

УДК 004.519.217

С.А. ЯРЕМЧУК

Одесский национальный политехнический университет, Украина

ПРОБЛЕМЫ АПРИОРНОЙ ОЦЕНКИ ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ ГАРАНТОСПОСОБНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ КРИТИЧЕСКОГО НАЗНАЧЕНИЯ

Предложена модель априорной оценки количества дефектов программных средств на основе метрик сложности в виде системы линейных уравнений. При верификации модели была достигнута большая точность оценки количества дефектов по сравнению с известными априорными моделями. Верификация и анализ коэффициентов модели выявили следующие проблемы применения модели: проблема отсутствия решений системы, проблема выбора метрик, проблема неустойчивых оценок, проблема интерпретации коэффициентов модели, проблема возникновения отрицательных коэффициентов и их влияния на точность оценки. Для исследования данных проблем было проведено имитационное моделирование, в результате которого были намечены пути их решения, что необходимо для эффективного использования предложенной модели на этапе разработки программных средств гарантоспособных информационных систем критического назначения. Рассмотрена возможность использования коэффициентов модели в качестве комплексного показателя качества разработки программных средств.

Ключевые слова: гарантоспособность, показатели надежности, априорная оценка, количество дефектов, метрики сложности, коэффициенты дефектосложности.

Введение

Гарантоспособность - это комплексное свойство информационных систем (ИС), включающее в себя понятия надежности и безопасности. Безопасность, определяемая в стандарте [1], как способность ИС достигать приемлемых уровней риска нанесения вреда людям, бизнесу, имуществу или окружающей среде, является важной характеристикой качества в использовании, и тесно связана с надежностью - характеристикой внешнего и внутреннего качества.

Надежность ПС напрямую зависит от количества дефектов.

Рост масштабов и сложности современных ПС приводит к увеличению количества дефектов, которые отрицательно влияют на гарантоспособность ИС.

Исследования, опубликованные в работе [2] показали, что разработчики допускают в среднем от 1 до 3 дефектов в час при проектировании, и от 5 до 8 дефектов в час при кодировании.

Среднее количество дефектов на тысячу строк исходного кода до начала его тестирования колеблется в пределах 10-50 [3]. В то же время оценки только экономических потерь экономики США от некачественных ИС составляют около 60 миллиардов долларов в год.

Снижение гарантоспособности, при одновре-

менном росте критичности функций ИС, резко увеличивает масштабы убытков и потерь от последствий проявления дефектов. Информационные системы управляющие критическими объектами и процессами, наиболее остро нуждаются в точной априорной оценке показателей надежности на ранних этапах разработки, задолго до начала эксплуатации. Точная априорная оценка позволит своевременно привлечь необходимые ресурсы для достижения требуемого уровня гарантоспособности информационных систем

Для количественной оценки надежности ИС в стандартах [4 – 6] определены меры в виде множества метрик. Одними из самых востребованных в практике программной инженерии являются: *прогнозное количество дефектов*, *плотность дефектов* (отношение количества дефектов к размеру ИС), и *степень выявления дефектов* (отношение выявленного при тестировании количества дефектов к прогнозируемому количеству). Эти показатели необходимы для оценки уровня гарантоспособности, и принятия управленческих решений на этапах разработки и эксплуатации ИС. Априорная модель оценки позволяет получить базовый показатель – прогнозное количество дефектов, необходимое для расчета перечисленных показателей.

Поэтому исследования проблемных вопросов использования данной модели, и определение путей их решения является актуальной задачей.

Мотивация исследований

Существующие априорные модели надежности, исследованные автором в работах [7,8], показали недостаточную точность оценки количества дефектов. Отклонения оценок от фактических значений составляли от 44 до 87%, что свидетельствует о необходимости разработки более точной модели. Поэтому автором в работе [7] предложена модель априорной оценки количества дефектов на основе метрик сложности, позволяющая оценить количество дефектов в ИС на этапе разработки, с большей точностью, чем существующие модели. Согласно модели, программная сложность является комплексным показателем, и может быть численно выражена в виде множества значений различных метрик сложности M_1, M_2, \dots, M_n .

Оценка количества дефектов в программном проекте N_d рассчитывается согласно выражению:

$$N_d = a_1 \cdot M_1 + a_2 \cdot M_2 + \dots + a_n \cdot M_n, \quad (1)$$

где a_1, a_2, \dots, a_n – неизвестные коэффициенты, названные коэффициентами дефектосложности (КД). КД определяются решением системы:

$$\begin{cases} a_1 \cdot M_{11} + a_2 \cdot M_{12} + \dots + a_n \cdot M_{1n} = N_{d1}; \\ a_1 \cdot M_{21} + a_2 \cdot M_{22} + \dots + a_n \cdot M_{2n} = N_{d2}; \\ \dots \\ a_1 \cdot M_{n1} + a_2 \cdot M_{n2} + \dots + a_n \cdot M_{nn} = N_{dn}, \end{cases} \quad (2)$$

где $N_{d1}, N_{d2}, \dots, N_{dn}$ – известное количество обнаруженных при тестировании дефектов в ранее созданных ИС данного разработчика, M_{ij} – значение сложности i -го проекта, рассчитанное по j -й метрике.

Полученные в результате решения (2) КД и значения метрик разрабатываемой ИС позволяют согласно (1) рассчитать априорную оценку количества дефектов.

В [7, 8] выполнена верификация модели на экспериментальных данных более двадцати программных систем (ПС) с использованием различных метрик сложности.

Полученная точность априорной оценки количества дефектов характеризуется средним отклонением 10%, которое значительно ниже отклонений оценок существующими априорными моделями.

При верификации модели были получены коэффициентами дефектосложности, представленные в табл. 1, 2.

В процессе верификации модели и анализе полученных КД были сформулированы следующие проблемные вопросы:

Таблица 1

КД, полученные при верификации функционально - ориентированных ПС

ПС	Значения КД для метрик			
	МакКейба	Чепина	Джилба	Сложности вычислений
1	4,09E-02	2,49E-03	1,55E-01	-
2	4,53E-02	-4,18E-04	1,43E-01	-
3	3,62E-02	2,27E-02	-	-1,03E-02
4	3,76E-02	1,67E-02	4,56E-02	-7,26E-03
5	4,55E-02	5,18E-04	1,44E-01	-4,30E-05
6	4,09E-02	1,05E-02	9,98E-02	-6,21E-03
7	3,77E-02	1,12E-02	-	-
8	4,18E-02	-	1,97E-01	-
9	3,92E-02	-	-	1,07E-02

Таблица 2

КД, полученные при верификации объектно-ориентированных ПС

ПС	Метрики	Значения КД	
1	LCOM, WMC	7,61E-02	-1,62E-01
2	LCOM, LOC	6,52E-02	-5,30E-03
3	LCOM, RFC	5,56E-02	-3,07E-02
4	LCOM, LOC	3,85E-02	-9,10E-03
5	LOC, WMC	1,36E-02	-2,93E-01
6	LOC, CBO	1,93E-02	-2,58E-01
7	LOC, RFC	2,17E-02	-1,29E-01
8	LOC, WMC	-4,00E-03	1,78E-01
9	LOC, NPM	-5,40E-03	2,44E-01
10	RFC, CBO	-3,94E-01	1,07E+00
11	RFC, CE	-7,00E-01	3,76E+00
12	RFC, WMC	-5,97E-01	1,64E+00

1) при каких условиях система (2) не имеет решений – *проблема отсутствия решений*;

2) какие метрики следует выбрать для количественного учета сложности и оценки дефектов – *проблема выбора метрик*;

3) при каких условиях оценка количества дефектов становится неточной или недостоверной – *проблема неустойчивых оценок*;

4) какой физический смысл имеют положительные и отрицательные КД – *проблема интерпретации КД*;

5) при каких условиях КД становятся отрицательными – *проблема условий возникновения отрицательных КД*;

6) влияют ли отрицательные КД на точность оценки, и в какой степени – *проблема влияния отрицательных КД на точность оценки*.

Перечисленные проблемы определяют цель исследования. *Целью работы* является исследование проблем использования априорной оценки количества дефектов, и определение путей их решений, что позволит эффективно применять предложенную модель в практике программной инженерии.

1. Проблема отсутствия решений

Важнейшим условием применения модели является существование единственного решения системы (2). Поэтому необходима проверка системы на невырожденность.

Согласно общеизвестным свойствам определителей, если элементы некоторого ряда основной матрицы системы пропорциональны соответствующим элементам параллельного ряда в виде выражения

$$\frac{M_{1k}}{M_{1m}} = \frac{M_{2k}}{M_{2m}} = \dots = \frac{M_{nk}}{M_{nm}} \quad (3)$$

то определитель системы $\Delta = 0$, система является вырожденной, и не имеет решения.

Если для системы (2) выражение (3) не выполняется, то $\Delta \neq 0$, система является невырожденной, и имеет единственное решение в виде искомым КД. Т.о., возможность применения модели определяется условием не пропорциональности векторов значений метрик системы (2), выбранных для количественного учета сложности.

В противном случае необходимо сформировать новую систему (2) на основе других метрик сложности, для значений которых условие коллинеарности векторов (3) не выполняется.

2. Проблема выбора метрик

Выбор множества не коллинеарных метрик, характеризующих различные виды программной сложности, является проблемным вопросом, поскольку на сегодняшний день не существует четких

критериев выбора метрик для количественной оценки сложности различных программных проектов.

Холстед первым в 1976 году предложил использовать метрики для оценки количества дефектов. Впоследствии метрики сложности создавались различными исследователями, в разное время, и для различных целей.

Бессистемность и хаотичность процесса создания метрик привела к появлению их большого количества, при отсутствии методов выбора метрик для оценки много аспектной сложности конкретных программных систем.

Отсутствие количественных и качественных критериев выбора метрик затрудняют практическое применение модели. Неправильный выбор метрик приведет к неполному учету сложности, и снижению точности оценки.

Для решения этой проблемы необходимо создание классификации метрик по видам программной сложности с учетом коллинеарности метрик между собой.

3. Проблема неточных оценок

Предположим, что при некоторых значениях метрик выражение (3) выполняется не точно, а приблизительно, с некоторой точностью (до десятых долей). При этом $|\Delta| \rightarrow 0$, и система *близка к вырождению*. В этом случае мы получим нестабильные КД с большим вариационным размахом значений, что повлечет за собой большие отклонения оценок.

Для проверки этого предположения было проведено имитационное моделирование в широком диапазоне значений различных метрик сложности десяти ПС. В качестве исходных данных выбраны значения метрик реальных ПС и составлена система (2). Далее значение каждой метрики изменялось от 0 до 50 000, при неизменных значениях остальных метрик. При этом вычислялись значения определителей системы, КД различных метрик, и оценок дефектов. В качестве примера в табл. 3 представлены результаты расчетов значений определителей, КД метрики Маккейба (оценивает сложность потока управления программы), метрики Чепина (оценивает сложность потока данных), и значений оценок при фактическом количестве *сорок дефектов*.

Таблица 3

Результаты расчетов определителей, КД и оценок

Значения метрики Чепина	Значение глав определителя системы Δ	КД метрики МакКейба	КД метрика Чепина	Оценка кол-ва дефектов
0	-18840	0,0500	0,0070	39,00
10	-18260	0,0494	0,0072	39,00
50	-15940	0,0465	0,0083	38,95

Окончание табл. 3

Значения метрики Чепина	Значение глав определителя системы Δ	КД метрики МакКейба	КД метрика Чепина	Оценка кол-ва дефектов
80	-14200	0,0438	0,0093	38,91
100	-13040	0,0416	0,0101	38,86
110	-12460	0,0403	0,0106	38,85
120	-11880	0,0389	0,0111	38,83
190	-7820	0,0233	0,0169	38,59
210	-6660	0,0153	0,0198	38,46
230	-5500	0,0040	0,0240	38,28
235	-5210	0,0004	0,0253	38,22
236	-5152	-0,0004	0,0256	38,21
240	-4920	-0,0037	0,0268	38,16
250	-4340	-0,0134	0,0304	38,00
300	-1440	-0,1792	0,0917	35,38
320	-280	-1,2071	0,4714	19,07
321	-222	-1,5405	0,5946	13,78
322	-164	-2,1098	0,8049	4,76
323	-106	-3,3019	1,2453	-14,15
324	-48	-7,3750	2,7500	-78,75
<u>325</u>	<u>10</u>	<u>35,8000</u>	<u>-13,2000</u>	<u>606,00</u>
326	68	5,3235	-1,9412	122,65
327	126	2,9048	-1,0476	84,29
328	184	2,0109	-0,7174	70,11
329	242	1,5455	-0,5455	62,73
330	300	1,2600	-0,4400	58,20
350	1460	0,3137	-0,0904	43,19
400	4360	0,1509	-0,0303	40,61
500	10160	0,1041	-0,0130	39,87
600	15960	0,0914	-0,0083	39,67
700	21760	0,0854	-0,0061	39,57
800	27560	0,0819	-0,0048	39,52
900	33360	0,0797	-0,0040	39,48
1000	39160	0,0781	-0,0034	39,46
2000	97160	0,0726	-0,0014	39,37
3000	155160	0,0713	-0,0009	39,35
4000	213160	0,0706	-0,0006	39,34
10000	561160	0,0696	-0,0002	39,32
20000	1141160	0,0693	-0,0001	39,31
30000	1721160	0,0692	-0,0001	39,31
50000	2881160	0,0691	-0,00005	39,30

Было установлено, что в некотором интервале значений метрик (в табл. 3 это интервал 320-330) определитель Δ имеет минимальных абсолютных значения (-280-300), КД ведут себя нестабильно, их абсолютные значения и знаки резко меняются. В этом интервале точность оценки значительно снижается, и оценка становится недостоверной. Строка таблицы с минимальным значением определителя 10 и оценкой дефектов 606 в $606/40=15$ раз превышает фактическое количество дефектов. После прохождения критического участка значения КД стабилизируются, оценка снова демонстрирует приемлемую точность. Замечено, что нестабильность КД наблюдается всего при 10 (330-320) из 50 000 значений метрики с низкой вероятностью $P=0,0002$.

Визуализация полученных числовых данных представлена на рис. 1, 2. На рис. 1 КД метрики МакКейба обозначен непрерывной линией, КД метрики Чепина обозначен пунктирной линией. Рис. 1 отражает поведение КД на всем диапазоне метрики

Чепина. Нестабильность КД отражены скачкообразными изменениями их значений. Рис. 2 показывает поведение оценок количества дефектов в том же диапазоне метрики. Схожесть двух кривых на разных графиках свидетельствует о большей значимости метрики МакКейба и ее КД (непрерывная линия на рис.1) для оценки дефектов. Метрика Чепина и ее КД (пунктирная линия на рис.1) выполняют уточняющую роль в оценке дефектов. Этот факт свидетельствует о различной значимости метрик для оценки дефектов.

Исследования десяти ПС с использованием различных метрик сложности подтвердили предположение области неустойчивых оценок. Проблема заключается в том, как до решения системы (2) определить эту область, и каковы количественные критерии неустойчивости. Для решения этой проблемы необходимо определить диапазон значений определителя в области неустойчивых оценок, во избежание неудовлетворительных результатов моделирования.

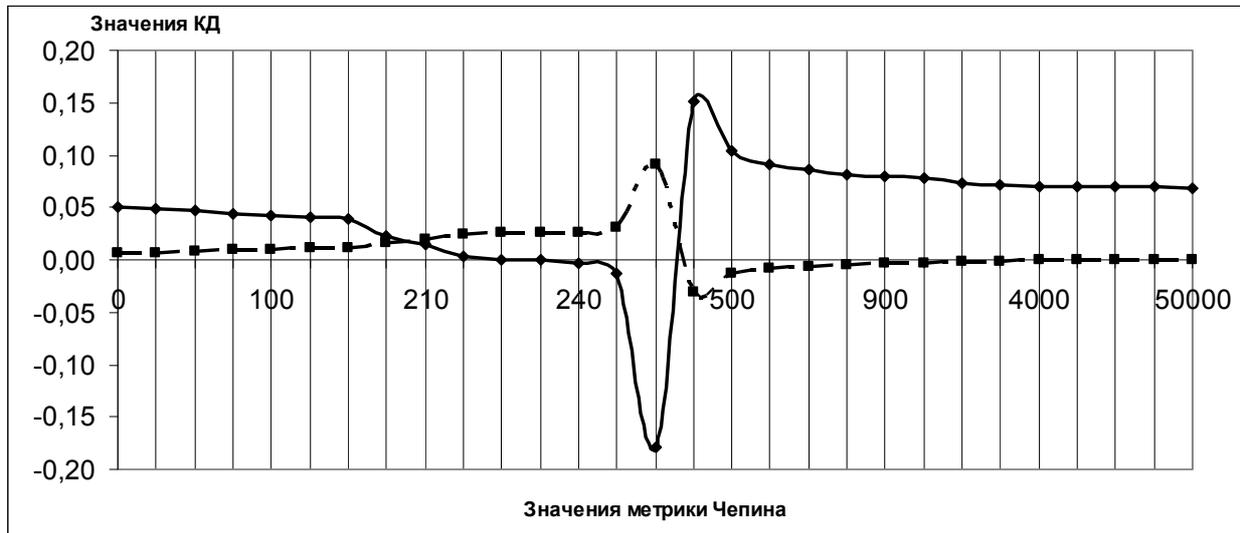


Рис. 1. График значений КД

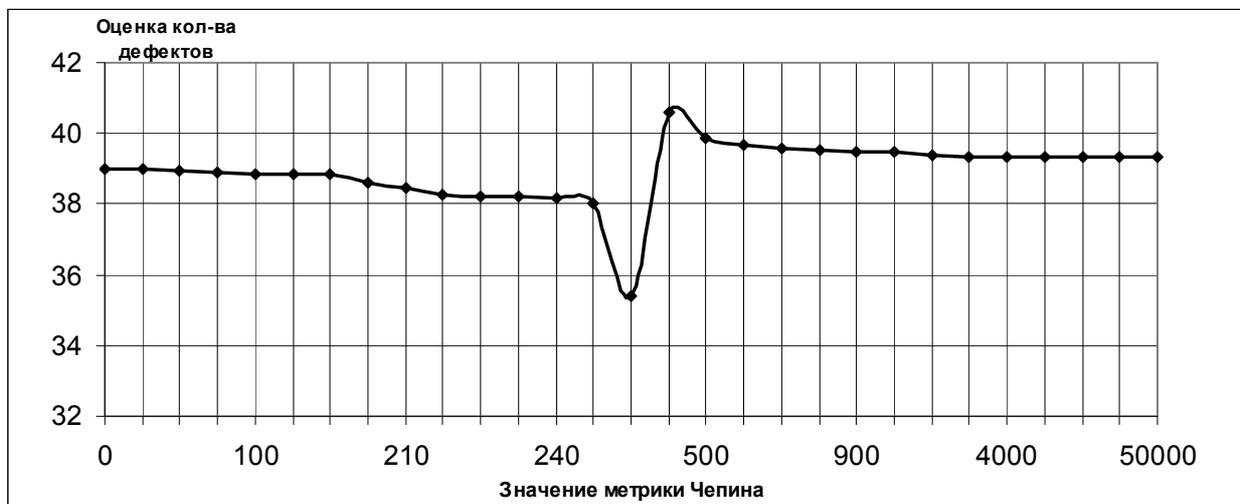


Рис. 2. График оценки количества дефектов

4. Проблема інтерпретації КД

Фізический смысл значений КД определенной метрики сложности будем интерпретировать, как долю дефекта на одну программную структуру, учитываемую данной метрикой. Значения КД различных метрик показывают степень дефектности различных программных структур.

Из табл. 1, 2 видны различные значения КД, что говорит о различной степени дефектности различных программных структур. В этой связи имеет смысл сравнение КД соответствующих метрик для ПС различных разработчиков. Например, КД метрики МакКейба равный 0.05, означает, что разработчик будет ошибаться 50 раз на каждую тысячу единиц данной метрики (количество циклов, условий, переходов программы).

Если для другого разработчика КД этой метрики составит 0,025, это означает двух кратное снижение количества внесенных дефектов в программные структуры, что при неизменности КД других метрик, вызовет уменьшение суммарного количества дефектов в ПС, повышение его надежности и качества.

Фізический смысл отрицательных КД заключается в уменьшении количества дефектов при увеличении значений соответствующих метрик, и объясняется следующими факторами:

1. Разбиение *одной* сложной программной конструкции в исходном коде *на десять* элементарных (например, замена *одного* условия IF со сложным предикатом *десятью* простыми условными переходами DO CASE), может увеличивать значения некоторой метрики, но уменьшить количество дефектов, за счет внесения полной ясности в каждую элементарную структуру, снижения вероятности внесения дефектов в каждую из них, что даст в итоге уменьшение количества дефектов.

2. В языках программирования существуют синонимичные конструкции (например, разворот цикла с небольшим количеством итераций в линейную последовательность команд), которые функционально одинаковые, но имеют различную сложность, и учитываются различными метриками. Например, замена одной конструкции *десятью* другими уменьшит на единицу значение метрики с КД=4,53E-02, при этом увеличит на десять единиц значение метрики с КД=5,18E-04, что в итоге уменьшит количество дефектов.

Проблемным остается вопрос, возможно ли множество значений КД считать новым комплексным показателем качества разработки, который характеризует уровень дефектности определенных программных структур, учитываемых конкретными метриками сложности.

5. Проблема возникновения отрицательных КД

Таблицы 1 – 3 показывают, что КД различных метрик имеют разные знаки. Поскольку КД системы (2) вычисляются выражением

$$a_i = \Delta_i / \Delta, \quad (4)$$

где Δ - главный определитель системы, Δ_i - определитель i -го дополнения системы, знаки КД определяются знаками двух определителей. Эти определители меняют свои знаки не одновременно, поэтому каждый КД на всем множестве значений метрик *меняет свой знак дважды* в виде последовательности 1 – плюс-минус-плюс (3 колонка табл.3), или последовательности 2 - минус-плюс-минус (в других расчетах). Проблемным остается вопрос, как физически интерпретировать факт смены знака КД, и какую информацию предоставляют разработчикам значения метрик в точках смены знаков КД.

6. Проблема влияния отрицательных КД на точность оценки

Экспериментальные результаты таблицы 3 показывают, что оценка количества дефектов имеет достаточную точность как при положительных, так и при разно знаковых КД. Однако необходимы дополнительное исследование и теоретическое обоснование положения о зависимости либо независимости точности априорной оценки количества дефектов от влияния отрицательных КД. Существует ли такая зависимость, и если да, то каков характер этой зависимости.

Выводы

Несмотря на полученные в ходе исследований результаты, остается ряд нерешенных вопросов:

1) необходимо создание *классификации метрик по видам программной сложности* с учетом факта коллинеарности векторов метрик, для обоснованного выбора метрик в случае отсутствия решений системы (2);

2) каковы численные *критерии неустойчивости оценок*, позволяющие определить область неустойчивости до решения системы (2);

3) возможно ли множество значений КД считать новым *комплексным показателем качества разработки*, который характеризует уровень дефектности программных структур;

4) как физически интерпретировать *факт смены знаков КД*;

5) существует ли зависимость точности оценок от знаков КД, и каков характер этой зависимости.

Литература

1. ISO/IEC TR 9126-1:2001, *Software engineering – Product quality. – Part 1: Quality model [Text]*.
2. Humphrey, W. *Introduction to the Personal Software Process [Text]* / W. Humphrey. – MA: Addison-Wesley, 1997. – 324 p.
3. Макконнелл, С. *Совершенный код. Мастер-класс [Текст]: пер. с англ. / С. Макконнелл. – СПб.: Питер, 2005. – 896 с.*
4. ISO/IEC TR 9126-2:2003, *Software engineering – Product quality – Part 2: External metrics [Text]*.
5. ISO/IEC TR 9126-3:2003, *Software engineering – Product quality – Part 3: Internal metrics [Text]*.
6. ISO/IEC TR 9126-4:2004, *Software engineering – Product quality. – Part 4: Quality in use metrics [Text]*.
7. Маевский, Д.А. *Оценка количества дефектов программного обеспечения на основе метрик сложности [Текст]* / Д.А. Маевский, С.А. Яремчук // *Электротехнические и компьютерные системы. – К.: Техника, 2012. – № 7. – С. 113 – 120.*
8. Маевский, Д.А. *Априорная оценка количества дефектов в программном обеспечении информационных систем [Текст]* / Д.А. Маевский, С.А. Яремчук // *Радиоэлектронные и компьютерные системы. – 2012. – № 4(56). – С. 73 – 80.*

Поступила в редакцию 7.02.2013, рассмотрена на редколлегии 6.03.2013

Рецензент: д-р техн. наук, проф., декан факультета В.М. Илюшко, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.

ПРОБЛЕМИ АПРІОРНОЇ ОЦІНКИ ПОКАЗНИКІВ НАДІЙНОСТІ ГАРАНТОЗДАТНИХ ІНФОРМАЦІЙНИХ СИСТЕМ КРИТИЧНОГО ПРИЗНАЧЕННЯ

С.О. Яремчук

Запропоновано модель априорної оцінки кількості дефектів програмних засобів на основі метрик складності у вигляді системи лінійних рівнянь. При верифікації моделі була досягнута більша точність оцінки кількості дефектів в порівнянні з відомими априорними моделями. Верифікація і аналіз коефіцієнтів моделі виявили наступні проблеми застосування моделі: проблема відсутності рішень системи, проблема вибору метрик, проблема нестійких оцінок, проблема інтерпретації коефіцієнтів моделі, проблема виникнення від'ємних коефіцієнтів і їх впливу на точність оцінки. Для дослідження даних проблем було проведено імітаційне моделювання, в результаті якого намічені шляхи їх розв'язання, що необхідно для ефективного використання запропонованої моделі на етапі розробки програмних засобів гарантоздатних інформаційних систем критичного призначення. Розглянута можливість використання коефіцієнтів моделі як комплексний показник якості розробки програмних засобів.

Ключові слова: гарантоздатність, показники надійності, априорна оцінка, кількість дефектів, метрики складності, коефіцієнти дефектоскладності.

PROBLEMS AN APRIORISTIC ESTIMATION OF INDICATORS RELIABILITY OF DEPENDABILITY INFORMATION SYSTEMS CRITICAL APPOINTMENT

S.A. Yaremchuk

The model of an aprioristic estimation of quantity of defects of software on the basis of metrics of complexity in the form of system of the linear equations has been described. Models verification ensures the big accuracy of an estimation of quantity of defects in comparison with known aprioristic models. Verification and the analysis of factors of model have been spent. Following problems of application of model have been revealed: a problem of absence of decisions of system, a problem of a choice of metrics, a problem of unstable estimations, a problem of interpretation of factors of model, a problem of occurrence of negative factors and their influence accuracy of estimation. For research of the given problems has been spent imitating modeling. As a result ways have been planned of the decision of problems. It is necessary for an effective widely adopted of the offered model at a development cycle of software of dependable information systems of critical appointment. Possibility of use of factors of model as a complex qualitative characteristic of software development is considered.

Keywords: dependability, reliability indicators, an aprioristic estimation, quantity of defects, complexity metrics, factors of deficiency and complexity.

Яремчук Светлана Александровна – аспірантка кафедри ТООЭ Одеського національного політехнічного університета, Одеса, Україна; стар. преп. кафедри ІУС Ізмайльського інститута водного транспорту, Ізмаїл, Одеська область, Україна, e-mail: svetlana397@yandex.ru.