



С.Н. Неборский, к.т.н., доцент кафедры программного обеспечения информационных технологий Белорусского государственного университета информатики и радиоэлектроники

Оценка надежности обучающих программных средств

В статье предлагаются новые метрики для оценки надежности обучающих программных средств, а также модификация метода оценки надежности по ГОСТ 28195-99. Предлагаемые метрики расширяют набор существующих метрик стандарта ISO/IEC 25023-2016 для более точной оценки качества по критерию надежности. Предлагаемая модификация метода оценки надежности программных средств позволяет автоматизировать процедуру оценки на основе предложенных метрик надежности обучающих программных средств.

В многообразии классов программных средств (ПС) все большую популярность набирают обучающие ПС. Связано это с ростом требований к образованию и стремлением профильных организаций предоставить лучшие инструменты по освоению предлагаемых дисциплин [1]. Технологически обучающие ПС не отличаются от других классов ПС, значит, и требования к их качеству сопоставимы с теми, что предъявляются к функционально аналогичным ПС.

Действующими на момент написания данной статьи стандартами в области качества ПС являются стандарты ISO/

IEC 250nn SQuaRE (Software product Quality Requirements and Evaluation – требования к качеству ПС и оценка качества ПС) [2]. Стандарт ISO/IEC 25010:2011 [3] предлагает 2 модели качества ПС: модель качества при использовании и модель качества продукта. Модель качества продукта сводит свойства качества к восьми характеристикам: функциональная пригодность, уровень производительности, совместимость, удобство пользования, надежность, защищенность, сопровождаемость и переносимость.

Предметом исследования в данной статье является надежность обучающих ПС. Надежность – это способность ПС поддерживать заданный уровень качества функционирования при его использовании в заданных условиях [4]. Актуальность проблемы надежности для обучающих ПС связана с двумя аспектами:

- ненадежное (в широком смысле) ПС не позволит слушателям курса освоить предлагаемую дисциплину,
- в силу экономических факторов обучающие ПС зачастую разрабатываются за счет внутренних ресурсов, что не позволяет применять современный инструментарий по управлению качеством ПС, а квалификация членов команды разработчиков уступает квалификации промышленных специалистов.

Жесткая, детерминированная, инерционная организация в большинстве образовательных учреждений не позволяет своевременно отслеживать конъюнктуру рынка и удовлетворять все возрастающие требования заказчиков образовательных услуг [5]. В таких условиях проблема обеспечения надежности обучающих ПС выходит на одно из первых мест.

В данной работе предлагается метод оценки надежности ПС и новые метрики оценки надежности, отражающие специфику обучающих ПС. На рисунке 1 отображена модель качества ПС по критерию надежности, предлагаемая стандартом ISO/IEC 25010:2011 [3].

Модель качества стандарта ISO/IEC 25010:2011 [3] ограничивается уровнем подхарактеристик. Конкретные метрики качества приведены в стандарте ISO/IEC 25023:2016 [6].

Завершенность – это степень соответствия ПС при нормальной работе требованиям надежности [7]. Для оценки данной подхарактеристики предлагается использовать следующие метрики: устранение отказов и наработка на отказ.

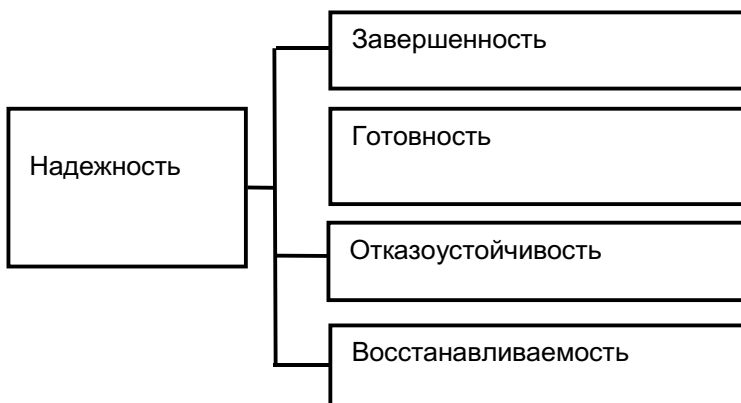


Рис. 1. Модель качества для оценки надежности ПС

В дополнение к ним предлагаются новые метрики: покрытие тестами, глубина обратной совместимости.

Устранение отказов [6]:

$$r_1 = \frac{a_1}{b_1}, \quad (1)$$

где a_1 – количество устраненных отказов ПС; b_1 – общее количество отказов ПС.

Наработка на отказ [6]:

$$r_2 = \frac{T}{N}, \quad (2)$$

где T – время работы ПС (время наблюдения); N – количество отказов ПС за данное время работы ПС.

В данной работе метрику наработки на отказ предлагается модифицировать с целью приведения к относительным единицам и нормирования:

$$r_2 = \begin{cases} 1, & \text{если } a_2 \geq b_2 \\ \frac{a_2}{b_2}, & \text{если } a_2 < b_2 \end{cases}, \quad (3)$$

где a_2 – полученное абсолютное значение наработки на отказ; b_2 – базовое абсолютное значение наработки на отказ.

При приведении абсолютных метрик к относительным в качестве базового значения может быть принят либо уровень худшего случая, либо требуемое значение. Для метрики наработки на отказ в качестве базового значения предлагается использовать требуемое, т. е. минимально допустимое, значение.

Готовность – это степень работоспособности и доступности ПС [7]. Для оценки готовности предлагается использовать такую метрику, как тестовая зрелость. В дополнение к ней предлагается новая метрика: резервное поведение.

Тестовая зрелость [6]:

$$r_3 = \frac{a_3}{b_3}, \quad (4)$$

где a_3 – количество успешно пройденных тестов ПС; b_3 – общее количество тестов ПС.

Отказоустойчивость – это способность ПС работать как предназначено, несмотря на наличие дефектов программного обеспечения или аппаратных средств [7]. Для оценки данной подхарактеристики предлагается использовать такую метрику, как предотвращение некорректных действий пользователя. В дополнение к ней предлагаются новые метрики: возможность отключения функций, однородность входных параметров.

Предотвращение некорректных действий пользователя [6]:

$$r_4 = \frac{a_4}{b_4}, \quad (5)$$

где a_4 – количество тестов, когда отказа ПС не произошло; b_4 – общее количество тестов, направленных на проверку типовых некорректных действий пользователя.

Восстанавливаемость – это способность ПС восстанавливать данные и требуемое состояние системы в случае прерывания или сбоя [7]. Для оценки данной подхарактеристики предлагается использовать такие метрики, как доступность и эффективность восстановления. В дополнение к ней предлагаются новые метрики: степень автоматического восстановления, возможность автоматического оповещения об отказе.

Доступность [6]:

$$r_5 = \frac{a_5}{a_5 + b_5}, \quad (6)$$

где a_5 – время работы ПС до наступления отказа; b_5 – время, необходимое ПС для восстановления работоспособности.

Эффективность восстановления [6]:

$$r_6 = \frac{a_6}{b_6}, \quad (7)$$

где a_6 – количество тестов, нарушивших работоспособность ПС, после которых работоспособность ПС была успешно восстановлена в заданное время; b_6 – общее количество тестов, нарушивших работоспособность ПС.

В таблице 1 приведены новые метрики, предлагаемые в данной работе.

Покрытие тестами:

$$r_7 = \frac{a_7}{b_7}, \quad (8)$$

где a_7 – количество функциональных требований, для которых определены тесты ПС; b_7 – общее количество функциональных требований

Данная метрика показывает, какова доля функций ПС, для которых разработаны тесты. Чем ближе значение данной метрики к 1, тем больше функций ПС покрыто тестами, и, следовательно, тем выше надежность ПС. Следует отметить, что в [6] присутствует одноименная метрика, однако она показывает отношение тестов, которые были выполнены при оценке качества, к общему количеству тестов. По сути данная метрика сводится к метрике тестовой зрелости и в чистом виде не представляет большого интереса при разработке.

Глубина обратной совместимости:

$$r_8 = 1 - \frac{a_8}{b_8}, \quad (9)$$

где a_8 – количество версий ПС, с которыми необходимо обеспечить совместимость текущей версии ПС; b_8 – общее количество выпущенных версий ПС.

Данная метрика показывает, сколько предыдущих версий ПС совместимо с текущей версией ПС. Рациональность состоит в том, что обеспечение обратной совместимости, как правило, связано с усложнением кода ПС. Одновременная поддержка различных версий ПС увеличивает риск допустить

Таблица 1

Предлагаемые метрики надежности ПС

Метрика	Подхарактеристика	Описание
покрытие тестами	завершенность	показывает, какова доля функциональных требований, для которых разработаны тесты
глубина обратной совместимости	завершенность	для заданного набора компонентов показывает, каково число предыдущих версий ПС, которые совместимы с текущей версией ПС
резервное поведение	готовность	определяет возможность настройки приложения на использование предопределенных значений вместо тех, что ведут к отказу ПС
возможность отключения функций	отказоустойчивость	определяет способность ПС автоматически устранять отказы путем отключения отдельных функций, работа которых привела к отказу, сохраняя при этом работоспособность всего ПС
однородность входных параметров	отказоустойчивость	показывает зависимость отказов от входных данных
степень автоматического восстановления	восстанавливаемость	определяет, насколько эффективно ПС может восстанавливать свою работоспособность без участия пользователя
возможность автоматического оповещения об отказе	восстанавливаемость	показывает степень, в которой ПС может автоматически оповещать об отказах доступными средствами информирования

ошибку, что сказывается на надежности ПС. Примером является разработка клиент-серверного решения. Пусть сервер предоставляет определенный набор интерфейсов с заранее известными входными и выходными параметрами. С развитием ПС может возникнуть необходимость в изменении устаревающих интерфейсов, однако для совместимости со старыми клиентами разработчик может принять решение о сохранении старых сигнатур вызовов функций. Чем ближе значение метрики к 1, тем выше надежность ПС.

Резервное поведение:

$$r_9 = \frac{a_9}{b_9}, \quad (10)$$

где a_9 – количество входных параметров ПС, которые могут быть автоматически переопределены при наступлении отказа; b_9 – общее количество входных параметров ПС.

Данная метрика определяет возможность настройки ПС на использование предопределенных значений вместо тех, что ведут к отказу. Разработчик может предусмотреть возможность автоматического переключения поведения системы на набор входных параметров по умолчанию, что исказит результат, но обеспечит отказоустойчивость системы. Например, если рыночная модель валютных курсов не может быть откалибрована по причине арбитражных входных данных, ПС может начать использовать предопределенное значение волатильностей курсов. Чем ближе значение метрики к 1, тем выше надежность ПС.

Возможность отключения функций:

$$r_{10} = \frac{a_{10}}{b_{10}}, \quad (11)$$

где a_{10} – количество функций ПС, которые могут быть отключены автоматически при наступлении отказа; b_{10} – общее количество функций ПС.

Данная метрика определяет способность ПС устранять отказы путем автоматического отключения отдельных функций, работа которых привела к отказу ПС. Примером является поведение Microsoft Excel: при возникновении необработанной ошибки в надстройке (add-in), данная надстройка отключается и не загружается по умолчанию в следующей раз, когда

пользователь запускает приложение. Чем ближе значение метрики к 1, тем выше надежность ПС.

Однородность входных параметров:

$$r_{11} = 1 - \frac{a_{11}}{b_{11}}, \quad (12)$$

где a_{11} – количество групп входных данных, которые могут привести к отказу ПС; b_{11} – общее количество групп входных данных ПС.

Данная метрика показывает зависимость отказов от структуры входных данных. Чем больше данных приходит на вход ПС, тем выше вероятность отказа. Однако надо группировать данные по категориям, чтобы сравнивать одинаковые сущности. Например, в ПС учета сотрудников отдела кадров информация о сотрудниках наподобие имени, адреса, даты рождения является однородной и некорректные данные в одном из полей могут привести к отказу с одинаковой вероятностью. В случае финансовой библиотеки категориями входных данных могут быть курсовые волатильности, кривые процентных ставок, спотовые цены на сырье, при этом каждая группа содержит однородный набор данных (например, спотовые цены на нефть, сахар, зерно). Чем ближе значение метрики к 1, тем выше надежность ПС.

Степень автоматического восстановления:

$$r_{12} = \frac{a_{12}}{b_{12}}, \quad (13)$$

где a_{12} – количество отказов, при которых ПС автоматически (без участия пользователя) восстановило свою работоспособность; b_{12} – общее количество зафиксированных отказов.

Данная метрика позволяет определить, насколько эффективно ПС может восстанавливать свою работоспособность без участия пользователя. Чем ближе значение данной метрики к 1, тем выше надежность ПС.

Возможность автоматического оповещения об отказе:

$$r_{13} = \frac{a_{13}}{b_{13}}, \quad (14)$$

где a_{13} – количество отказов, при наступлении которых произошло автоматическое уведомление пользователей; b_{13} – общее количество зафиксированных отказов.

Данная метрика показывает степень, в которой ПС может автоматически оповещать об отказах доступными средствами информирования, например, с помощью электронного письма или текстового сообщения SMS. Метрика актуальна для неинтерактивных режимов работы ПС либо систем с ограниченными возможностями анализа фактов отказа. Чем ближе значение данной метрики к 1, тем выше надежность ПС.

Для оценки надежности в данной работе предлагается модификация метода оценки качества по ГОСТ 28195-99 [8]. Предлагаемый модифицированный метод предполагает, что в модели качества все метрики приведены к относительным единицам и нормированы с тем, чтобы значения изменялись на отрезке [0; 1].

Оценка надежности выполняется в следующей последовательности.

1. Исходя из специфики разрабатываемого ПС и условий проекта, выбирается набор участвующих в оценке подхарактеристик и метрик надежности. Множество результатов расчета метрик задается как $R = \{r_{ij}, i = \overline{1, I}, j = \overline{1, J}\}$, где i – индекс подхарактеристики, j – индекс метрики в данной подхарактеристике, I – общее количество выбранных подхарактеристик, J – количество метрик надежности по подхарактеристике, J – количество метрик надежности с максимальным количеством выбранных метрик.

2. Для каждой метрики надежности r_{ij} определяется весовой коэффициент v_{ij} . Если некоторая метрика в оценке не участвует, ее весовой коэффициент устанавливается в 0.

3. Производится оценка каждой подхарактеристики надежности P_i . Для этого используется следующая формула:

$$P_i = \frac{\sum_{j=1}^J (r_{ij} \cdot v_{ij})}{\sum_{j=1}^J v_{ij}}. \quad (15)$$

4. Для каждой подхарактеристики надежности определяется весовой коэффициент V_i . Если некоторая подхарактеристика в оценке не участвует, ее весовой коэффициент устанавливается в 0.

5. Рассчитывается интегральная оценка надежности P :

$$P = \frac{\sum_{i=1}^I (P_i \cdot V_i)}{\sum_{i=1}^I V_i} \quad (16)$$

Интегральная оценка надежности P принимает значения на отрезке $[0; 1]$, причем чем она ближе к 1, тем выше надежность ПС. Величина P сравнивается с экспертной (базовой) оценкой надежности, заданной в спецификации, и на основе разницы между полученной оценкой и требуемой оценкой принимается решение о том, в какой степени соответствует реализованное ПС установленному уровню качества по критерию надежности.

Весовые коэффициенты метрик и подхарактеристик качества вычисляются экспертным путем, по статистическим данным организации, исходя из требований заказчика.

Подводя итог, в данной работе предложены новые метрики для оценки надежности обучающихся ПС: покрытие тестами (8), глубина обратной совместимости (9), резервное поведение (10), возможность отключения функций (11), однородность входных параметров (12), степень автоматического восстановления (13) и возможность автоматического оповещения об отказе (14). Данные метрики просты в использовании, а их применение позволяет более точно оценить качество обучающихся ПС по критерию надежности. Помимо новых метрик в данной статье предложена модификация метода оценки надежности, преимуществами которого являются возможность автоматизации процедуры оценки на основе обозначенных метрик надежности, а также способность накапливать статистические данные о свойствах ПС. Предложенные новые метрики надежности обучающихся ПС в совокупности с методом оценки надежности позволяет получить продукт более высокого качества.

Литература

1. Захарова И.Г. Информационные технологии обучения и развитие учебных навыков // Открытое образование. – 2002. – №1. – С.24-30.
2. СТБ ISO/IEC 25000-2009: Разработка программного обеспечения. Требования к качеству и оценка программного продукта (SQuaRE). Руководство по SQuaRE.
3. ISO/IEC 25010-2011: Systems and software engineering – Systems and software product Quality Requirements and Evaluation (SQuaRE) – System and software quality models.
4. СТБ ИСО/МЭК 9126-2003: Информационные технологии. Оценка программной продукции. Характеристики качества и руководства по их применению.
5. Голенков, В.В. Виртуальные кафедры и интеллектуальные обучающие системы / В.В. Голенков, В.В. Емельянов, В.Б. Тарасов. – Новости ИИ. – 2001. – №4. – С. 1-19.
6. ISO/IEC 25023-2016: Systems and software engineering – Systems and software product Quality Requirements and Evaluation (SQuaRE) – Measurement of system and software product quality.
7. ГОСТ Р ИСО/МЭК 2010-2015: Информационные технологии. Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов.
8. ГОСТ 28195-99. Оценка качества программных средств. Общие положения.

S.N. Neborsky Evaluation of Reliability of Educational Software

This paper proposes both new metrics to evaluate reliability of educational software, and a modified method of software reliability evaluation under the standard 28195-99. The proposed metrics extend the existing framework of metrics under ISO/IEC 25023-2016. They cater for more exact evaluation of the reliability of educational software. The modified software assessment method proposed in this paper allows automating of the reliability evaluation.

Статья поступила 25.09.2016

