

Мигуценко Р. П., канд. техн. наук, доц.
Национальный технический университет «Харьковский политехнический институт»

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ОГРАНИЧЕННОСТИ АПРИОРНОЙ ИНФОРМАЦИИ НА ВИД И РАЗМЕР ДОСТОВЕРНОСТИ ДИАГНОСТИКИ

mrp1@bk.ru

Рассмотрена задача математического моделирования вероятностных свойств достоверности диагностики, осуществляемой на базе дискриминантной (решающей) функции при заданных ограничениях на объем обучающей выборки. Представлены варианты возможных проявлений ограниченности объема обучающей выборки, влияющей на количество ожидаемой диагностической информации на выходе ИИС диагностики.

Ключевые слова: ошибка, дисперсия, достоверность, погрешность, контроль.

Постановка проблемы. В любых информационно-измерительных системах (ИИС) контроля, диагностики и идентификации, которые осуществляют преобразование первичной измерительной многомерной информации в информацию вторичную одномерную, представленную логическими решениями, эффективность последних зависит от степени адекватности математических моделей информационных преобразований изменениям физических свойств объекта вибродиагностики. Степень такой адекватности зависит от корректности выполнения следующих условий планирования диагностического эксперимента [1, 2]:

- обеспечение заданной точности оценивания коэффициентов математической модели информационно-логического преобразования результатов первичных измерений во вторичные статистические решения;

- выбор наиболее информативных составляющих вектора входных измерительных сигналов ИИС контроля и диагностики;

- выбор показателей эффективности работы ИИС, которые учитывают априорную неопределенность оценок коэффициентов математической модели преобразования, оптимальных, по максимуму получаемой диагностической информации;

- выбор размерности вектора входных сигналов и объема измерений в реальных условиях эксплуатации ИИС.

Выполнение перечисленных условий связано, во многом, с изучением объекта диагностики, построения его вероятностной модели на этапе обучения, что обеспечивает максимально возможное количество ожидаемой диагностической информации при эксплуатации ИИС диагностики. Такое изучение вероятностных свойств объекта диагностики и максимизация информационно-диагностических возможностей проектируемых ИИС функциональной диагностики требует решения следующих задач:

- статистическое обоснование выбора вида математической модели измерительно-

логических преобразований для процедуры функциональной диагностики с учетом ограниченности априорной информации о свойствах объекта диагностики;

- построение и анализ информационной модели процедуры диагностики, учитывающей требования плана диагностического эксперимента и вероятностные свойства математической модели этой процедуры;

- построение информационных моделей процедур комплексного преобразования многомерных измерительных сигналов в альтернативные диагностические решения, которые учитывают нормативные риски возможных вероятностей ошибок.

Цель статьи. Раскрытие возможности математического моделирования вероятностных свойств достоверности диагностики, осуществляемой на базе дискриминантной (решающей) функции при заданных ограничениях на объем обучающей выборки.

Количественное влияние объема обучающей выборки на достоверность диагностики. При использовании для альтернативной функциональной диагностики линейной решающей функции

$$g(x) = \sum_{i=1}^n \frac{(m_i^{(0)} - m_i^{(1)})}{\sigma_i^2} \left[x_i - \frac{(m_i^{(0)} + m_i^{(1)})}{2} \right],$$

где $m_i^{(0)}$, $m_i^{(1)}$ – оценки условных средних для i -той составляющей x_i входного вектора \bar{X} , $i = \overline{1, n}$;

σ_i^2 – оценка дисперсии x_i , ее коэффициенты $m_i^{(0)}$, $m_i^{(1)}$, σ_i , $i = \overline{1, n}$, оценивают в ходе обучения системы диагностики. Такое обучение осуществляют по выборкам N_0 и N_1 объектов с верифицированными состояниями S_0 и S_1 .

Если в ходе обучения системы диагностики измерения составляющих x_1, \dots, x_n входного вектора \bar{X} сопровождается случайными по-

грешностями, то возникает дополнительная неопределенность в значениях коэффициентов решающей функции. Обнаружим влияние такой неопределенности на количество ожидаемой диагностической информации, если значение входного вектора \bar{X} представлено фиксированным множеством $\{x_1, \dots, x_n\}$ результатов измерения его случайных составляющих (x_1, \dots, x_n) .

Используя известные выражения [3]:

$$\sigma_{\xi}^2 = \delta^2 + \delta^4 q_0 q_1 \text{ и } \sigma_{\xi}^2 = \delta^2 (2N)^{-1},$$

где δ^2 – квадрат геометрического расстояния между векторами $\bar{X}^{(0)}$ и $\bar{X}^{(1)}$;

$$I = \frac{1}{2} \log_2 \left\{ 1 + \frac{(1 + \delta^2 q_0 q_1)}{\delta^2 [\alpha(1 - \alpha)q_0 + \beta(1 - \beta)q_1] + (2N)^{-1}} \right\}.$$

Если количество ожидаемой диагностической информации равно L бит ($I = L$), то, с учетом того, что для линейной решающей функции:

$$\begin{cases} \alpha = \beta; \\ 1 - \alpha = 1 - \beta = P_D, \end{cases}$$

$$-P_D^2(2^{2L} - 1)\delta^2 + P_D(2^{2L} - 1)\delta^2 + \left[\frac{(2^{2L} - 1)}{2N} - 1 - \delta^2 q_0 q_1 \right] = 0. \quad (1)$$

Решая квадратичное, относительно P_D , уравнение (1), можно найти зависимость максимально возможной достоверности диагностики $P_{D \max}$ от количества ожидаемой информации L , при условии $N \rightarrow \infty$:

$$P_{D \max} = 0.5 + 0.5 \sqrt{1 - \frac{4}{\delta^2(2^{2L} - 1)} - \frac{4q_0 q_1}{(2^{2L} - 1)}}.$$

Качественное влияние объема обучающей выборки на достоверность диагностики. Кроме количественного имеет место и качественное влияние объема обучающей выборки N на достоверность диагностики. Это влияние выражается в неверном выборе вида решающей функции и нарушении априорных предположений о модели условных законов распределения составляющих вектора \bar{X} , включая априорные предположения об их взаимно корреляционных свойствах. На рис. 1 Представлены варианты возможных проявлений ограниченности объема обучающей выборки, влияющей на количество ожидаемой диагностической информации на выходе ИИС функциональной диагностики.

Рис. 1 показывает, что для конкретной системы диагностики величина достоверности принятия решений – это величина постоянная, которую можно повысить, увеличивая объем

q_0, q_1 – априорные вероятности состояний S_0, S_1 объекта диагностики,

для дисперсии решающей функции до и после диагностики, найдем из уравнения [4]

$$I = \log_2 \sqrt{1 + \frac{\sigma_x^2}{\sigma_{\Delta x}^2}}$$

количество ожидаемой диагностической информации с учетом случайных погрешностей в оценке коэффициентов решающей функций при ограниченных обучающих выборках ($N \ll \infty$):

можно получить следующее уравнение, связывающее достоверность диагностики P_D , геометрическое расстояние δ , количество информации L , априорные вероятности q_0, q_1 и объем выборки N :

обучающей выборки. Если рассмотреть множество ИИС диагностики, которые используют одинаковую модель решающей функции, но обученных и калиброванных по разным, причем одинаковым по объему N , обучающим выборкам, то реализации достоверности диагностики этих ИИС будут различными проявлениями случайной достоверности P_N , математическое ожидание которой $M[P_N]$ и дисперсия $D[P_N]$ будут функциями объема обучающей выборки N . При увеличении N выполняется:

$$\begin{cases} a) M[P_N] \xrightarrow{N \rightarrow \infty} P_{D \max}; \\ б) D[P_N] \xrightarrow{N \rightarrow \infty} 0. \end{cases} \quad (2)$$

Достоверность $M[P_N]$ – это средняя (ожидаемая по среднему) достоверность диагностики, а $P_{D \max}$ – достоверность асимптотическая.

Если P_{\max} – это предельно максимальная, при отсутствии качественных и количественных проявлений ограниченности объема обучающей выборки, достоверность, то

$$P_{D \max} \leq P_{\max}. \quad (3)$$

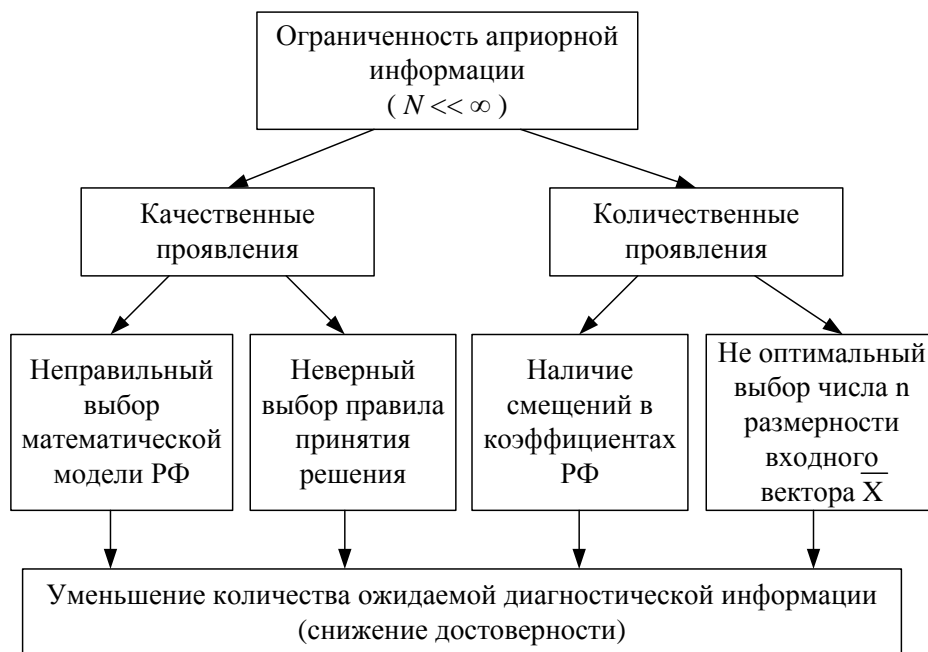


Рис. 1. Варианты воздействия ограниченности объема обучающей выборки N на причины снижения достоверности функциональной диагностики

В неравенстве (3) предполагается, что для $P_{Д\max}$, имеют место качественные проявления ограниченности объема обучающей выборки N (например, возможен неверный выбор модели

решающей функции). Приведенные рассуждения характерны и для вероятностей ошибок диагностики [5].

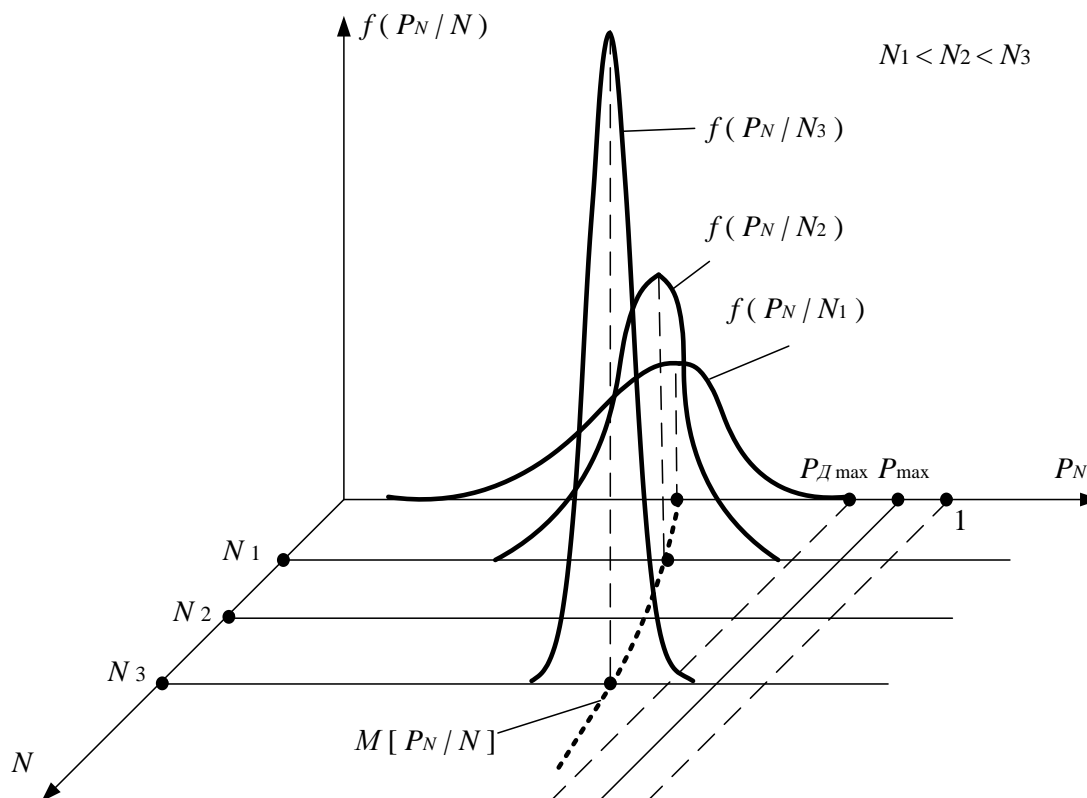


Рис. 2. Вероятностные графические модели видов достоверности диагностики

Из условий (2) следует, что для случайных достоверностей P_N существует условная плотность $f(P_N|N)$ распределения их вероятностей, которая зависит от объема обучающей выборки

N , поэтому достоверность P_N будем называть условной. На рис. 2 представлены плотности $f(P_N|N)$, математическое ожидание $M[P_N]$ и достоверности $P_{Д\max}$ и P_{\max} , графически ил-

люстрируя вероятностные свойства самой достоверности P_D при ограничениях на объем обучающей выборки N и сложность ее оценки.

Из рис. 2 и выражений (2) следует, что в реальных условиях (при $N \ll \infty$) эффективность работы любой ИИС диагностики определяется вероятностями P_N и $M[P_N]$, причем существует статистическая возможность не только точечного оценивания этих видов достоверности, но и оценка интервальная, учитывающая объем обучающей выборки N и заданную доверительную вероятность оценивания.

Выводы. 1. Предложена новая классификация видов достоверности функциональной диагностики, учитывающая качественные и количественные проявления нарушений при синтезе решающих функции из-за ограничения объема обучающих выборок. 2. Избраны конкретные виды достоверности диагностики, учитывающие объемы обучающих выборок и допускающие статистическое оценивание их величин при заданной доверительной вероятности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Раудис Ш. Ограниченность выборки в задачах классификации // Сборник «Статистические проблемы управления» – Вильнюс, 1976. Вып. 18. С. 1 – 185.
2. Бурданова Е.В., Денисов А.П. Моделирование решающих функций при обнаружении объектов по радиолокационным данным // Научные ведомости белгородского государственного университета. № 9-1-1. Т 1, 2009. С. 51 – 55.
3. Неделько В. М. О точности интервальных оценок вероятности ошибочной классификации, основанных на эмпирическом риске // Всероссийская конференция ММРО-14. М.: МАКС Пресс, 2009. С. 56–59.
4. Орнатский П.П. Теоретические основы информационно-измерительной техники. К.: Вища школа, 1983. 455 с.
5. Володарский Е.Т., Малиновский Б.Н., Туз Ю.М. Планирование и организация измерительного эксперимента. К.: Вища шк. Головное изд-во, 1987. 280 с.

1.