УДК 621.396

МОХАММАД РЕДА АБД АЛЬ-ЛАТИФ АЛЬ-ХИННАВИ, аспирант, НТУ «ХПИ» (г. Харьков)

АНАЛИЗ МЕТОДОВ ПОСТРОЕНИЯ МОДЕЛИ РЕГРЕССИИ ПОКАЗАТЕЛЯ КАЧЕСТВА ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ ПЕРЕДАЧИ ДАННЫХ

В статті розглянуті питання забезпечення безупинної роботи систем передачі даних за рахунок підвищення точності прогнозування моменту відмови по результатах вимірів вектору контрольованих параметрів. Проаналізовано методи побудови моделі регресії показника якості функціонування системи, що забезпечує мінімальну дисперсію помилки визначення моменту відмови.

Problems of providing of uninterrupted functioning of data transmission systems due to increasing accuracy in forecasting of failure moment according to measuring results of the vector of controlled parameters are considered in the article. Methods of model construction of regression quality factor of system functioning providing minimum error dispersion in definition failure moment are analysed.

Постановка проблемы. В настоящее время по-прежнему актуальной является задача обеспечения бесперебойной работы систем передачи данных [1]. Поскольку в силу объективных причин, исключить отказы каналов связи невозможно, прибегают к их резервированию. Однако, полное решение данной задачи возможно лишь при введении в систему передачи данных дополнительных контролирующих устройств и организации гибкой системы технического обслуживания, учитывающей фактическое состояние обслуживаемых средств. Эффективность указанных мер зависит с одной стороны от глубины и достоверности контроля, и с другой стороны – от точности прогнозирования момента наступления отказа.

Анализ литературы. Вопросы оптимизации контроля технического состояния и прогнозирования работоспособности технических систем широко освещены в [2 –5].

Математическая модель любой системы как объекта контроля является зависимостью показателя качества функционирования от параметров при фиксированных входных сигналах системы: $G = f(\theta_1, \theta_2, \ldots, \theta_m)$. Для системы передачи сигналов [6] в качестве комплексного показателя используют коэффициент взаимной корреляции контролируемого и эталонного сигналов

$$G = \left| \frac{1}{2\sqrt{P_k P_э} T} \int_0^T Z_k(t) Z_э^*(t) dt \right|^2, \; t \in [0, T], \qquad (1)$$

где $Z_k(t)$, $Z_э(t)$ – комплексные формы представления соответственно контролируемого и эталонного (с параметрами, равными номинальным значениям) сигналов; $Z_э^*(t)$ – функция, комплексно сопряженная с $Z_э(t)$; $P_k, P_э$ – мощности соответственно контролируемого и эталонного сигналов; T – длительность элемента сигнала.

Выбор комплексного показателя системы передачи сигналов в виде (1) дает возможность установления с его помощью функциональной зависимости между параметрами θ, характеризующими техническое состояние системы, и вероятностью P ошибочного приема элемента сигнала, определяющей степень соответствия системы целевому предназначению

$$P = \varphi(\theta, W, g, H), \tag{2}$$

где W – величина, определяющая статистические свойства канала; g – параметр, характеризующий влияние различного рода помех; H – энергетические параметры сигнала.

Задача оптимизации контроля фактического состояния системы сводиться к выбору номенклатуры контролируемых параметров $\theta = \{\theta_i\}_{i=\overline{1,m}}$, установлению эксплуатационных допусков $\Delta\theta = \{\Delta\theta_i\}_{i=\overline{1,m}}$ и формулированию правила принятия решения по результатам контроля, позволяющего максимизировать достоверность контроля D. Таким образом, задача сводится к построению целевой функции вида

$$\max_{G_Д \in G(\theta)} D(G_Д) = \min_{\Delta\theta}\{\alpha(\Delta\theta), \beta(\Delta\theta)\}, \tag{3}$$

где достоверность контроля D является функцией от допусков на показатель качества системы $G_Д$ [7] и определяется как $D = 1 - \alpha - \beta$. Здесь α и β – ошибки соответственно первого и второго рода, обусловленные неправильным выбором допусков на параметры $\Delta\theta$.

Правило принятия решения о возможности дальнейшей эксплуатации системы по результатам контроля имеет вид

$$\rho(G) = \begin{cases} 1, \text{ если } G > G_Д; \\ 0, \text{ если } G \le G_Д, \end{cases} \tag{4}$$

а при наличии модели (2) –

$$\rho(P)=\begin{cases}1, \text{ если } P<P_Д;\\ 0, \text{ если } P\geq P_Д.\end{cases} \tag{5}$$

Постановка задачи прогнозирующего контроля системы сводится к стандартной задаче экстраполяции векторного случайного процесса по результатам наблюдения на некотором интервале времени [4]. В качестве векторного случайного процесса можно рассматривать обобщенный P или комплексный G показатели качества. В общем виде задача прогнозирования формулируется следующим образом. Пусть работоспособность системы передачи сигналов определяется значениями ее параметров $\{\theta_i(t_i)\}$, $i=\overline{1,m}$, $l=\overline{1,n}$, полученными в моменты времени $t_l\in[0,\tau']$. Требуется найти такой момент времени $t_k\in[\tau',\tau'']$, при котором показатель системы P (или G) первый раз выйдет за пределы допусковой области, т.е. $P(t_k)\geq P_Д$ (или $G(t_k)\leq G_Д$). Иными словами необходимо предсказать наиболее вероятный момент отказа системы (при этом полагается, что понятие отказа системы передачи сигналов определено значениями показателей качества $P_Д$ или $G_Д$).

Математическая постановка задачи прогнозирования требует точного знания статистических законов распределения $W_1(G,t_l)$, $W_2(G,t_k)$, $W_3(G,t_l,t_k)$, $l=\overline{1,n}; k=\overline{n+1,d}$. Только при этих условиях можно решать задачу прогнозирования в пределах множества моментов времени $t\in[0,\tau'']$, на котором определены случайные функции $W_1(t), W_2(t), W_3(t)$.

Рассмотренные подходы к оптимизации контроля технического состояния системы и прогнозирования ее работоспособности имеют существенный недостаток – необходимость точного знания условий эксплуатации в моменты времени проведения контроля. Ограниченность выборочных данных о результатах контроля приводит к большим ошибкам определения момента наступления отказа.

Целью статьи является выбор структуры модели системы передачи данных, обеспечивающей минимальную дисперсию ошибки прогнозирования момента наступления ее отказа.

Анализ методов построения модели регрессии. Рассмотрим выборку $W=\left(\theta_i^{(1)},\theta_i^{(2)},\ldots,\theta_i^{(m)};G_i\right)_{i=1,2,\ldots,n}$ наблюдений вектора контролируемых параметров $\Theta=\left(\theta^{(1)},\theta^{(2)},\ldots,\theta^{(m)}\right)^T$ и комплексного показателя качества G в виде $n\times m$ матрицы X и $n\times 1$ вектора y. Матрица X считается детерминированной полного ранга $rank\ X=m$, а вектор y содержит шум:

$$y=y_o+\xi,\quad y_o=Ey=X\Lambda_o,\quad E\xi=0,\quad E\xi\xi^T=\sigma^2 I_n, \tag{6}$$

где y_o – точное (незашумленное) значение комплексного показателя качества G; Λ_o – точный (неизвестный) вектор параметров модели объекта контроля, связывающий между собой значения вектора контролируемых параметров и комплексный показатель качества; ξ – вектор шума с независимыми, одинаково распределенными компонентами; E – символ математического ожидания по всем возможным реализациям вектора шума; I_n – единичная $n\times n$ матрица; σ^2 – неизвестная конечная дисперсия шума.

Пусть исследуемый вектор переменных $(\Theta;G)^T$ подчиняется $(m+1)$-мерному нормальному распределению. Тогда, согласно [8], неизвестная истинная модель $f(X;\Lambda_o)$ (функция регрессии y_o по X) принадлежит к классу моделей, линейных по параметрам. Задача состоит в выборе такой линейной функции $\hat{f}_a(\Theta;\Lambda)=\lambda^{(0)}+\sum_{k=1}^{m}\lambda^{(k)}\theta^{(k)}$, которая бы аппроксимировала данную зависимость с минимальной дисперсией ошибки.

Подходы к решению данной задачи описаны в [8]. В частности, предлагается процедура определения оптимального состава и числа предикторов θ, основанная на поиске максимума нижней доверительной границы $\left(R^2\right)_P^{\min}$ для истинного значения коэффициента детерминации R^2 (при заданной доверительной вероятности P):

$$\left(R^2\right)_P^{\min}\approx\hat{R}^2(k)-\lambda(P)\cdot\frac{2\cdot m(n-m-1)}{(n-1)(n^2-1)}\left(1-\hat{R}^2(k)\right), \tag{7}$$

где

$$\hat{R}^2(k)=\max_{i_1,i_2,\dots,i_k}\hat{R}^2_{G\cdot\theta^{(i_1)}\theta^{(i_2)}\dots\theta^{(i_k)}},$$

$R_{G\cdot\theta^{(i_1)}\theta^{(i_2)}\dots\theta^{(i_k)}}$ – множественный коэффициент корреляции; $\lambda(P)$ – коэффициент пропорциональности, зависящий от заданной величины доверительной вероятности P.

Ошибка в предсказании G в этом случае определяется выражением

$$\sigma_G^2=\sigma^2\left(1-\left(R^2\right)_P^{\min}\right). \tag{8}$$

В [8] описан также метод структурной минимизации теоретического критерия адекватности модели $f_a(\Theta;\Lambda)$

$$\Delta(f_a) = E\rho\left(\hat{\varepsilon}_{f_a}(\Theta)\right), \tag{9}$$

где $\rho\left(\hat{\varepsilon}_{f_a}(\Theta;\Lambda)\right)$ – функция потерь, измеряющая убытки от неточности восстановления $G(\Theta;\Lambda)$ с помощью функции $f_a(\Theta;\Lambda)$; $\hat{\varepsilon}_{f_a}(\Theta;\Lambda) = G(\Theta;\Lambda) - f_a(\Theta;\Lambda)$; $E(\cdot)$ – операция усреднения по всем возможным значениям случайной величины $\hat{\varepsilon}_{f_a}(\Theta;\Lambda)$ (при каждом фиксированном Θ) и по всем возможным значениям Θ.

Показано, что для ограниченного объема исходных данных n он позволяет установить компромисс между сложностью выбираемой модели регрессии и качеством приближения к выборочным данным (величиной $\hat{\Delta}_n(f_a)$)

$$\hat{\Delta}_n(f_a) = \frac{1}{n}\sum_{i=1}^{n}\rho\left(\hat{\varepsilon}_{f_a}(\Theta_i;\Lambda)\right), \tag{10}$$

при котором достигается наименьшая гарантированная оценка теоретического критерия адекватности.

Принципиально новый подход к решению задачи выбора аппроксимирующей функции $f_a(\Theta;\Lambda)$, использование метода группового учета аргументов (МГУА), изложенного в [9]. Полученные по этому методу модели оптимальной сложности отображают неизвестные закономерности функционирования исследуемого объекта (процесса), информация о которых неявно содержится в выборке данных. В МГУА для построения моделей применяются принципы автоматической генерации вариантов, неокончательных решений и последовательной селекции лучших моделей по внешним критериям. Такие критерии основаны на делении выборки на части, при этом оценивание параметров и проверка качества моделей выполняются на разных подвыборках. Это позволяет избежать обременительных априорных предположений, поскольку деление выборки дает возможность автоматически (неявно) учесть разные виды априорной неопределенности.

Описанный в [9] метод критических дисперсий, являющийся аналитическим аппаратом теории МГУА, позволяет детально исследовать закономерности изменения сложности оптимальных структур в зависимости от уровня шума и других показателей неполноты априорной информации. При этом метод гарантирует минимальную дисперсию ошибки прогнозирования

$$J(s) = E\left\|y_0 - \hat{y}_s\right\|^2 = E\left\|y_0 - X_s\hat{\Lambda}_s\right\|^2, \tag{11}$$

где s – число оцениваемых параметров; X_s – подматрица из s произвольных (например, первых) столбцов матрицы X; $\hat{\Lambda}_s$ – соответствующая МНК-оценка по выборке W.

Указанные достоинства метода критических дисперсий позволяют надеяться на получение в рамках теории МГУА оценки момента наступления отказа системы передачи данных с минимальной дисперсией ошибки.

Выводы. Обеспечение бесперебойной работы системы передачи данных требует повышения точности прогнозирования момента наступления ее отказа по результатам измерения вектора контролируемых параметров. Известные методы решения задачи прогноза опираются на точное знание статистических законов распределения показателя качества функционирования системы. Ограниченность выборочных данных о результатах контроля приводит к возникновению больших ошибок в определении момента отказа, что снижает эффективность мероприятий, призванных обеспечить бесперебойную работу системы. В статье проанализированы методы построения модели регрессионной зависимости между показателем качества функционирования системы и вектором контролируемых параметров, которая бы обеспечивала минимальную дисперсию ошибки в определении момента отказа. Отмечено, что решение данной задачи наиболее целесообразно проводить в рамках теории МГУА. Аналитическим аппаратом этой теории является метод критических дисперсий, который позволяет детально исследовать закономерности изменения сложности модели в зависимости от уровня шума и других показателей неполноты априорной информации.

Список литературы: **1.** *Мохаммад Реда Абд Аль-Латиф Аль-Хиннави.* Обеспечение бесперебойной работы систем передачи данных. Вестник Национального технического университета «Харьковский политехнический институт». Сборник научных работ. Тематический выпуск: Информатика и моделирование. – Харьков: НТУ «ХПИ», 2003. – № 19. – С. 107 – 110. **2.** *Дружинин Г.В.* Надежность автоматизированных систем. – М.: Энергия, 1977. – 536 с. **3.** *Волков Л.И.* Управление эксплуатацией летательных комплексов. – М.: Высшая школа, 1981. – 368 с. **4.** *Гаскаров Д.В., Голинкевич Т.А., Мозгалевский А.В.* Прогнозирование технического состояния и надежности радиоэлектронной аппаратуры. – М.: Сов. радио, 1974. – 224 с. **5.** Промышленные комплексы моделирования процессов эксплуатации сложных технических систем / Под ред. *В.В. Литвинова, Т.П. Марьяновича*. – Киев: Наук. думка, 1994. – 242 с. **6.** *Федоренко В.В.* Математическая модель системы передачи сигналов для решения задач контроля // Электрон. моделирование. – 1991. –13. – № 6. – С. 85 – 88. **7.** *Федоренко В.В.* Модель оптимизационных задач технического обслуживания систем передачи сигналов по фактическому состоянию // Электрон. моделирование. – 1994. – 16. – № 1. – С. 47 – 51. **8.** *Айвазян С.А., Енюко И.С., Мешалкин Л.Д.* Прикладная статистика: Исследование зависимостей: Справ. изд. / Под ред. *С.А. Айвазяна.* – М.: Финансы и статистика, 1985. – 487 с. **9.** *Степашко В.С.* Теоретические аспекты МГУА как метода индуктивного моделирования // УСиМ. –2003. – № 2. – С. 39 – 44.

Поступила в редакцию 14.04.04