УДК 621.317

МЕТОД ОПРЕДЕЛЕНИЯ КОЭФФИЦИЕНТОВ ЗАВИСИМОСТИ КОНТРОЛИРУЕМЫХ ПАРАМЕТРОВ НА ОСНОВЕ АТРИБУТИРОВАННОГО БИНАРНОГО ДЕРЕВА

С.В. Герасимов[1], П.С. Слупский[1], А.А. Феклистов[1], Д.В. Чуйков[2]
([1]Объединенный научно-исследовательский институт Вооруженных Сил, Харьков,
[2]Научный метрологический центр (военных эталонов), Харьков)

В статье предлагается метод определения коэффициентов зависимости параметров контроля сложного технического комплекса, основанный на построении и обработке атрибутированного бинарного дерева, который может быть использован при определении оптимального состава контролируемых параметров сложных технических комплексов.

коэффициенты зависимости, параметры контроля, атрибутированное бинарное дерево, сложные технические комплексы

Постановка проблемы. В обеспечении высоких качественных показателей сложного технического комплекса (СТК) на всех этапах жизненного цикла важная роль отводится его метрологическому обслуживанию (Моб) [1]. Эффективность проведения метрологического обслуживания СТК зависит от количества и номенклатуры контролируемых параметров (КП). В тоже время необоснованно большое количество КП приводит к увеличению времени проведения Моб и, как следствие, к повышению затрат на эксплуатацию СТК. Поэтому для обеспечения оперативности проведения измерения и контроля параметров технического комплекса необходимо определить минимальный состав его КП, который обеспечивает требуемую достоверность Моб СТК [2].

При выборе КП необходимо ориентироваться на контроль независимых параметров. Независимыми считаются параметры, выход за пределы допуска которых не вызывает выход за допуск других параметров, т.е. параметры не связанные между собой [3]. Во многих случаях КП сложного технического комплекса могут зависеть друг от друга, тогда достаточно проконтролировать один из них для получения информации о техническом состоянии СТК. Поэтому при выборе оптимального состава КП сложного технического комплекса актуальной является задача определения зависимых параметров, решение которой позволяет сократить число параметров контроля.

Анализ литературы. Для определения зависимости между параметрами используются экспертные методы и метод, основанный на определении корреляционной связи между параметрами [3] – [6]. Экспертные методы просты в реализации, но основаны на использовании субъективных оценок экспертов, что может существенно влиять на оценку зависимости параметров. Недостаток метода, основанного на определении корреляционной связи между параметрами, заключается в том, что он не позволяет учитывать погрешности измерения параметров. Исключить указанные недостатки позволяет метод, который базируется на построении атрибутированного бинарного дерева.

Целью статьи является разработка метода определения коэффициентов зависимости параметров контроля сложного технического комплекса, основанного на построении и анализе атрибутированного бинарного дерева.

Основная часть. Одним из подходов к определению связей между контролируемыми параметрами является формирование гипотез о наличии зависимостей между параметрами с квантором ассоциации [6]. Под бинарной ассоциацией понимается такое отношение между двумя параметрами x_i и x_j $(i \neq j)$ некоторого объекта, при котором “совпадение сильнее различия”, то есть количество случаев, в которых параметры x_i и x_j одновременно имеют либо одновременно не имеют определенные значения, больше количества остальных случаев, в которых один из них не имеет данного значения. Параметры объекта представляются множеством $X=\{x_1,\ldots,x_{n_x}\}$, которому на основе функций соответствия ставятся унарные предикаты $\varphi(\tau)=\{\varphi_1(\tau),\ldots,\varphi_n(\tau)\}$, принимающие три логических значения: 1 – “истина”, 0 – “ложь” и x – “не определено” [7] – [9]. Логическое значение записывается перед предикатом в круглых скобках (например, $(1)\varphi(\tau)$ обозначает, что унарный предикат $\varphi(\tau)$ имеет значение “истина”). Переменная “τ” соответствует номеру случая. Введем функцию $|\{M\}|$, которая определяет мощность множества $\{M\}$. Коэффициенты количества случаев, для которых предикаты обладают заданными значениями, рассчитываются следующим образом:

$$\begin{aligned} a_{11} &= \left|\left\{\tau \,\middle|\, (1)\varphi_i(\tau) \wedge (1)\varphi_j(\tau)\right\}\right|; \quad a_{10} = \left|\left\{\tau \,\middle|\, (1)\varphi_i(\tau) \wedge (0)\varphi_j(\tau)\right\}\right|; \\ a_{01} &= \left|\left\{\tau \,\middle|\, (0)\varphi_i(\tau) \wedge (1)\varphi_j(\tau)\right\}\right|; \quad a_{00} = \left|\left\{\tau \,\middle|\, (0)\varphi_i(\tau) \wedge (0)\varphi_j(\tau)\right\}\right|. \end{aligned} \tag{1}$$

В качестве численной меры ассоциации предлагается использовать величину “обратная сопряженность” r, которая рассчитывается следующим образом [6]:

$$r = \frac{p_{11} \cdot p_{00}}{p_{10} \cdot p_{01}}, \tag{2}$$

где $p_{ij} = a_{ij} / m$ – вероятность выполнения формулы $(c)\,\varphi_i(\tau) \wedge (d)\,\varphi_j(\tau)$; $c, d \in \{0, 1\}$; $i \neq j$; m – количество случаев.

Считается, что всякая разумная мера зависимости является строго монотонной функцией r [9]. При этом, если взять логарифмическую сопряженность $\delta = \log(\Delta)$, то при $\delta > 0$ зависимость положительна, при $\delta < 0$ – отрицательна, а при $\delta = 0$ – отсутствует. Численная мера квантора бинарной ассоциации $\gamma_{\approx_2}$ вводится как величина, обратная "обратной сопряженности" Едвардса [8]:

$$\gamma_{\approx_2} = \frac{1}{r} = \frac{p_{10} \cdot p_{01}}{p_{11} \cdot p_{00}} = \frac{a_{10} \cdot a_{01}}{a_{11} \cdot a_{00}}. \tag{3}$$

Величина $\gamma_{\approx_2} \in [0..+\infty]$; она не определена в двух случаях: числитель и знаменатель равны ($\gamma_{\approx_2} = 1$); вычисление числителя либо знаменателя невозможно из-за равенства 0 одного из коэффициентов. Гипотеза с квантором бинарной ассоциации имеет следующий вид:

$$\approx_{\gamma\approx_n}(\varphi_i(\tau),\ \varphi_j(\tau)).$$

Под многомерной ассоциацией понимается гипотеза $\approx_{\gamma\approx_n}(\varphi_1(\tau),\ \varphi_2(\tau),\ \ldots, \varphi_n(\tau))$, которая описывает такую ассоциативную связь между n параметрами $x_1,\ x_2,\ \ldots,\ x_n$, при котором количество случаев, в которых эти параметры одновременно имеют либо не имеют определенные значения, больше количества случаев, в которых один либо некоторые из них не имеет данных значений.

Для вычисления численной меры ассоциативной зависимости параметров предлагается метод формирования гипотез с квантором многомерной ассоциации $\gamma_{\approx_n}$, основанный на представлении коэффициентов a в виде узлов атрибутированного бинарного дерева (рис. 1).

Каждый узел имеет обязательный атрибут "код", представляющий последовательность нулей и единиц, которые соответствуют логическим значениям унарных предикатов (например, код узла a_{1101} – "1101"). Код узла используется при классификации узлов дерева на два равных множества: множество $\{a^+\}$, характеризующее ассоциативное сходство параметров; множество $\{a^-\}$, характеризующее ассоциативное различие.

Численная мера квантора $\gamma_{\approx_n}$ многомерной ассоциации КП СТК определяется следующим образом:

$$\gamma_{\approx_n} = \frac{\prod\limits_{i=1}^{n^-} a_i^-}{\prod\limits_{j=1}^{n^+} a_j^+}, \tag{4}$$

где Π – операция произведения; $n = |\{a\}|$ – общее количество узлов дерева; $n^+ = |\{a^+\}$ – количество узлов ассоциативного сходства; $n^- = |\{a^-\}|$ – количество узлов ассоциативного различия; $\{a\} = \{a^+\} \cup \{a^-\}$; $n = n^+ + n^-$; $n^+ = n^- = 1/2 \cdot n$.

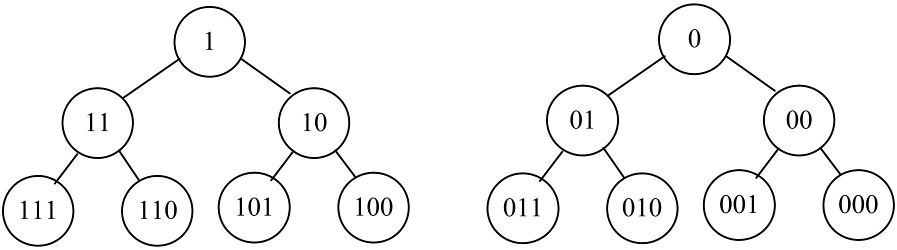

Рис. 1. Представление коэффициентов a в виде узлов атрибутированного бинарного дерева

Соотношение (4) позволяет определить наличие (отсутствие) связи между параметрами, т.е. коэффициент зависимости контролируемых параметров СТК. Для расчета выражения (4) необходимо найти значение коэффициентов множества $\{a\}$.

Анализ формул (1) – (4) показывает, что для классификации узлов дерева достаточно определения одного из множеств $\{a^+\}$ или $\{a^-\}$. Рассмотрим определение множества $\{a^+\}$, учитывая, что $\{a^-\} = \{a\} \backslash \{a^+\}$. Определение множества $\{a^+\}$ основано на введении набора атрибутов для узла и связанных с ними коэффициентов.

Объединение всех узлов $\{a\} = \{a_{=1}\} \cup \{a_{=0}\} \cup \{a_{>1}\} \cup \{a_{>0}\} \cup \{a_{0=1}\}$. Коэффициент n равен: $n = n_{=1} + n_{=0} + n_{>1} + n_{>0} + n_{0=1}$. Обозначим множество узлов $\{a_{=1}\}$, $\{a_{=0}\}$, $\{a_{>1}\}$ и $\{a_{>0}\}$ как $\{a_\Sigma\}$:

$\{a_\Sigma\} = \{a_{=1}\} \cup \{a_{=0}\} \cup \{a_{>1}\} \cup \{a_{>0}\}$. Мощность множества $\{a_\Sigma\}$ равна $n_\Sigma = n_{=1} + n_{=0} + n_{>1} + n_{>0}$.

Классификация узлов дерева основана на анализе двух условий: $n_\Sigma = 1/2\,n$ либо коэффициент $n_\Sigma > 1/2\,n$.

Если коэффициент $n_\Sigma = 1/2\,n$, то узлы множеств $\{a_{=1}\}$, $\{a_{=0}\}$, $\{a_{>1}\}$, $\{a_{>0}\}$ являются узлами ассоциативного сходства, а узлы $\{a_{0=1}\}$ – узлами ассоциативного различия

$$\{a^+\} = \{a_{=1}, a_{=0},\ a_{>1}, a_{>0}\};\ \ \{a^-\} = \{a_{0=1}\}.$$

Если коэффициент $n_\Sigma > 1/2\,n$, то выполняются следующие операции. Объединим множества $\{a_{>1}\}$ и $\{a_{>0}\}$ во множество $\{a_?\}$: $\{a_?\} = \{a_{>1}\} \cup \{a_{>0}\}$). Из $\{a_?\}$ выберем узлы, которые характеризуют ассоциативное сходство; все остальные узлы классифицируем как узлы ассоциативного различия. Определения узлов, характеризующих ассоциативное сходство, основано на предположении о том, что чем чаще параметры встречаются совместно, тем выше ассоциативное сходство между ними. Таким образом, в качестве a^+ выбираются узлы с максимальными значениями коэффициентов. Определим их как узлы "возможного ассоциативного сходства" $(a_?^+)$. Количество данных узлов $n_?^+ = |\{a_?^+\}| = 1/2\,a - n_{=1} - n_{=0}$. Все остальные узлы определим как узлы "возможного ассоциативного различия" $(a_?^-)$. Количество данных узлов есть: $n_?^- = |\{a_?^-\}| = 1/2 \cdot a$. К множеству $\{a^-\}$ относятся также все узлы $\{a_{0=1}\}$

$$\{a^+\} = \{a_{=1}^+,\ a_{=0}^+,\ a_?^+\};\ \ \ \{a^-\} = \{a_{0=1},\ a_?^-\}.$$

При рассмотрении узлов арности 5 и более было замечено, что среди узлов $a_{>1}$ и $a_{>0}$ есть такие, в которых по количеству единиц (нулей) одни узлы более подходят на отбор в множество $\{a^+\}$, чем другие. Например, узел a_{11101} имеет больше 1, чем a_{11100} и, следовательно, в большей степени характеризует "ассоциативное сходство". Введем дополнительное ограничение на вычисление численной меры ассоциации для коэффициентов арности 5 и более при отборе коэффициентов в множество $a_?^+$. Предлагается следующим образом определить коэффициенты Δ_1 и Δ_0. Коэффициент Δ_1 характеризует разность между количеством единиц и средним количеством знаков (1 и 0) в коде узла: $\Delta_1 = \text{abs}((0{,}5\,c(a) - c_1(a))$, где abs(x) – функция определения абсолютного значения числа x, $c(a)$ – количество знаков в коде узла, $c_1(a)$ – количество единиц в коде узла. Коэффициент Δ_0 характеризует разность между количеством нулей и средним количеством знаков в коде узла: $\Delta_0 = \text{abs}\,[(0{,}5c(a) - c_0(a)]$, где $c_0(a)$ – количество нулей в коде узла. Так как $\Delta_1 = \Delta_0$, то для них можно использовать одно общее обозначение Δ.

Рассмотрим коэффициент a_{11101}. Значения коэффициентов Δ для него равны: $c(a) = c(a_{11101}) = 5$; $c_1(a) = 4$; $c_0(a) = 1$; $\Delta_1 = \text{abs}(5/2 - 4) = 1{,}5$; $\Delta_0 = \text{abs}(5/2 - 1) = 1{,}5$. Введение коэффициента Δ позволяет уточнить обозначение узлов $a_{>1}$ и $a_{>0}$ следующим образом: $a^{\Delta}_{>1}$ и $a^{\Delta}_{>0}$. Предлагается считать, что чем больше значение Δ, т.е., чем больше разница между количеством единиц и нулей в коде, тем выше ассоциативное сходство между параметрами. Таким образом, в множество $a^{+}_{?}$ сначала должны отбираться узлы с большими значениями Δ. Обозначим минимальное значение Δ как $\Delta^{\min}$. Для квантора ассоциации размерности 4 коэффициент Δ = {0, 1}; для квантора ассоциации размерности 5 – Δ = {0,5; 1,5}. Для учета разницы в количестве единиц и нулей в коде предлагается в первую очередь выбирать коэффициенты с $\Delta^{*} > \Delta^{\min}$: $\{a^{+}_{?}\} = \{a^{\Delta^{*}}_{>1}\} \cup \{a^{\Delta^{*}}_{>0}\}$. Таким образом, после определения коэффициентов множества $\{a\}$ с помощью атрибутированного бинарного дерева рассчитываются значения коэффициентов зависимости контролируемых параметров СТК согласно выражения (4).

Рассмотрим практический пример применения предлагаемого метода определения коэффициента зависимости между параметрами.

Пример. Дано два контролируемых параметра. Параметр x_1 описывает номинальное напряжение прибора в диапазоне $10\,\text{В} \pm 1\,\text{В}$, параметр x_2 характеризует силу току прибора в диапазоне $1000\,\text{mA} \pm 50\,\text{mA}$. Необходимо оценить величину ассоциативной зависимости параметров.

Параметр x_1 может принимать следующие значения:

$$x_1^1 = 10\,\text{В} - 1\,\text{В} = 9\,\text{В},\quad x_1^2 = 10\,\text{В},\quad x_1^3 = 10\,\text{В} + 1\,\text{В} = 11\,\text{В}.$$

Параметрам с данными значениями соответствуют три унарных предиката: $\varphi_1^1(\tau)$, $\varphi_1^2(\tau)$, $\varphi_1^3(\tau)$, где τ – номер измерения. Предикат

$\varphi_1^2(\tau)$ принимает значение "истина", когда $x_1 = x_1^2 = 10\,\text{В}$; при $x_1 = x_1^1 = 9\,\text{В}$ или $x_1 = x_1^3 = 11\,\text{В}$ $\varphi_1^2(\tau)$ принимает значение "ложь".

Аналогично для параметра x_2:

$$x_2^1 = 1000\,\text{mA} - 50\,\text{mA} = 950\,\text{mA},$$

$$x_2^2 = 1000\,\text{mA},$$

$$x_2^3 = 1000\,\text{mA} + 50\,\text{mA} = 1050\,\text{mA}.$$

Параметрам с данными значениями соответствуют унарные предикаты: $\varphi_2^1(\tau)$, $\varphi_2^2(\tau)$, $\varphi_2^3(\tau)$. Предикат $\varphi_2^2(\tau)$ принимает значение "истина", когда $x_2 = x_2^2 = 1000\,\text{mA}$; при $x_2 = x_2^1 = 950\,\text{mA}$ или $x_2 = x_2^3 = 1050\,\text{mA}$ $\varphi_2^2(\tau)$ принимает значение "ложь".

Предположим, что было проведено 15 измерений, в результате которых было установлено, что в 5 измерениях $(1)\,\varphi_1^2(\tau) \wedge (1)\,\varphi_2^2(\tau)$; в 4 измерениях $(0)\,\varphi_1^2(\tau) \wedge (0)\,\varphi_2^2(\tau)$; в 2 измерениях $(1)\,\varphi_1^2(\tau) \wedge (0)\,\varphi_2^2(\tau)$; в 4 измерениях $(0)\,\varphi_1^2(\tau) \wedge (1)\,\varphi_2^2(\tau)$.

Таким образом, коэффициенты равны:

$$a_{11} = 5;\ a_{00} = 4;\ a_{10} = 2;\ a_{01} = 4,$$

т.е. $\{a\} = \{a_{11}, a_{00}, a_{10}, a_{01}\} = \{5, 4, 2, 4\}$, $n = 4$, а множество

$$\{a_\Sigma\} = \{a_{=1}\} \cup \{a_{=0}\} \cup \{a_{>1}\} \cup \{a_{>0}\} = \{a_{11}\} \cup \{a_{00}\} \cup \{\varnothing\} \cup \{\varnothing\},$$

где $\varnothing$ – символ пустого множества.

Так как коэффициент $n_\Sigma = 1/2 \cdot n = 2$, то множество узлов $\{a_{=1}\}$, $\{a_{=0}\}$, $\{a_{>1}\}$ и $\{a_{>0}\}$ являются узлами ассоциативного сходства $\{a^+\} = \{a_{11}, a_{00}\}$, множество узлов $\{a_{0=1}\}$ – узлами ассоциативного различия $\{a^-\} = \{a_{10}, a_{01}\}$. Численная мера коэффициента зависимости параметров x_1 и x_2 определяется согласно соотношения (4)

$$\gamma_{\approx_2} = \frac{a_{10} \cdot a_{01}}{a_{11} \cdot a_{00}} = \frac{2 \cdot 4}{5 \cdot 4} = 0{,}4.$$

Значение ассоциативной зависимости $\gamma_{\approx_2} = 0{,}4 < 1$ показывает, что между параметрами x_1 и x_2 существует ассоциативное сходство (зависимость). Таким образом, при определении технического состояния прибора достаточным является измерение или контроль одного параметра.

Выводы. В работе предложен метод определения коэффициентов зависимости КП СТК, основанный на построении и анализе атрибутированного бинарного дерева, который позволяет сократить количество КП за счет контроля только независимых параметров. Предлагаемый подход использования атрибутированного бинарного дерева для автоматического поиска зависимостей между КП представлен впервые. Полученные результаты являются основой для проведения дальнейших исследований по разработке математического обеспечения экспертных систем выбора оптимального состава КП сложных технических комплексов при проведении их метрологического обслуживания.

ЛИТЕРАТУРА

1. *ДСТУ 2682–94. Метрологічне забезпечення. Основні положення. Чинний з 1.01.1995. – К.: Держстандарт України, 1994. – 16 с.*
2. *Чинков В.Н., Герасимов С.В. Комплексная методика оптимизации контролируемых параметров сложных технических объектов // Український метрологічний журнал, 2003. – № 1. – С. 11 – 15.*
3. *Савин С.К. Точность и работоспособность радиоэлектронных систем летательных аппаратов. – М.: Машиностроение, 1988. – 191 с.*
4. *Тульчин Л.Г., Хаскин А.М., Шаповалов В.Д. Оценка качества электроизмерительных приборов. – Л.: Энергоатомиздат, 1982. – 216 с.*
5. *Рубичев Н.А., Фрумкин В.Д. Достоверность допускового контроля. – М.: Изд-во стандартов, 1990. – 171 с.*
6. *Гаек П., Гавранек Т. Автоматическое образование гипотез: математические основы общей теории. – М.: Наука, Главная редакция физико-математической литературы, 1984. – 280 с.*
7. *Amir A., Feldman R., Kashi R. A new and versatile method for association generation // Information systems. – 1997. – Vol. 22, No. 6/7. – P. 333 – 347.*
8. *Петров В.Л., Феклістов А.О., Феклістов О.О. Метод автоматичного утворення гіпотез з квантором багатовимірної асоціації в інтелектуальних системах підтримки прийняття рішень // Системи обробки інформації. – Х.: ХВУ. – 2004. – Вип. 1. – С. 13 – 19.*
9. *Феклистов А.А. Метод определения коэффициентов логической неопределенности в системах поддержки принятия решений // Зб. наук. пр. Інституту проблем моделювання в енергетиці ім. Г.Є.Пухова. – К.: ІПМЕ, 2003. – Вип. 22. – С. 232 – 237.*

Поступила 22.03.2005

Рецензент: доктор технических наук профессор Е.Л. Казаков,
Объединенный научно-исследовательский институт ВС, Харьков.