

1 Правове забезпечення захисту інформації. Проблеми розвитку нормативної та методичної баз системи захисту інформації

УДК 004.681.003

АЛГОРИТМ ПРОГНОЗИРОВАНИЯ ТЕХНИЧЕСКОГО СОСТОЯНИЯ КОМПЛЕКСНЫХ СИСТЕМ ЗАЩИТЫ ИНФОРМАЦИИ

Евгения Иванченко, Владимир Хорошко, Юлия Хохлачева
Национальный Авиационный Университет

Аннотация: Представлен алгоритм прогнозирования технического состояния систем защиты информации и его особенности. Приведена классификация методов прогнозирования по степени формализации, общему принципу действия и способу получения прогнозной информации. Представлена конструктивная классификация методов прогнозирования. Также разработаны подходы и принципы для построения модели с использованием метода группового учета аргументов и определены этапы решения задачи структурно-параметрической идентификации моделей.

Annotation: The article presents a prediction algorithm technical state of information security systems, and its features. A classification of methods for predicting the degree of formalization, the general principle of action and a process for producing predictive information. Presented structural classification of forecasting methods. Also developed approaches and principles to build the model using the method of data handling and groupovogo defined steps of the solution of the problem of structural-parametric identification models.

Ключевые слова: Защита информации, системы защиты информации, алгоритм прогнозирования, методы прогнозирования, информационная безопасность.

1 Введение

Прогнозирование технического состояния сложных систем – сложная и многоплановая задача. Системы защиты информации являются сложными техническими системами, включающими в себя различные подсистемы для обеспечения требуемого уровня защищенности. Поэтому процессы, происходящие в сфере информационной безопасности (ИБ), относятся к сложным, многомерным, динамическим, нестационарным, активным (целенаправленным), что чрезвычайно затрудняет задачу прогнозирования показателей комплексных систем защиты информации (КСЗИ).

Анализ публикаций, посвященных прогнозированию, показывает, что известно около двухсот методов прогнозирования [1]. На рис. 1 и 2 приведены две классификационные схемы методов прогнозирования [1, 2].

В литературе достаточно много аналогичных схем, многие из которых не отвечают принципам классификации (достаточная полнота охвата методов прогнозирования, единство классификационного признака на каждом уровне членения, непересекаемость разделов классификации, возможность дополнения новыми методами).

По степени формализации методы прогнозирования можно разделить на интеллектуальные (применяются в случаях, когда невозможно учесть влияние многих факторов из-за значительной сложности объекта прогнозирования) и формализованные (рис. 1).

Интеллектуальные методы прогнозирования используют оценки экспертов. Различают индивидуальные и коллективные экспертные оценки.

В группу формализованных методов входят методы экстраполяции (наименьших квадратов, экспоненциального сглаживания, скользящих средних, адаптивного сглаживания) и моделирования (структурного, сетевого, матричного, имитационного).

Конструктивная классификация (рис. 2) позволяет наглядно изобразить совокупность методов прогнозирования в виде иерархического дерева и характеризовать каждый уровень иерархии своим классификационным признаком [3].

На первом уровне все методы по признаку «информационное основание метода» делятся на три класса: фактографические, комбинированные и экспертные.



Рисунок 1 – Классификация методов прогнозирования по степени формализации, общему принципу действия и способу получения прогнозной информации

Фактографические методы базируются на фактической информации об объекте прогнозирования и его прошлом развитии.

В экспертных методах используется информация, полученная в процессе систематизированных процедур выявления и обобщения мнений специалистов – экспертов.

Сложность выбора наиболее эффективного метода прогнозирования заключается в определении относительно классификации методов прогнозирования характеристик каждого метода, перечня требований к ретроспективной информации и прогнозному фону.

В процессе систематизированного научно обоснованного прогнозирования функционирования КСЗИ происходило совершенствование методологии прогнозирования как совокупности методов, приемов и способов мышления, позволяющих на основе анализа ретроспективных данных, связей внутри объекта

прогнозирования, а также измерения параметров его в рамках рассматриваемого явления или процесса вынести решение определенной достоверности относительно его будущего развития.

Прогнозирование представляет собой достаточно сложную задачу, что подтверждается рассмотренным анализом причин и факторов, потенциально влияющих на изменение прогнозируемого показателя. Решение такой задачи, как и всякой другой сложной задачи, требует системного подхода, который помогает понять суть проблемы, выбрать адекватные методы ее решения, оценить причины возможных неудач.

Прогнозирование технического состояния СЗИ является одним из решающих научных факторов формирования стратегии и тактики проведения выбора показателей прогнозирования.



Рисунок 2 – Конструктивная классификация методов прогнозирования

Для прогнозирования и моделирования процессов, происходящих в СЗИ, наиболее применимы статистические модели, которые основываются на существующих тенденциях и изменениях показателей прогнозирования.

Модели прогнозирования могут выступать как долгосрочными, так и краткосрочными. Вследствие высокой степени неопределенности в состоянии объекта прогнозирования приоритетное значение отдается краткосрочным прогнозам.

Неотъемлемой составной частью задач управления (оптимального, автоматизированного или на уровне принятия решений) является построение моделей, которые описывают или прогнозируют поведение подсистем, процессов и системы в целом. В общем случае для получения математической модели необходимо выбрать ее структуру и оценить параметры, то есть решить задачу структурной идентификации.

Структурная идентификация нами будет рассматриваться как задача поиска структуры модели с минимальной дисперсией ошибки прогнозирования.

В общем случае процесс решения задачи структурно-параметрической идентификации включает в себя следующие основные этапы.

Этап 1. Задание выборки данных и априорной информации.

Этап 2. Выбор или задание класса базисных функций и преобразование данных.

Этап 3. Генерация различных структур моделей в выбранном классе.

Этап 4. Оценивание параметров генерируемых структур и формирование множества моделей.

Этап 5. Минимизация заданного критерия и выбор оптимальной модели.

Этап 6. Проверка адекватности полученной оптимальной модели.

Этап 7. Принятие решения о завершении процесса.

Перечисленные этапы описывают произвольный процесс построения моделей, причем в зависимости от априорной информации и цели моделирования те или иные этапы могут отсутствовать.

Следует отметить, что задачи моделирования и прогнозирования тесно связаны между собой, но не тождественны. Задачей моделирования является построение математических моделей для количественного описания связи между выходными показателями моделируемых процессов и входными переменными. А прогнозирование – количественное предсказание развития процесса по моделям, полученным тем или иным методом. При этом задача прогнозирования может выступать как экстраполяция закономерного развития моделируемых процессов на будущий промежуток времени или как имитационное моделирование, то есть применение математических моделей для исследования изменения моделируемых процессов с учетом разных вариантов поведения независимых переменных.

Большинство методов построено на разных подходах, что усложняет сравнительный анализ для определения условий их эффективного применения. Для задач управления в КСЗИ важно создавать модели с меньшей ошибкой прогнозирования. Это является основой для сравнения эффективности существующих подходов.

Наиболее распространенными являются две группы методов, на основе которых создаются модели для прогнозирования:

- описание взаимодействия процессов функционирования КСЗИ разного уровня на основе изучения внутренних механизмов функционирования моделируемых процессов и представления полученной информации;

- получение моделей с помощью оценивания параметров по выборкам наблюдений с применением различных методов анализа.

При этом построение математических моделей по выборкам наблюдений с помощью автоматического генерирования и сравнения большого количества различных моделей с выбором наилучшей из них, соответствующей той информации, которая неявно содержится в поступающей информации.

Цель такого подхода – выявления скрытых закономерных причинно-следственных связей между элементами исследуемых процессов на основе статистических данных, в которых эти связи объективно отражены.

Для сложных КСЗИ и процессов, проходящих в них, типичной является неопределенность информации, касающейся механизма их функционирования, степени информативности измерительных переменных и свойств. При этом необходимо применять формализованные методы и автоматизированные способы моделирования.

Поэтому актуальной является проблема разработки алгоритма структурной идентификации прогнозирующих моделей с целью создания автоматизированных способов для оптимального выбора структуры моделей сложных объектов по выборкам ограниченного объема в условиях неполноты информации.

II Основная часть

Рассматриваем статический объект с m входами и одним выходом, результаты наблюдения за которым могут быть представлены матрицей $X[N \times m]$ и вектором $y[N \times 1]$.

По данным N наблюдений требуется определить структуру зависимости одномерной выходной переменной y от набора выходных переменных X в условиях, когда априорно неизвестно, какие именно переменные из набора входных переменных участвуют в формировании выходной переменной, а также неизвестна дисперсия ошибки в наблюдениях выходной переменной y .

Пусть искомая модель объекта может принадлежать некоторому множеству G , в котором каждая модель имеет вид:

$$\hat{y} = f(X, \hat{\theta}_f), \quad (1)$$

где $\hat{\theta}_f$ – вектор параметров модели, оцениваемый тем или иным способом по данным наблюдений.

Задача сводится к поиску минимума заданного критерием J качества модели

$$f^* = \arg \min J(y_1 f(X, \hat{\theta}_f)). \quad (2)$$

Данная постановка задачи показывает, что все методы ее решения могут отличаться, по крайней мере, по следующим признакам:

- 1) методу формирования структуры моделей из множества G ;
- 2) методу оценивания параметров этих моделей;
- 3) самим критерием J оценки их качества;
- 4) организацией движения к минимуму J .

Методы оценивания параметров моделей, критерии их качества и методы поиска минимума критериев в основном независимы между собой и могут применяться в различных сочетаниях. Поэтому может быть предложено много разных методов решения задачи (1) и (2) [4, 5].

Рассмотрим эту задачу в соответствии с идеологией метода группового учета аргументов (МГУА) [6]. Целесообразность использования МГУА объясняется тем, что в нем реализуются итерационные схемы усложнения моделей. Усложнение моделей от ряда к ряду селекции происходит за счет «скрещивания» лучших моделей предыдущего ряда.

Основные принципы МГУА:

- а) промежуточное решение не единственно;
- б) внешнее дополнение (использование промежуточной выборки);
- в) самоотбор промежуточных решений;
- г) единственность окончательного решения.

Представление о наилучшей структуре модели (1) довольно условно и зависит от требований исследователя. Обычно предполагают, что у можно аппроксимировать полиполями. Если степень полиполю не выше k , то задачу выбора структуры модели можно представить в виде поиска гиперкуба $[0,1]^m$. Из всех вершин нужно найти такую, в которой достигается экстремум $g_r(V,k): g_r(V,k) \rightarrow \min$, где r – число выбранных переменных, $1^V \leq r \leq m$, $V=(V_1, V_2, \dots, V_m)$, $V_j=1$, если переменная x_j входит в модель, $V_j=0$ – в противном случае.

При выборе структуры модели следует руководствоваться двумя принципами:

- 1) стремиться к тому, чтобы модель была проще, то есть включала как можно меньше коэффициентов полиполю (принцип экономичности);
- 2) улучшать модель, проверяя ее адекватность (принцип адекватности).

Для минимизации заданного критерия (2), определенного на множестве моделей G , применяются как переборные, так и итерационные алгоритмы.

Анализ приведенных выше критериев и методов показал, что для построения математической модели нелинейной регрессионной зависимости вида (1) и (2) целесообразно использовать [7]:

- а) МГУА в качестве метода перебора моделей;
- б) метод наименьших квадратов и метод модулей в качестве методов оценки параметров моделей;
- в) критерии остаточной суммы квадратов, регулярности, скользящего контроля для оценки качества получаемых моделей.

Исходя из всего сказанного, примем, что закон функционирования исследуемого объекта имеет вид

$$y = y^0 + \xi = \sum_{j=1}^{m^0} \theta_j^0 x_j + \xi \quad (3)$$

где y – наблюдаемый выход объекта; y^0 – ненаблюдаемый незашумленный выход объекта; ξ – ненаблюдаемая случайная ошибка измерения; x_j^0 – j -й вход объекта из множества входов χ^0 , участвующих в формировании выхода объекта; m^0 – число входов, принадлежащих множеству χ^0 ; $\theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_m^0)$ – вектор не равных нулю неизвестных коэффициентов.

Множество входов χ^0 неизвестно. Известно лишь, что $\chi^0 \subset \chi$, где χ – некоторое множество точно измеряемых m входов объекта.

Пусть в результате наблюдения объекта получены:

- 1) X – матрица (размерности $N \times m$) N наблюдений m входов множества χ , $X = m$;
 - 2) y – вектор (размерности N) соответствующих наблюдений выходной величины y .
- В соответствии с законом функционирования объекта (3) должно выполняться равенство:

$$y = y^0 + \xi = X^0 \theta^0 + \xi \quad (4)$$

где y^0 – вектор (размерности N) значений ненаблюдаемого незашумленного выхода объекта; X^0 – матрица (размерности $N \times m^0$) наблюдений входов объекта, принадлежащих множеству χ^0 ; ξ – вектор (размерности N) ненаблюдаемых случайных ошибок измерения.

Пусть относительно ξ выполнены предположения

$$E\{\xi\} = O_N, \quad E\{\xi, \xi^T\} = \sigma^2 I_N,$$

где $E\{\cdot\}$ – знак математического ожидания; T – знак транспонирования; O_N – нулевой вектор (размерности $N \times 1$); σ^2 – неизвестная конечная величина; I_N – единичная матрица (размерности $N \times N$).

Требуется найти:

- 1) множество χ^0 ;
- 2) оценку вектора коэффициентов θ^0 ;
- 3) оценку дисперсии ошибки измерений σ^2 .

Рассмотрим алгоритм структурной идентификации. Пусть класс синтезируемых моделей имеет вид:

$$\hat{Z} = \sum_{q=1}^S \theta_q \cdot \prod_{j=1}^m x_j^{\alpha_{qj}} \quad (5)$$

где \hat{Z} – выходная функция, S – число членов в модели; θ_q , $q=1, 2, \dots, S$ – коэффициенты; x_j , $j=1, 2, \dots, m$ – входные переменные; m – число входных переменных; α_{qj} – показатель степени, с которой переменная x_j входит в q -й член.

Структурой частного описания называем набор параметров α_{qj} и S , определяющих \hat{Z} в представлении (5).

Для построения итерационного алгоритма МГУА необходимо:

- 1) указать начальную матрицу частных описаний \hat{Z}^0 ;
- 2) определить оператор R , осуществляющий отображение $\hat{Z}^{r-1} \xrightarrow{R} \hat{Z}^r$, $r=1, 2, \dots$ – номер итерации;
- 3) указать правило завершения итераций.

Общий вид матрицы частных описаний \hat{Z}^r определим следующим образом:

$$\hat{Z}^r = [\hat{Z}_1^r : \hat{Z}_2^r : \dots : \hat{Z}_{F+2+m+2S}^r] \quad (6)$$

где \hat{Z}_j^r , $j=1, 2, \dots, F+2+m+2S$ – векторы (размерности N) частного описания; F – число лучших частных описаний, передаваемых от итерации к итерации; S – число членов в структуре лучшего частного описания итерации ($r-1$).

Обозначим:

$$\begin{aligned} \hat{G}^r &= [\hat{Z}_1^r : \hat{Z}_2^r : \dots : \hat{Z}_F^r], \\ \hat{C}^r &= [\hat{Z}_{F+1}^r : \hat{Z}_{F+2}^r : \dots : \hat{Z}_{F+2+m}^r], \\ \hat{D}^r &= [\hat{Z}_{F+2+m+1}^r : \hat{Z}_{F+2+m+2}^r : \dots : \hat{Z}_{F+2+m+2S}^r]. \end{aligned} \quad (7)$$

Алгоритм состоит из следующих шагов.

Шаг 1. Указывается начальная матрица частных описаний \hat{Z}^0 (для нее полагаем $S=0$):

$$\hat{Z}^0 = [O : o : X : I] = [O : \hat{C}^0], \quad (8)$$

где O – нулевая матрица (размерности $N \times F$); o – нулевой вектор (размерности N); I – единичный вектор (размерности N); $X = [x_1 : x_2 : \dots : x_m]$ – матрица наблюдений входных переменных (размерности $N \times m$).

Шаг 2. Определение вектора R .

Пусть векторы \hat{Z}^r строятся по правилу:

$$\hat{Z}^r(i) = \hat{a}\hat{Z}_{j_1}^{r-1}(i) + \hat{b}\hat{Z}_{j_2}^{r-1}(i) \cdot \hat{Z}_{j_3}^{r-1}(i), \quad (9)$$

где $r = 1, 2, \dots$ – номер итерации; $i = 1, 2, \dots, N$ – номер наблюдения, который для удобства записи взят в скобки; $j_1, j_2, j_3 = 1, 2, \dots, F+2+m+S$, ($j_1 \geq j_2$) – номера частных описаний из матрицы \hat{Z}^{r-1} ; \hat{a}, \hat{b} – коэффициенты, определяемые на обучающей подвыборке наблюдений (A).

Значения коэффициентов \hat{a}, \hat{b} находятся как решение задачи минимизации

$$\hat{a}, \hat{b} = \arg \min_{a,b} \Phi(a,b), \quad \Phi(a,b) = \sum_{i=1}^{N(A)} e_A^2(i), \quad (10)$$

где $e_A(i)$, $i = 1, 2, \dots, N(A)$, – остатки в регрессии y_A по двум переменным:

$$y_A(i) = a \cdot \hat{Z}_{j_1 A}^{r-1}(i) + b \cdot \hat{Z}_{j_2 A}^{r-1}(i) \hat{Z}_{j_3 A}^{r-1}(i) + e_A(i). \quad (11)$$

Из всех генерируемых по правилу (9) – (11) частных описаний отбираются F описаний, лучших по минимуму квадратичной нормы остатков на проверочной подвыборке наблюдений (B):

$$J = \frac{1}{N(B)} \sum_{i=1}^{N(B)} \hat{e}_B^2(i),$$

$$\hat{e}_B(i) = y_B(i) \hat{Z}_B^r(i), \quad (12)$$

$$\hat{Z}_B^r(i) = \hat{a} \cdot \hat{Z}_{j_1 B}^{r-1}(i) + \hat{b} \cdot \hat{Z}_{j_2 B}^{r-1}(i) \hat{Z}_{j_3 B}^{r-1}(i),$$

где $N(B)$ – объем проверочной подвыборки B .

Отобранные лучшие частные описания, ранжированные по уменьшению величины J , используются при формировании матрицы

$$\hat{G}^r = [\hat{Z}_1^r : \hat{Z}_2^r : \dots : \hat{Z}_F^r] \quad (13)$$

Матрица \hat{C}^r не изменяется: $\hat{C}^r = \hat{C}^{r-1}$. Матрица \hat{D}^r формируется с учетом структуры лучшего из F отобранных частных описаний (\hat{Z}_F^r). Первые S столбцов матрицы \hat{D}^r заполняются отдельными членами лучшего частного описания по правилу:

$$\hat{d}_h^r(i) = \hat{Z}_{F+2+m+h}^r(i) = \hat{\theta}_h \prod_{j=1}^m x_j^{\alpha_{hj}}(i), \quad (14)$$

где $h = 1, 2, \dots, S$ – номер члена в структуре; S – число членов в структуре лучшего частного описания.

Вторые S столбцов матрицы \hat{D}^r формируются поочередным исключением отдельных членов из структуры лучшего частного описания по правилу:

$$\hat{d}_{S+h}^r(i) = \hat{Z}_{F+2+m+S+h}^r(i) = \sum_{q=1(q \neq h)}^S \hat{\theta}_q \prod_{j=1}^m x_j^{\alpha_{hq}}(i), \quad (15)$$

Таким образом определяется оператор R преобразования

$$\hat{Z}^{r-1} \xrightarrow{R} \hat{Z}^r.$$

Правило остановки для итерационной схемы: вычисления заканчиваются на итерации r^* , если выполняется условие

$$J(\hat{Z}_F^{r^*-1}) - J(\hat{Z}_F^{r^*}) < \delta_r, \quad (16)$$

где $J(\hat{Z}_F^{r^*})$ – значение критерия для лучшего частного описания итерации r ; δ_r – заданное число.

Особенностью алгоритма является его многоэтапность. Номер текущего этапа определяет максимально возможное число членов в моделях. Синтез моделей начинается с этапа с номером $p=1$ или с любого заданного номера p^* .

Каждый этап представляет собой итерационную схему (8) – (16). Начальная матрица частных описаний этапа с номером p задается конечной матрицей частных описаний предыдущего этапа

$$\hat{Z}_p^0 = \hat{Z}_{p-1}^{r^*}, \quad (17)$$

а для $p=p^0$ она совпадает с (8). Вычисления заканчиваются на этапе p^* , если выполнено условие

$$J(\hat{Z}_{F,p^*-1}^r) - J(\hat{Z}_{F,p^*}^r) < \delta_p, \quad (18)$$

где $J(\hat{Z}_{F,p}^r)$ – значения критерия для лучшего частного описания r -ой итерации этапа p ; δ_p – заданное число.

III Выводы

Особенностью предлагаемого алгоритма является:

- 1) многоэтапность поиска модели;
- 2) поиск модели как в классе линейных, так и в классе нелинейных по входным переменным моделей;
- 3) приемы исключения отдельных членов лучшего частного описания и на основе этого расширение базисного набора аргументов;
- 4) оптимальная по вычислительным затратам для итерационных алгоритмов МГУА схема расчета критерия скользящего экзамена;
- 5) возможность оценивать коэффициенты в моделях как по методу наименьших квадратов, так и по методу наименьших модулей.

Кроме того, алгоритм позволяет прогнозировать техническое состояние КСЗИ, а это в свою очередь дает возможность обеспечивать требуемый уровень информационной безопасности объектов различных классов, сложности и назначения.

Литература: 1. Тихонов А. Н. Методы решения некорректных задач / Тихонов А. Н., Арсенин В. Я. – М.: Наука, 1974. – 458 с. 2. Основы экономического и социального прогнозирования / Под ред. Мосина Н. – М.: Высшая школа, 1985. – 386 с. 3. Бегма Т. В. Математичні моделі функціонування складних систем / Бегма Т. В., Капустян М. В., Хорошко В. О. // Вісник СНУ ім. В. Даля, №7(161), 2.1, 2011. – С. 252–263. 4. Вучков И. Прикладной линейный регрессионный анализ / Вучков И., Болджнева Л., Салаков Е. – М.: Финансы и статистика, 1987. – 239 с. 5. Степашко В. С. Методы и критерии решения задач структурной идентификации / Степашко В. С., Кочерга Ю. Л. // Автоматика, №5, 1985. – С. 29–37. 6. Сарычев А. П. Решение проблемы разбиения в МГУА при расчете критерия регулярности в условиях активного эксперимента / Сарычев А. П. // Автоматика, №4, 1989. – С. 19–27. 7. Головань С. М. Основи надійності інформаційних систем / Головань С. М., Корнейко О. В., Петров О. С., Хорошко В. О., Щербак Л. М. – Луганськ: Вид. «Наулідж», 2012. – 335 с.

УДК 004.924:534.6.08

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ИЗМЕНЕНИЯ ПАРАМЕТРОВ ВОДНОЙ СРЕДЫ КАК КАНАЛА ТРАНСЛЯЦИИ СНИМАЕМОЙ РЕЧЕВОЙ ИНФОРМАЦИИ

Елена Азаренко, Михаил Дивизинюк, Юлия Гончаренко, Дмитрий Гончаренко
Севастопольский национальный университет ядерной энергии и промышленности

Анотация: Запропонована математична модель зменшення швидкості звуку сторонньою домішкою у водному середовищі, використовуваному для трансляції мовної інформації, що знімається.

Summary: The mathematical model of diminishing of speed of sound is Offered by an extraneous admixture in the water environment used for translation of the taken off speech information.

Ключевые слова: Акустический сигнал, скорость звука, речевая информация, упругая среда, примесь.

I Введение

Решение задачи съема речевой информации в общем случае сводится к решению ряда частных задач, таких как регистрация речевой информации приемным (микрофонным) устройством, трансляция преобразованных зарегистрированных сигналов к приемному устройству, преобразование полученных сигналов в виде, приемлемом для злоумышленника [1]. Одним из методов решения частной задачи скрытой трансляции данных на значительные расстояния является метод передачи высокочастотных акустических сигналов в водной среде, модулированной снятой речевой информацией [2]. Для реализации подобного метода достаточно обязательных конструктивных систем, которыми оснащаются административные здания и жилые дома, спортивные сооружения и увеселительные заведения, загородные особняки, кемпинги и т. п.