УДК 621.391

Д.М. Ненадович

АЛГОРИТМ АДАПТИВНОГО ПРОГНОЗИРОВАНИЯ ЗНАЧЕНИЙ ВРЕМЕННЫХ РЯДОВ В ЗАДАЧАХ ЭКСПЕРТНОГО ОЦЕНИВАНИЯ КАЧЕСТВА ПРОЕКТНО-ТЕХНИЧЕСКИХ РЕШЕНИЙ, ПРЕДЛАГАЕМЫХ К РЕАЛИЗАЦИИ В ХОДЕ ПРОЕКТИРОВАНИЯ СЛОЖНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

При организации экспертной деятельности, большой интерес вызывает подход к формированию прогнозных значений временного ряда значений экспертных показателей качества (ЭПК) проектируемой сложной информационной системы (СИС). Так, например, для эвентуальной модели временного ряда прогнозные значения ЭПК на l - шагов вперед могут быть получены при выполнении условий

$q \geq p + d$, $l > q - p - d$ [1].

Общее выражение для модели АРПСС, временного ряда (m-го индикатора значения элемента векторного ЭПК, состояния модели системы, (θ_m) и т.д.), может быть представлено в следующем виде:

$$\phi(B)\Delta^d \theta_m(k) = \lambda(B)V_t + \phi(B)\Delta^d w_t; \qquad (1)$$

где $\phi(B)$ - оператор авторегрессии, определяемый в соответствии с выражением:

$$\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p; \qquad (2)$$

B - оператор сдвига назад ($B\theta_t = \theta_{t-1}$), Δ - разностный оператор со сдвигом назад ($\Delta = \theta_t - \theta_{t-1}$), w_t - значения шума наблюдения за индикаторами (погрешность измерения); V_t - значение шума возбуждения процесса θ_t , $\lambda(B)$ - оператор скользящего среднего, определяемый в соответствии с выражением:

$$\lambda(B) = 1 - \lambda_1 B - \ldots - \lambda_q B^q.$$

Очень часто для получения оценок коэффициентов авторегрессии и скользящего среднего, а также текущих значений дисперсии шума возбуждения в модели (p,q) используют раздельные субоптимальные процедуры их определения [2, 3]. Вначале оценивают коэффициенты АР на основе реализации процедуры наименьших квадратов или решения уравнения Юла-Уолкера, затем на их основе формируют модель временного ряда и получают разностный ряд между принятым и смоделированным. Последовательность остаточных ошибок (разностный ряд) используется в дальнейшем для определения коэффициентов СС.

Однако, необходимость реализации высокой степени автоматизации процесса экспертной деятельности на основе экспертных систем требует построения алгоритмов определения параметров авторегрессии в виде удобном для реализации на ЭВМ. С этой целью, воспользуемся рекуррентным алгоритмом наименьших квадратов (РНК), который позволяет производить анализ ошибок фильтрации по выборке наблюдаемых значений объемом N и при поступлении новых текущих данных $\theta(N+1)$ переходить от вектора коэффициентов линейного предсказания $\vec{\phi}_{p,N}$ к вектору $\vec{\phi}_{p,N+1}$, не решая уравнение Юла- Уолкера [3]. Для получения алгоритма РНК выражение для ошибки линейного предсказания вперед при использовании выборки размером N для k-го временного шага и глубины регрессии p запишем в следующем виде:

$$e'_{p,N}(k) = \vec{\theta}^{T}_{p,\mathrm{N}}(k)\vec{\phi}^{T}_{p,N}(k), \qquad (3)$$

где $\vec{\theta}^{\mathrm{T}}_{\mathrm{p,N}} = \left(\theta(k);\ \ \theta(k-1)\ \ \ldots\ \ \theta(k-p)\right)$ - вектор значений временного ряда размерностью p+1;

$\vec{\phi}_{\mathrm{p,N}}(r) = \left(1\ \ ;\ \vec{\phi}_{\mathrm{p,N}}(k-1)\ \ \ldots\ \ \vec{\phi}_{\mathrm{p,N}}(k-p)\right)^{T}$ - вектор значений коэффициентов авторегрессии размерностью p+1.

Так как суммирование в выражении для ошибки осуществляется с учетом отрицательного знака при коэффициентах авторегрессии, то в результате получаем разность между значением ряда в момент времени k и взвешенными значениями регрессии на прошлые значения ряда глубиной p.

Введем понятие суммы экспоненциально взвешенных квадратов ошибок предсказания на всей длине выборки $\rho'_{p,N} = \sum_{k=1}^{N} W^{N-k}(e'_{p,N}(n))^2$.

Тогда основу базового РНК-алгоритма составляют следующие выражения для векторов коэффициентов предсказания, коэффициентов усиления и дисперсии ошибки фильтрации:

$$\vec{\phi}_{p,N+1} =$$

$$= \begin{cases} \vec{\phi}_{p,N} - P_{N\,\theta_N}\vec{\theta}_{m\,p-1}(N)(\vec{\theta}^{T}_{p-1}(N)\vec{\phi}_{p,N} + \\ +\vec{\theta}_{m\,p-1}(N+1)) = \vec{\phi}_{p,N} - \\ -\vec{e}^{\,f}_{p,N}(N+1)P_{N\theta_N}\vec{\theta}_{m\ p-1}(N) = \\ = \vec{\phi}_{p,N} - \vec{e}^{\,f}_{p,N}(N+1)\vec{c}_{p-1,N}; \\ \vec{c}_{p-1,N} = P_{N-1}\vec{\theta}_{p-1}(N)/ \\ /\left(W + \vec{\theta}^{T}_{m\,p-1}(N)P_{N-1\,\theta_{N-1}}\vec{\theta}_{m\,p-1}(N)\right); \\ P_N = W^{-1}\left(\vec{I} - \vec{c}_{p-1,N}\vec{\theta}^{T}_{m\,p-1}(N)\right)P_{N-1\theta_{N-1}}, \end{cases} \qquad (4-6)$$

где $\vec{e}^{\,f}_{p,N}(N+1) = \vec{\theta}_{m\ p-1}{}^{T}(N)\vec{\phi}_{p,N} + \vec{\theta}_{m\,p-1}(N+1)$ - вектор остаточных ошибок фильтрации, т.к. в отличие от ошибки предсказания здесь используется вектор $\vec{\phi}_{p,N}$, а не $\vec{\phi}_{p,N+1}$; $\vec{c}_{p-1,N} = P_N\vec{\theta}_{m\,p-1}(N)$ - вектор коэффициентов компенсации остаточной ошибки фильтрации-экстраполяции; $\vec{P}_{N-1} = \vec{R}^{-1}_{p-1,N-1}$ - матрица дисперсий ошибок фильтрации; $R_{p-1,N-1} = \sum_{n=1}^{N-1} W^{N-1-n}\vec{\theta}_{m\,p-1}(n)\vec{\theta}^{T}_{m\,p-1}(n)$ - матрица размерностью $(p-1\times p-1)$ взвешенных с весом $0 < W < 1$ вторых моментов процесса на шагах $p-1$, усредняемых по выборке объемом $N-1$.

Практика показывает, что для формирования оценочных и прогнозных значений в общем случае нестационарных процессов изменения значений ЭПК моделей проектируемой ТКС в телекоммуникационной ЭС целесообразно реализовать (в качестве базовой с возможностью подстройки) модель АРПСС (1, 2, 4). Данная модель обладает свойствами безызбыточности и адекватности для большинства процессов, которые необходимо оценивать в ходе экспертизы различных этапов проекта ТКС.

Реккурентные соотношения для определения значений операторов скользящего среднего для модели АРПСС (1, 2, 4) могут быть представлены следующим образом:

$$\lambda_{1t} = -(c_1^{/} / \sigma_V^2 - \lambda_{1(t-1)}\lambda_{2(t-1)} - \lambda_{2(t-1)}\lambda_{3(t-1)} - \\ -\lambda_{3(t-1)}\lambda_{4(t-1)}), \quad (7)$$

$$\lambda_{2t} = -(c_2^{/} / \sigma_V^2 - \lambda_{1(t-1)}\lambda_{3(t-1)} - \lambda_{2(t-1)}\lambda_{4(t-1)}), \quad (8)$$

$$\lambda_{3t} = -(c_2^{/} / \sigma_V^2 - \lambda_{3(t-1)}\lambda_{4(t-1)} - \lambda_{2(t-1)}\lambda_{4(t-1)}), \quad (9)$$

$$\lambda_{4t} = -(c_4^{/} / \sigma_V^2 - \lambda_4), \quad (10)$$

где $c^{/}_q$ - параметры автоковариации, определяемые в соответствии с выражениями:

$$c_0 = \phi_1, \quad c_1 = \phi_1 c_0, \quad c_2 = \phi_1 c_1,$$

$$c_3 = \phi_1 c_2, \quad c_4 = \phi_1 c_3, \quad c_5 = \phi_1 c_3, \quad (11)$$

$$c_0^{/} = \phi_0^2 c_0 + \phi_1^2 c_0 + \phi_0\phi_1, \quad (12)$$

$$c_1^{/} = \phi_0^2 c_1 + \phi_1^2 c_1 + \phi_0\phi_1(c_2 + c_0), \quad (13)$$

$$c_2^{/} = \phi_0^2 c_1 + \phi_1^2 c_2 + \phi_0\phi_1(c_3 + c_1), \quad (14)$$

$$c_3^{/} = \phi_0^2 c_3 + \phi_1^2 c_3 + \phi_0\phi_1(c_4 + c_2), \quad (15)$$

$$c_4^{/} = \phi_0^2 c_4 + \phi_1^2 c_4 + \phi_0\phi_1(c_5 + c_3), \quad (16)$$

а $\sigma_v{}^2$ -остаточная дисперсия шума возбуждения, определяемая в соответствии с выражением;

$$\sigma_V^2 = c_0^{/} / 1 + \lambda_{1t}^2 + \lambda_{2t}^2 + \lambda_{3t}^2 + \lambda_{4t}^2 \quad (17)$$

Необходимо отметить, что формирование оценочного значения $\dot{\theta}_{m\,(p-1)}$ возможно только при выполнении условия:

$$\sigma_V^2 \le \sigma_{V\,ДОП}^2, \quad (18)$$

определяемого требуемой точностью прогнозирования. В случае не выполняются этого условия по завершению начального этапа оценивания, производится корректировка порядка операторов авторегрессии. Если требования к остаточной дисперсии по каким-то причинам не заданы, с достаточной для практики точностью процесс корректировки порядка операторов авторегрессии может быть приостановлен исходя из анализа соотношения:

$$(\sigma_V^2(k-1)/\sigma_V^2(k))/(\sigma_V^2(k-2)/\sigma_V^2(k-1)) \le 1{,}3, \quad (19)$$

Кроме того, при определении порядка модели важное значение имеет объем и расположение участка наблюдаемой при этом выборки. Так Ульрихом и Клейтоном эксперементально, а Лангом и Макклелланом аналитическими средствами показано, что выбор порядка модели наиболее эффективен на участке выборки составляющем от $N/2$ до $N/3$ [3].

Отдельно следует остановиться на учете шумов наблюдения для случаев оценки наблюдаемых временных последовательностей поступающих, например, из систем-аналогов ТКС. В этом случае, учитывая, что мощность шума наблюдения влияет только на член автокорреляционной последовательности, соответствующий нулевому сдвигу и дополнительную погрешность оценочного значения можно предотвратить, удаляя из оценки автокорреляционной последовательности член соответствующий нулевому сдвигу, ограничивая при этом пределы изменения остаточной дисперсии. Соответствующие компенсационные схемы представлены в [3].

Необходимо отметить, что вначале итеративной процедуры вычисления оценочных значений операторов авторегрессии они могут полагаются равными нулю, а $\phi_0 = -1$. По завершению процесса определения порядка разности определяется первоначальное

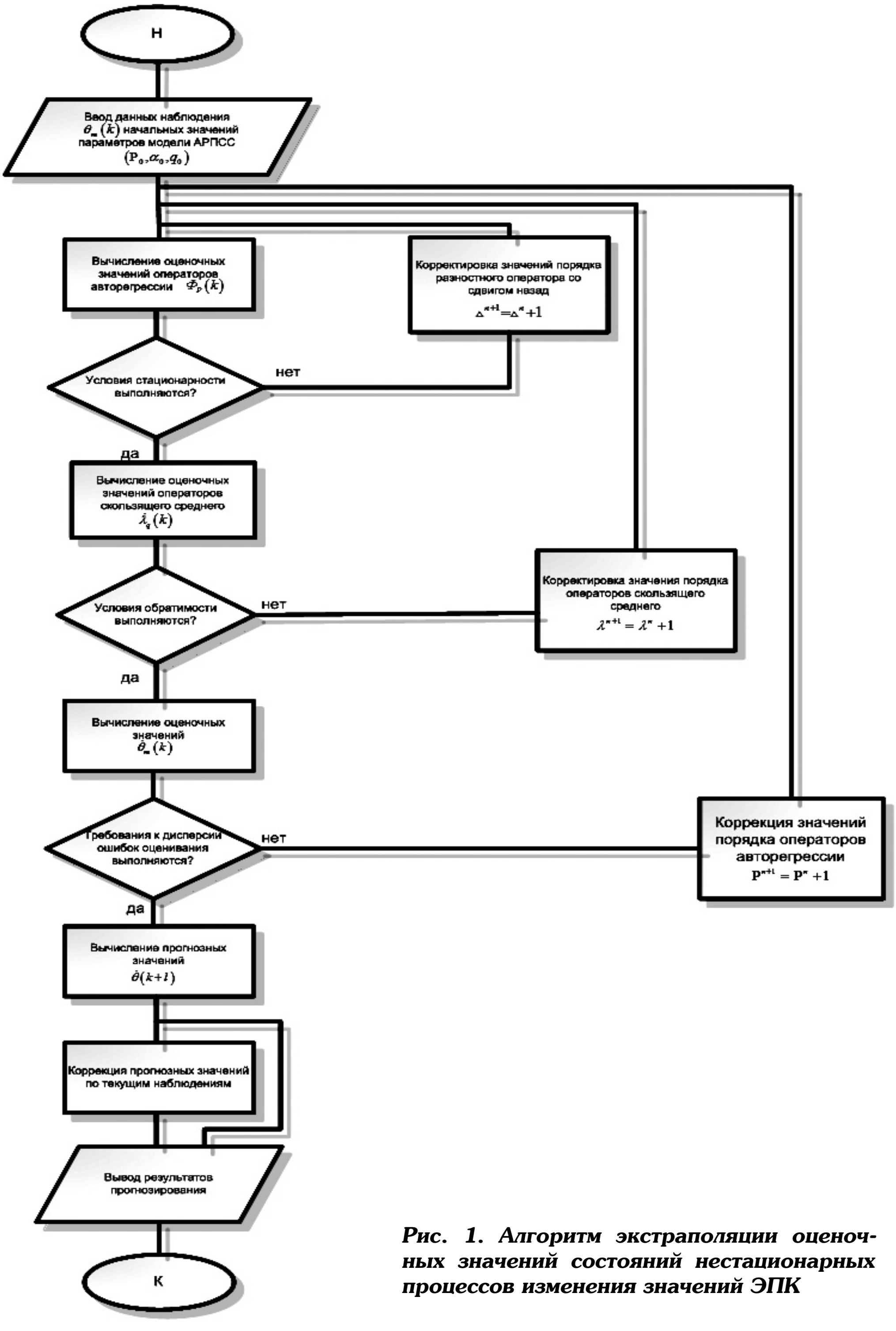

Рис. 1. Алгоритм экстраполяции оценочных значений состояний нестационарных процессов изменения значений ЭПК

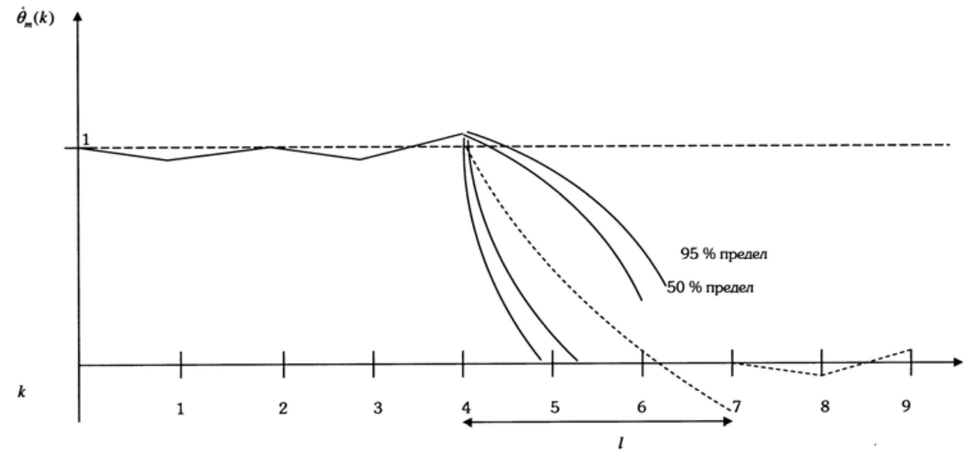

Рис. 2. Пример прогноза изменения значений индикатора:
—— наблюдаемые значения;
- - - - - - прогнозные значения

прогнозное значение индикатора значений ЭПК, моделируемой ТКС в соответствии с выражением:

$$\vec{\theta}_{m(l)} = \phi_1[\vec{\theta}_{m\,(t+l-1)}] + V_{t+l} - \lambda_1[V_{t+l-1}] - \ldots - \lambda_4[V_{t+l-3}]\,. \tag{20}$$

Подправление прогнозного значения (состояния модели ТКС, значения ПК) на $t+1$ -ом шаге осуществляется в соответствии с выражением:

$$\theta_{m\,(t+1)}(l+1) = \theta_{m\,(t)}(l) + \varphi_l V_{t+1}\,, \tag{21}$$

где φ- линейный оператор, преобразующий V_t в $\theta_{cm(t)}$, определяемый в соответствии с выражениями:

$$\varphi_1 = \phi_1 - \lambda_1 \tag{22}$$

$$\varphi_2 = \phi_1\varphi_1 + \phi_2 - \lambda_2$$

$$\varphi_l = \phi_1\varphi_{l-1} + \ldots + \phi_{p+d}\varphi_{l-p-d} - \lambda_l$$

Доверительный интервал значений прогноза определяется в соответствии с выражением:

$$z_{t+l}(\pm) = z_t(l) \pm u_{\varepsilon/2}\{1 + \sum_{j=1}^{l-1}\varphi_j^2\}^{1/2} s_v \tag{23}$$

где $u_{\varepsilon/2}$ - квантиль уровня $1-\varepsilon/2$, s_v - текущее значение σ_V.

Условия стационарности, с достаточной для практики точностью, определяются размещением значений операторов авторегресии внутри единичного круга:

$$-1 < \phi < 1\,. \tag{24}$$

В случае не выполнения условий стационарности процесса производится коррекция порядка оператора разности со сдвигом назад.

Важным этапом реализации алгоритма формирования оценочных значений операторов СС является этап проверки выполнения условий обратимости - свойства обеспечивающего «разумную» [1, 2] связь между настоящими и предшествующими значениями временного ряда.

Условия обратимости, в рассматриваемом случае, состоят в том, что корни характеристического уравнения

$$\lambda B = 1 - \lambda_1 B - \ldots - \lambda_q B^q = 0\,, \tag{25}$$

должны лежать вне единичного круга. Тогда условие, подлежащее контролю в ходе формирования прогнозных значений для рассмотренной базовой модели можно записать

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 < 1, \qquad (26)$$

$$\lambda_4 - \lambda_3 - \lambda_2 - \lambda_1 < 1, \qquad (27)$$

$$-1 < \lambda_q < 1, \qquad (28)$$

при не выполнении условия производится изменение порядка операторов скользящего среднего. Структурная схема обобщенного алгоритма экстраполяции значений индикаторов наблюдаемых в ходе моделирования процессов функционирования СИС и пример реализации процесса формирования прогнозных значений индикаторов ЭПК представлены на рис. 1 и рис. 2 соответственно. Следует отметить, что представленный алгоритм позволяет производить настройку модели в ходе начального этапа экстраполяции (фильтрации) и подстройку в ходе дальнейшего наблюдения за процессом изменения значений ЭПК. В отличии от критериев типа «окончательная ошибка предсказания» или информационных критериев выбора порядка модели предлагаемых в методах Берга, Акаике, Парзена [3] предложенный критерий более прост в реализации и позволяет учесть динамику скорости снижения дисперсии при увеличении порядка модели с достаточной для практики точностью [4].

СПИСОК ЛИТЕРАТУРЫ

1. *Бокс Дж., Джекинс Г.* Анализ временных рядов. Прогноз и управление - М.: Мир, 1974.- 402 с.
2. *Статистические методы* прогноза временных рядов (Предварительный анализ и модели прогноза). М.: МГУ, Монография. 2002, 320 с.
3. *Марпл. – мл. С.Л.* Цифровой спектральный анализ и его приложения - М.:Мир, 1990.- 584 с.
4. *Ненадович Д.М.* Методологические аспекты экспертизы телекоммуникационных проектов. – М. Горячая линия – Телеком, 2008 – 272 с. ГИАБ

Коротко об авторе

Ненадович Д.М. – кандидат технических наук, менеджер Департамента эксплуатации ОАО «Ростелеком».

Рецензент д-р техн. наук, проф. *С.А. Редкозубов.*