

УДК 519.248

Ю.А. ДОЛГОВ, Т.Г. ДАНИЛИНА

Придністровський державний університет ім. Т.Г. Шевченка, Молдова

ОПРЕДЕЛЕНИЕ ВИДОВ ЧАСТНЫХ И ОБОБЩЕННЫХ КОЛИЧЕСТВЕННЫХ ОЦЕНОК НАДЕЖНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Статья продолжает серию статей по исследованию количественных характеристик надежности программного обеспечения в соответствии с международным стандартом ISO/IEC 9126. Для получения обобщенных оценок надежности на основе частных оценок метрик методами Харрингтона-Менчера необходимо знать веса этих метрик, которые предлагается находить экспертным методом весовых коэффициентов важности (ВКВ). Помимо повышенной точности метода ВКВ (процедура попарного сравнения объектов с тремя числовыми оценками) он обладает четырьмя степенями защиты: от некомпетентности эксперта, расхождения экспертов по отдельным вопросам, традиционного коэффициента конкордации и соответствия полученной ранжировки законам природы (с помощью закона Ципфа). Приведен алгоритм и дан пример составления экспертной анкеты (опросного листа) и всей процедуры ранжировки.

Ключевые слова: количественная оценка надежности ПО, метод Харрингтона-Менчера, метод весовых коэффициентов важности.

Введение

Экспертные оценки отражают опыт и знания специалистов относительно исследуемого объекта. Суть экспертных методов заключается в том, чтобы используя опыт, знания, интуицию специалистов извлечь из субъективных суждений объективную истину. Существует много разновидностей экспертных методов, но большинство из них могут быть сведены к двум классам: методам прямого ранжирования и методам попарного сравнения. Наилучшими с точки зрения точности выводов являются методы прямого ранжирования, однако они ограничены человеческими возможностями: при числе объектов сравнения 12 – 15 никакой эксперт не в состоянии проранжировать их правильно. Поэтому при большом количестве объектов сравнения прибегают к психологически более комфортным методам попарного сравнения, при котором эксперт отдает предпочтение одному из факторов с точки зрения его влияния на параметр оптимизации. При этом в случае ошибки эксперта неопределенность каждого вывода, если воспользоваться энтропийной оценкой

$$H = -\sum_{i=1}^n P_i \log_2 P_i, \quad (1)$$

составляет 1 бит.

Предлагаемый метод весовых коэффициентов важности (ВКВ) обладает меньшей неопределенностью и более удобен для эксперта с психологической точки зрения, так как эксперту нужно сравни-

вать между собой только 2 фактора.

В случае равновероятного исхода событий, когда эксперт затрудняется отдать предпочтение какому-либо фактору, и с учетом

$$H_1 \in H_0 \cap H_2$$

(исход событий (сравнения двух факторов) может быть $H_0=0$; $H_1=1$; $H_2=2$) из формулы энтропийной оценки меры неопределенности (1) получим [1, 2]

$$H = -\frac{1}{2} \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} \log_2 \left(\frac{1}{2} \right) - \left(-\frac{1}{4} \log_2 \left(\frac{1}{4} \right) \right) = \frac{1}{2}. \quad (2)$$

Это значит, что при использовании метода ВКВ в случае ошибки или незнания эксперта может произойти потеря информации в два раза меньшая, чем при использовании метода парных сравнений. То есть достоверность выводов при использовании метода ВКВ выше, чем при использовании метода парных сравнений.

1. Результаты исследований

Для реализации метода весовых коэффициентов важности необходимо соблюдение таких правил:

1. Опрос экспертов производится только письменно и только в виде заполнения специально разработанной анкеты.

2. Анкета должна состоять из пунктов, в которых сформулированы некоторые утверждения (не вопросы).

3. Пункты анкеты должны быть сформулированы таким образом, чтобы каждый эксперт мог ответить на них однозначно.

4. Отбор экспертов производится исследователем по возможности из разнородных групп.

5. Опрос экспертов производится **индивидуально**.

6. Обработка анкет должна вестись объективными методами. Должны быть некоторые контрольные критерии проверки.

7. После обработки анкет должно быть достаточно убедительное представление результатов.

После составления опросного листа эксперт заполняет экспертную таблицу-матрицу по следующему правилу:

$$a_{ij} = \begin{cases} 2, & \text{если по мнению эксперта} \\ & \text{фактор } i \text{ важнее фактора } j; \\ 1, & \text{если оба фактора одинаковы} \\ & \text{или эксперт не знает, что сказать;} \\ 0, & \text{если фактор } i \text{ уступает фактору } j. \end{cases} \quad (3)$$

Эксперт заполняет только верхнюю треугольную часть матрицы, на диагонали которой стоят единицы, а нижнюю треугольную часть матрицы заполняет эксперт по правилу

$$a_{ij} = 2 - a_{ji}. \quad (4)$$

В конечном виде ранжирование объектов происходит по величине весовых коэффициентов важности k -го порядка [3]

$$b_i(k) = \frac{p_i(k)}{\sum_{i=1}^n p_i(k)}, \quad (5)$$

где $p_i(k)$ – итерированная важность k -го порядка для i -го объекта;

n – число сравниваемых объектов.

Конкретно величины $p_i(k)$ можно найти по следующим формулам:

$$p_i(1) = \sum_{j=1}^n a_{ij}, \quad (6)$$

$$p_i(2) = \sum_{f=1}^n \psi_f \cdot p_f(1); f = \overline{1, n}, \quad (7)$$

где
$$\psi_f = \begin{cases} 2, & \text{если } p_f(1) < p_i(1); \\ 1, & \text{если } p_f(1) = p_i(1); \\ 0, & \text{если } p_f(1) > p_i(1). \end{cases}$$

Практика показала, что условие стабильности ранжирования соблюдается уже при $k = 1$, и всегда при $k = 2$, поэтому считать итерированные важности более высоких порядков нецелесообразно.

Правильность заполнения матрицы и вычисления величин проверяется по следующему равенству:

$$\sum_{i=1}^n p_i(1) = n^2. \quad (8)$$

В отличие от других методов экспертных оценок метод весовых коэффициентов важности позволяет оценить внутреннюю непротиворечивость ответов экспертов. **Коэффициент внутренней непротиворечивости ответов** 1 -го эксперта можно определить по формуле

$$q_1 = \frac{n^3 - \left\{ \sum_{i=1}^n p_i(2) \right\}_1}{\frac{1}{3}(n^3 - n)}. \quad (9)$$

Если величина q_1 меньше некоторого граничного значения, например, $q_{гр} = 0,5$, то мнение такого эксперта не следует учитывать в дальнейших расчетах, так как эксперт сам себе противоречит. В противном случае с мнением эксперта следует считаться.

Известно, что любые выводы, сделанные любым экспертным методом, не могут быть приняты во внимание, если не доказана значимость коэффициента конкордации (согласия экспертов). Однако коэффициент конкордации нельзя искать без предварительной очистки экспертных данных от факторов, мнения по которым резко разошлись, и от мнения тех экспертов, которое по большинству факторов не совпадает с мнением остальных экспертов. При достаточно большом количестве экспертов (более 10) их мнение, выраженное в количественной форме, можно считать распределенным по нормальному закону.

Итак, по данным таблиц ответов экспертов вычисляются весовые коэффициенты важности, которые заносятся в сводную таблицу. Она является основной для вычисления средних величин $b_i(k)$ и дисперсий $S^2\{b_i(k)\}$.

Для выделения факторов, вызывающих непримиримые разногласия экспертов, предлагается воспользоваться критерием Кохрена, при нахождении которого требуется знать только выборочную дисперсию:

$$G = \frac{\max\{S_i^2(k)\}}{\sum_{i=1}^n S_i^2(k)}, \quad (10)$$

где $S_i^2(k) = \frac{1}{m-1} \sum_{i=1}^m [b_{il}(k) - b_i(k)]^2$ – выборочная дисперсия весовых коэффициентов важностей, вычисленная для всех m экспертов по i -му фактору;

$\max\{S_i^2(k)\}$ – максимальное числовое значение

одной из выборочных дисперсий $S_i^2(k)$, вычисленных для всех n исследуемых факторов.

Полученное расчетное значение критерия Кохрена G сравнивается с табличным $G_{\text{табл}}(q; \nu_1; \nu_2)$ для q уровня значимости; ν_1 – число степеней свободы числителя (равное числу экспертов m без единицы); ν_2 – число степеней свободы знаменателя (равное числу ранжируемых объектов n). При $G > G_{\text{табл}}$ фактор, которому принадлежит максимальная дисперсия $\max\{S_i^2(k)\}$, должен быть изъят из дальнейших расчетов и вопрос о его роли должен решаться дополнительным исследованием. При невыполнении неравенства считается, что ни по какому объекту эксперты не высказали противоречивых суждений.

Последней проверкой правильности выводов экспертизы является вычисление коэффициента конкордации (согласия экспертов). Вычисление коэффициента конкордации производится по следующей формуле:

$$W = \frac{\sum_{i=1}^n \left[\sum_{l=1}^m r_{il}(1) - mn \right]^2}{\frac{1}{3} m \left[m(n^3) - n \right] - \sum_{i=1}^n \sum_{l=1}^m (t_{il}^3 - t_{il})}, \quad (11)$$

где t_{il} – число повторений (одинаковых значений) величин $r_{il}(1)$, сделанных l -тым экспертом.

Для проверки значимости коэффициента конкордации формируется критерий χ^2 Пирсона:

$$\chi^2 = m(n-1)W, \quad (12)$$

который сравнивается с табличным значением $\chi^2_{\text{табл}}(q; \nu=n-1)$, и при выполнении условия $\chi^2 > \chi^2_{\text{табл}}$ найденный коэффициент конкордации W признается значимым, то есть считается, что эксперты высказались в основном согласованно, противоречий в их мнениях нет, и полученное ранжирование можно принять за окончательное решение.

2. Практическая часть

В ходе работы был применен метод количественной оценки надежности программного обеспечения на основе обобщенной функции Харрингтона-Менчера, предложенный в [4] и [5]. При этом для определения весов частных показателей качества использовался описанный выше метод ВКВ.

Экспертам было предложено заполнить 4 анкеты. В первой из них (рис. 1) факторами, значимость которых предлагалось сравнить, являются характеристики надежности ПО: завершенность, отказоустойчивость, восстанавливаемость и согласованность надежности (в соответствии с международным стандартом ISO/IEC 9126).

ОПРОСНЫЙ ЛИСТ 1

Надежность

Эксперт заполняет верхнюю треугольную часть матрицы по правилу

$$a_{ij} = \begin{cases} 2, & \text{если фактор } i \text{ важнее фактора } j; \\ 1, & \text{если оба фактора одинаковыми или эксперт не знает, что сказать;} \\ 0, & \text{если фактор } i \text{ уступает фактору } j. \end{cases}$$

Факторы влияния X_i	Факторы, j				$P_i(1)$	$P_i(2)$
	X_1	X_2	X_3	X_4		
X_1 - завершенность;	1					
X_2 - отказоустойчивость;		1				
X_3 - восстанавливаемость;			1			
X_4 - соответствие надежности;				1		
Σ						

Рис. 1. Пример анкеты для заполнения экспертами

Аналогичные анкеты заполнялись для каждой из первых трех характеристик. Факторами для сравнения являлись метрики ПО, определяющие данную характеристику [6]. Для характеристики «соответствие надежности» анкета не составлялась, так как она содержит только одну метрику. Эксперты заполняли анкеты согласно правилу (3). Заполненные анкеты обрабатывались в соответствии с правилом (4), затем рассчитывались важности $P_i(1)$ и $P_i(2)$ по формулам (6) и (7) для каждого фактора (рис.2).

Эксперт заполняет верхнюю треугольную часть матрицы при попарном сравнении объекта i (строка объектом j (столбец) элементами a_{ij}

$$a_{ij} = \begin{cases} 2, & \text{если по мнению эксперта объект } i \text{ предпочтительнее объекта } j; \\ 1, & \text{если оба объекта качественно равны друг другу;} \\ 0, & \text{если объект } i \text{ уступает объекту } j. \end{cases}$$

	1	2	3	4	5	6	7	8	9	10	11
1	1	0	0	0	0	0	0	0	2	4	
2	1	1	0	0	0	0	0	0	2	4	
3	2	2	1	0	0	0	0	0	5	13	
4	2	2	2	1	2	1	1	1	12	32	
5	2	2	2	0	1	2	2	1	12	32	
6	2	2	2	1	0	1	2	0	10	46	
7	2	2	2	1	0	0	1	1	9	27	
8	2	2	2	1	1	2	1	1	12	32	
9									64	370	
10											
11											

Рис.2. Пример заполненной и обработанной анкеты

Затем были отсеяны некомпетентные эксперты путем вычисления коэффициента внутренней непротиворечивости по формуле (9). После дальнейшего применения метода ВКВ с использованием формул (10) – (12) были получены весовые коэффициенты важности для характеристик надежности ПО: для характеристики «Завершенность» – 0, 107; для «Отказоустойчивости» – 0, 562; для «Восстанавливаемости» – 0, 292; для «Соответствия надежности» – 0, 388.

Заключение

С использованием метода ВКВ были определены весовые коэффициенты качества для расчета

количественной оценки надежности программного обеспечения по методике, основанной на обобщенной функции Харрингтона-Менчера [4, 5]. Расчеты проводились для приложений, к надежности функционирования которых не предъявляются повышенные требования.

Литература

1. Гулка З.Н. Энтропийная оценка методов обработки измерительной информации / З.Н. Гулка, Ю.А. Долгов // Информационно-измерительные системы: тез. докл. 7-й Всесоюзной научно-технической конференции. – Винница, 1985. – С. 16-17.
2. Долгов Ю.А. Энтропийная и квазиэнтропийная оценка методов обработки измерительной

информации / Ю.А. Долгов, З.Н. Гулка. – Кишинев, 1987 – 5 с.

3. Долгов Ю.А. Статистическое моделирование / Ю.А. Долгов. – Тирасполь: РИО ПГУ, 2002. – 280 с.

4. Долгов Ю.А. Количественная оценка некоторых характеристик надежности программного обеспечения / Ю.А. Долгов, Т.Г. Данилина // Радиоелектронні і комп'ютерні системи. – Х.:ХАИ, 2006. – № 7 (19). – С. 152-155.

5. Долгов Ю.А. Количественная оценка надежности программного обеспечения / Ю.А. Долгов, Т.Г. Данилина // Радиоелектронні і комп'ютерні системи. – 2007. – № 8 (27). – С. 152-155.

6. ISO/IEC 9126. Software engineering – Software product quality.

Поступила в редакцию 2.02.2009

Рецензент: д-р техн. наук, проф., проф. кафедры Б.М. Конорев, Национальный аэрокосмический университет им. Н.Е. Жуковского “ХАИ”, Харьков.

ВИЗНАЧЕННЯ ВИДІВ ПРИВАТНИХ І УЗАГАЛЬНЕНИХ КІЛЬКІСНИХ ОЦІНОК НАДІЙНОСТІ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Ю.О. Долгов, Т.Г. Даниліна

Стаття продовжує серію статей по дослідженню кількісних характеристик надійності програмного забезпечення відповідно до міжнародного стандарту ISO/IEC 9126. Для здобуття узагальнених оцінок надійності на основі приватних оцінок метрик методами Харрінгтона-Менчера необхідно знати ваги цих метрик, які пропонується знаходити експертним методом вагових коефіцієнтів важливості (ВКВ). Окрім підвищеної точності методу ВКВ (процедура попарного порівняння об'єктів з трьома числовими оцінками) він володіє чотирма мірами захисту: від некомпетентності експерта, розбіжності експертів з окремих питань, традиційного коефіцієнта конкордації і відповідності отриманого ранжування законам природи (за допомогою закону Ципфа). Також описані алгоритм і приклад складання експертної анкети (опитного аркуша) і всієї процедури ранжування.

Ключові слова: кількісна оцінка надійності ПО, метод Харрінгтона-Менчера, метод вагових коефіцієнтів важливості.

DEFINITION OF SORTS OF PRIVATE AND GENERALIZED QUANTITATIVE ESTIMATIONS OF SOFTWARE RELIABILITY

Y.A. Dolgov, T.G. Danilina

The article continues the series of articles on research of quantitative characteristics of reliability of the software according to the international standard ISO/IEC 9126. For reception of the generalized estimations of reliability on the basis of private estimations of metrics methods of Kharrington-Mencher is necessary for knowing weight of these metrics which are offered to be found an expert method weight factors importance (WFI). Besides the raised accurancy of method WFI (procedure of paired matching of objects with three numerical estimations) it possesses four degrees of protection: from incompetence of expert, discrepancy of experts on the individual questions, tradicional factor of concordation and correspondences of the received ranging to nature laws (by means of the law of Tsipf). The algorithm is resulted and the example of compilation of the expert questionnaire and all procedure of ranging is set.

Keywords: quantitative estimation of software reliability, Kharrington-Mencher method, weight factors importance method

Долгов Юрий Александрович – д-р техн. наук, проф., зав. кафедрой информационных технологий и автоматизированного управления производственными процессами, Приднестровский государственный университет им. Т.Г.Шевченко, Тирасполь, Молдова, e-mail: dolax@mail333.com.

Данилина Татьяна Геннадьевна – ст. преп. кафедры информационных технологий и автоматизированного управления производственными процессами, Приднестровский государственный университет им. Т.Г.Шевченко, Тирасполь, Молдова, e-mail: tandan57@mail.ru.