processing on the basis of mathematical apparatus of soft calculating, in particular, neural networks, fuzzy set and fuzzy conclusions.

Дата поступления — 29.10.2015