

АНАЛИЗ СУЩЕСТВУЮЩИХ МЕТОДИК ОЦЕНКИ КАЧЕСТВА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

И.И. Дегтярева, Л.П. Володько

Полесский государственный университет, innaid@rambler.ru, lyudvik@tut.by

В настоящее время рынок программных продуктов весьма обширен и наполнен огромным количеством программ, сходных по своему назначению и заявленным продавцами функциям, но не по качеству реализации этих функций. Покупателю программного продукта очень сложно сделать выбор, ведь он может основываться только на заявленном в документации перечне функций программы и требованиях к техническим средствам, необходимым для ее функционирования, а никак не на количественной оценке качества программного продукта и сравнении экономической эффективности имеющихся на рынке программных средств. Если бы подобная оценка была возможной, то совершалось бы значительно меньше ошибок при покупке программного обеспечения (ПО). В то же время неудачный выбор программы может принести фирме-покупателю убытки, которые обусловлены:

- потерями ожидаемого эффекта от купленного ПО;
- затратами денежных средств и времени на приобретение программного продукта (ПП), оказавшегося неэффективным, на его освоение и внедрение, включая наполнение баз данных, справочников и т. д.;
- дополнительными текущими расходами на эксплуатацию ПО;
- потерями темпа в конкурентной борьбе при отказе от уже купленного ПО средства и повторными расходами времени и средств на приобретение, освоение и внедрение нового ПО того же назначения.

В настоящее время существует несколько методик оценки качества программного обеспечения, позволяющего избежать подобной ситуации. Эти методики можно разделить следующим образом.

1. Оценка качества ПО по критерию функциональной полноты.
2. Экспертная оценка качества ПО.
3. Реляционный подход к измерению и оценке качества ПО.
4. Оценка факторов, влияющих на время выполнения ПО отдельных функциональных операций, и работы программы в целом.

Рассмотрим подробнее данные методики.

Оценка качества программного обеспечения по критерию функциональной полноты проводится следующим образом.

Для осуществления оптимального выбора программного обеспечения необходимо располагать количественной оценкой того, в какой степени программы соответствуют требованиям пользователя. Кроме того, нужно определить, какие из необходимых пользователю функций не реализованы в программном продукте. Также важно выявить системы, лучшие по критерию функциональной полноты, и определить перечень функций, реализуемых всеми представленными на рынке программными продуктами. Ситуация осложняется тем, что в современных программных продуктах количество выполняемых ими функций может достигать нескольких сотен и более. Очевидно, что при таких условиях вручную оценить степень соответствия программного продукта требованиям пользователя не представляется возможным. Поэтому для решения поставленных задач был предложен формализованный подход, описанный в [1].

Эта методика позволяет с наименьшими затратами финансовых ресурсов и интеллектуальных усилий осуществить следующее:

- составить полный перечень функций, реализуемых всем представленным на рынке ПО;
- систематизировать сведения о составе и функциональной полноте существующего ПО;
- количественно оценить степень соответствия того или иного ПО требованиям пользователя к функциональной полноте;
- проранжировать ПО по критерию функциональной полноты;
- на стадии предварительного анализа исключить из дальнейшего рассмотрения информационные и программные продукты, в которых не реализуются нужные пользователю функции;
- сформировать группу ПО, имеющего одинаковую функциональную полноту, сопоставить их цены и другие характеристики;
- расширить для потребителя-пользователя возможности оптимального выбора на рынке программного обеспечения, предоставив перечень выполняемых каждым ПП функций, а разработчику показать место его продукта среди существующих программных систем и одновременно дать первоначальную оценку конкурентным рыночным позициям фирм-разработчиков ПП.

Однако с помощью этой методики нельзя оценить другие показатели качества ПО, например, такие как завершенность разработки проекта, быстрое действие программной системы, степень и простота настройки программы на техническую среду, возможность работы в сети, трудоемкость освоения и внедрения программного продукта и прочие. Таким образом, даже удостоверившись в том, что программный продукт полностью удовлетворяет пользователя по составу функций, нельзя утверждать, что этот программный продукт удовлетворит пользователя по своему качеству. Также с точки зрения цели данного исследования, к отрицательным характеристикам применения данной методики можно отнести тот момент, что с ее помощью можно определить наличие функции в программном пакете, но не качество ее реализации. Кроме того, эта методика не предоставляет возможности определить, как и какие факторы влияют на функционирование программного пакета. Самой же главной проблемой является быстрое устаревание программных продуктов, появление новых версий программ с большим количеством новых функций, что вызывает необходимость проводить исследования рынка программного обеспечения с большой частотой. Проанализировав все положительные и отрицательные стороны применения этой методики, можно утверждать, что она не в состоянии служить основой для создания универсальной автоматизированной системы оценки потребительского качества программного обеспечения, но может являться ценным ее дополнением.

Другая методика оценки качества программного обеспечения изложена в [2] – это методика экспертных оценок.

В основе метода групповой экспертизы лежат, как известно, следующие постулаты:

- экспертная оценка носит вероятностный характер;

- считается, что истинное значение исследуемой характеристики находится внутри диапазона оценок отдельных экспертов, т.е. обобщенное коллективное мнение более достоверно (исключение составляет случай, когда группа экспертов существенно неоднородна по уровню компетентности);

- методика отбора экспертов, процедура общения с ними, методы обработки полученных экспертных оценок, зачастую, оказывают решающее влияние на результаты экспертизы. При этом нельзя допускать попыток упростить установленный порядок проведения экспертизы, сократить число туров опросов, изменить содержание анкет, разработанных в соответствии с принятой методикой экспертного опроса.

Применение этой методики на практике вызывает определенные трудности. Если выполняется условие транзитивности, то для упорядочения ПО и оценки согласованности результатов экспертизы можно использовать методы ранговой корреляции. Однако трудно рассчитывать на то, что найдется хотя бы один эксперт, обладающий достаточной компетентностью, чтобы сделать заключение обо всех представленных на рынке ИП. Поэтому в лучшем случае речь может идти о сопоставлении в процессе экспертизы лишь нескольких сложных программных систем. Кроме того, достаточно очевидно, что найти эксперта, хорошо знающего всего две-три системы, тоже значительно проще. Несомненно и другое: дать обоснованную количественную оценку качеству сложной программной системы можно, только имея определенный опыт работы с ней. При этом следует учитывать, что на достоверность результатов сравнения наиболее существенное влияние оказывают факторы, характеризующие индивидуальные свойства эксперта: компетентность (степень пересечения тезауруса эксперта с содержанием исследуемой предметной области, решаемой задачи), природные, генетические способности (степень развитости ассоциативного мышления, глубина памяти, скорость реакции и др.). Причем на результатах экспертизы могут отразиться и такие характеристики эксперта, как общая эрудиция, продолжительность работы с конкретной программной системой и наличие опыта работы с другими аналогичными системами, характер задач, решавшихся с использованием данной программы и другие. Будет сказываться и так называемый «эффект очередности», заключающийся в том, что оценка, даваемая экспертом, зависит от того, в какой последовательности испытывались (осваивались) сравниваемые программные системы.

Следует отметить, что сравнение информационных и программных систем по многим из показателей качества, таким, как качество интерфейса, документации, удобство сопровождения и др., обычно проводится только на основании результатов экспертных опросов, хотя и в этом случае экспертная информация может быть дополнена объективными данными натурального эксперимента, представленными в шкале интервалов или отношений.

Следующий, реляционный подход к оценке качества программных продуктов представляет собой философию, охватывающую разработку и применение методов и средств измерения и оценки качества программного обеспечения и базируется на трех основных принципах:

- система базовых, внутренних свойств любого ИП строится на основе анализа лексических и семантических свойств символов и конструкций каждого языка описания этого продукта. Результатом анализа для каждого языка является система свойств, в терминах которых может быть определено любое внутреннее свойство ИП, написанного на данном языке в рамках конкретной его реализации;

- результаты любых: прямых и непрямых измерений внутренних и внешних свойств ИП представляются в виде кортежей значений мер этих свойств;

- непрямым образом измеряемые внутренние свойства ИП, все его внешние свойства, рейтинги метрик и критерии оценивания продуктов определяются формально. По существу, это формальное описание знаний в прикладной области «качество ПО».

Следует отметить, что реляционный подход к измерению и оценке качества программного обеспечения ориентирован не на оценку качества реализации функций программного продукта, а на оценку качества кода программы. Поэтому использование этого подхода для создания автоматизированной системы оценки именно потребительского качества ПО не представляется целесообразным.

Четвертая методика оценки качества программного обеспечения основана на оценке степени влияния различных факторов на время работы программы. Эта методика основана на проведении активных экспериментов с ИП. Часто время реакции системы на запрос пользователя становится критическим фактором, очень сильно снижающим качество функционирования программной системы. Программа, успешно прошедшая тестирование даже на реальных данных заказчика через

какое-то время может существенно замедлить свою работу, что может быть связано с банальным увеличением количества записей в таблицах базы данных. Указанная методика основана на проведении экстремальных экспериментов с оцениваемой программной системой и выяснении степени влияния на качество работы программы таких факторов, как количество записей в основной (рабочей) таблице (таблицах) базы данных, количество записей в каждом из используемых справочников и так далее.

Применение этой методики позволяет:

- выявить и систематизировать важнейшие особенности информационных систем описать свойства факторов, влияющих на выходные характеристики информационных систем;
- определить основные требования к свойствам моделей, формируемым по результатам активных экспериментов;
- выделить критерии плана эксперимента при исследовании информационных систем.

Нельзя не отметить, что проведение активных экспериментов дает совершенно правдивую и, самое важное, объективную информацию о выходных характеристиках работы оцениваемой программной системы. Кроме того, при условии автоматизации проведения активных экспериментов с программным обеспечением время оценки качества функционирования программы будет на порядок меньше, чем проведение оценки качества программы по любой другой методике.

Для оценки потребительского качества программного продукта с помощью активного экспериментирования достаточно располагать только той информацией, которая содержится в документации, предоставляемой разработчиком ПП и нет необходимости предварительно долго изучать программу, как это, например, нужно при использовании экспертной методики оценки качества ПП; также не нужны исходные тексты программы как это необходимо для оценки качества программного обеспечения при реляционном подходе.

Каждая из описанных выше методик обладает своими особенностями, положительными и отрицательными свойствами. Без применения таких методик, как оценка качества программного обеспечения по критерию функциональной полноты и методика экспертных оценок невозможно получить полную оценку качества программной системы, в связи с уникальностью данных методик, и оценить программную систему по некоторым параметрам можно только с их использованием.

Литература:

1. Хубаев Г.Н. Информационные и программные системы как объекты активного экспериментирования // Программные продукты и системы. – 1999. – № 2. – С. 2–7.
2. Липаев В.В. Методы обеспечения качества крупномасштабных программных средств / В.В. Липаев. – М.: СИНЕГ, 2003. – 520 с.