

МЕТРОЛОГИЧЕСКАЯ АТТЕСТАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ СРЕДСТВ ИЗМЕРЕНИЙ (СОСТОЯНИЕ И ПЕРСПЕКТИВЫ)

Ю.А. Кудеяров, Ю.Е. Лукашов, А.А. Сатановский

Практически в любом современном средстве измерений в том или ином виде присутствует программное обеспечение (ПО). Оно может быть встроенным, установленным на компьютер, входящий в состав измерительной системы, автономным. Использование ПО для обработки измерительной информации является неотъемлемой частью получения конечных результатов измерений. Однако использование ПО, без определения его точностных характеристик сопряжено с риском получения недостоверных результатов.

Если обратиться к состоянию вопроса об метрологической аттестации ПО, то здесь оказывается уже сделаны первые шаги по выработке подходов к решению этой задачи. В первую очередь можно отметить программу SSFM 1и 2 (Software Support For Metrology) Национальная физическая лаборатория (NPL, Великобритания) (см., например, [1, 5-7]. Определенные наработки по этому вопросу имеются и у Национального института по стандартизации и технологиям (NIST, США). Некоторые нормативные разработки имеются и в России [2-4].

Прежде чем переходить к обзору подходов к аттестации ПО, представленных в указанных публикациях, необходимо определить само понятие аттестации ПО и в первую очередь понятие его метрологической аттестации.

В этой работе под метрологической аттестацией ПО будет пониматься признание метрологической службой законности и правомерности его применения, основанное на исследовании степени влияния ПО на метрологические характеристики средств измерений и измерительных систем, в которых это программное обеспечение используется. В Директиве по измерительным приборам [8] степень влияния ПО на метрологические характеристики средств измерений характеризуется параметром, называемым **критичность**. Под критичностью при этом понимается полная либо частичная ответственность ПО за метрологические характеристики приборов и систем. К дальнейшему обсуждению этого параметра мы вернемся ниже.

Любая процедура аттестации основана на установлении соответствия аттестуемого объекта нормативным или техническим требованиям. При этом необходимо отметить, что согласно международному руководству [9] метрологические требования должны предъявляться не ко всему ПО, а только к тем его частям, которые контролируются законодательной метрологией. В качестве таких частей могут быть, например, подпрограммы считывания результатов измерений с датчиков, обработки результатов измерений, отображения этих результатов и т.п. Из сказанного следует, что и метрологической аттестации должны подлежать только те части ПО, которые контролируются законодательной метрологией.

Из всех имеющихся рекомендаций и руководств наиболее полно подходы к аттестации ПО представлены в работах NPL [1,5-7]. Эти документы нацелены на метрологическое ПО, т.е. на ПО, используемое при обработке измерительной информации. Авторы разделили работу [5] для удобства читателей на две части. Первая часть ориентирована на пользователей ПО и рассматривает следующие основные вопросы:

- построение модели измерительной системы, использующей программные средства;
- оценка рисков использования ПО;
- оценка уровня целостности ПО.

Вторая часть представляет интерес, прежде всего, для разработчиков ПО и посвящена, в основном, рассмотрению методик аттестации ПО. Следует подчеркнуть, что авторы работы из NPL придерживаются мнения, что аттестация ПО является частью жизненного цикла ПО и должна осуществляться разработчиками программного продукта.

В упомянутом документе употреблен термин «целостность ПО», который требует пояснения.

Если отвлечься от деталей, то под целостностью ПО авторы понимают идентичность ПО правильной исходной версии (например, утвержденной), т.е. ПО не должно быть изменено преднамеренно или неумышленно.

В Национальной физической лаборатории считают, что самым эффективным методом аттестации ПО, используемого в средствах измерений, является метод «черного ящика» (black box testing). В этом методе аттестация ПО происходит на основе сопоставления результатов обработки так

называемых «эталонных» данных самим тестируемым ПО и «эталонных» результатов, полученных обработкой тех же данных «эталонным» ПО. Проблема, однако, заключается в том, что при таком подходе необходимо иметь в своем распоряжении «эталонное» ПО, т.е. ПО, записанное с использованием программных средств, отвечающих предельно высоким стандартам. Далеко не все испытательные центры таким ПО располагают. Как показывает практика, решение может быть найдено на пути использования так называемых генераторов «эталонных» данных. Условная схема использования таких генераторов изображена на рисунке. Основная особенность схемы заключается в том, что на ее входе не «эталонные» данные, а «эталонные» результаты. Утверждается, что в этом случае усилия, необходимые для разработки генераторов «эталонных» данных составляют ничтожную часть усилий, затрачиваемых на разработку «эталонного» ПО. Кроме того, такие генераторы легко тестируются. Совокупность «эталонных» данных и «эталонных» результатов в работах NPL называют «эталонными» парами.

Со схемой генерации «эталонных» пар можно работать только в том случае, когда математическое решение соответствующей метрологической задачи *априори* известно. К счастью, большинство метрологических задач относится именно к этому классу.

На самом деле даже при известной математической постановке задачи проблема генерации «эталонных» данных по известным «эталонным» результатам не является однозначной, т.е. в результате такой генерации можно получить целые семейства или классы «эталонных» данных, удовлетворяющим условиям математической задачи. Не вдаваясь в математические тонкости, отметим, что, в принципе, для решения задачи аттестации ПО можно работать с любым полученным набором «эталонных» данных при условии, что они соответствуют исследуемой измерительной задаче.

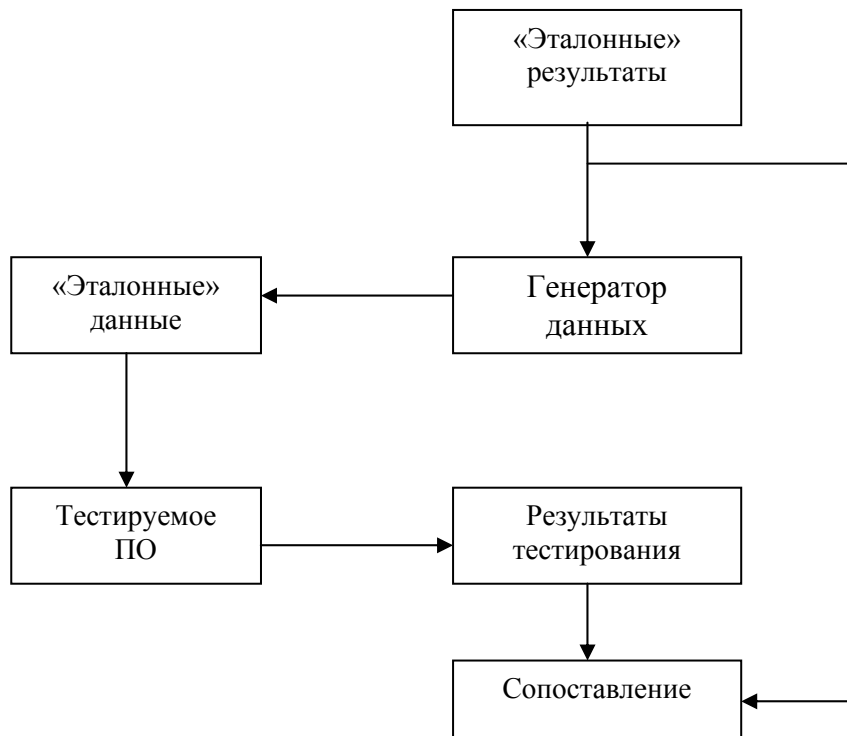


Схема испытаний ПО с использованием генерации «эталонных» данных

В работах NPL аттестацию ПО методом «черного ящика» предлагается проводить в шесть этапов:

1. Спецификация испытываемого ПО.

Спецификация может быть проведена в математических или нематематических терминах, но в каждом случае соответствующая информация должна быть достаточной для избежания неоднозначных толкований. Цель такой спецификации заключается в том, чтобы убедиться в самых общих чертах в адекватности используемых моделей и алгоритмов решаемой задаче, при этом один из способов спецификации заключается в идентификации входа и выхода программного продукта и функциональных соотношений или моделей, связывающих вход и выход.

В качестве примера спецификации испытываемого программного продукта рассмотрим задачу нахождения программным способом по методу наименьших квадратов (МНК) гауссовой кривой на массиве экспериментальных данных $y_i, i = 1, 2, \dots, m$, каждое из которых отвечает экспериментально найденной абсциссе из набора $x_i, i = 1, 2, \dots, m$. [13]. В терминах входной/выходной модели спецификация выглядит следующим образом:

А). Вход. Значения абсцисс x_i , $i = 1, 2, \dots, m$ и соответствующие значения ординат y_i , $i = 1, 2, \dots, m$.

Б) Выход. Значения параметров A , \bar{x} , s в модели гауссового пика

$$y(x) = A \cdot \exp\left\{-\frac{(x - \bar{x})^2}{2s^2}\right\},$$

а также значения остаточных величин e_i , $i = 1, 2, \dots, m$, где $e_i \equiv y_i - y(x_i)$.

В) Входная/выходная модель. Выходные значения определяются из входных решением следующей нелинейной задачи по МНК

$$\min_{A, \bar{x}, s} \sum_{i=1}^m [y_i - y(x_i)]^2.$$

С использованием выписанного выше функционала метода максимального правдоподобия, частным случаем которого является метод наименьших квадратов, определяются также оценки остаточных членов.

2. Исполнение (реализация) испытываемого ПО.

Первые два этапа обычно выполняются на стадии разработки ПО.

3. Спецификация наборов «эталонных» данных.

Эти последовательности наборов данных в большинстве своем определяются значениями управляющих параметров, которые должны быть идентифицированы, при этом каждому значению управляющего параметра отвечает своя последовательность наборов данных. В процессе испытаний ПО исследуется его функционирование для разных последовательностей наборов данных.

Часто управляющему параметру придаются функции контрольного параметра, или параметра, который ранжирует последовательности наборов данных по степени их трудности для обработки испытываемым ПО. Использование таких ранжированных в указанном смысле наборов данных позволяет выявить случаи, когда тестируемое ПО основано на неверном выборе математического алгоритма. К «эталонным» наборам данных часто предъявляется требование максимального соответствия наборам, которые могут быть получены при практических измерениях на базовом средстве

измерений. Спецификация наборов «эталонных» данных должна устанавливать наличие перечисленных требований.

4. Спецификация выполняемых измерений и требований испытаний.

При испытаниях ПО необходимо пользоваться показателями качества (или, как говорят, метриками качества) для установления количественных различий между обработками, производимыми испытываемым ПО над разными наборами «эталонных» данных. Посредством спецификации требований к ПО и пользователь, и разработчик испытывают программный продукт в терминах этих метрик, что в конечном итоге дает возможность объективной оценки ПО в плане соответствия цели, которой оно должно достигать.

Для иллюстрации вернемся к примеру, описанному в п.1. Для оценки качества ПО предлагается использовать два показателя качества (две метрики). Первый основан на оценке коэффициента обусловленности [1] алгоритма вычисления остаточных погрешностей, используемого тестируемым ПО при обработке «эталонных» остатков. Этот коэффициент для рассматриваемого случая имеет вид

$$P(\vec{e}) = \log_{10} \left(1 + \frac{\|\vec{e}_t - \vec{e}_r\|}{\|\vec{y}\| \eta} \right),$$

где \vec{e}_t - вектор остаточных погрешностей, полученный испытываемым ПО, \vec{e}_r - вектор «эталонных» остатков, \vec{y} - вектор значений «эталонных» данных и η - численная точность генератора «эталонных» результатов. По сути, коэффициент P характеризует потерю точности испытываемого алгоритма по сравнению с алгоритмом, обладающим оптимальной стабильностью. Аналогичные коэффициенты обусловленности могут быть вычислены и по отношению к другим рассчитываемым величинам таким как A , \bar{x} , s и т.п.

Второй показатель качества алгоритма имеет прямое отношение к числу N итераций, необходимых исследуемому ПО для обеспечения сходимости решения (понятно, что с $N = \infty$ в пределах максимального числа принимаемых в расчет итераций никакой сходимости не наблюдается).

Более детальное описание соответствующих метрик можно найти, например, в работе [6].

5. Генерация «эталонных» пар.

Уже говорилось, что «эталонные» пары, т.е. совокупность наборов «эталонных» данных и соответствующих «эталонных» результатов, могут быть получены, в частности, методом генерации «эталонных» данных. Здесь нет возможности подробно останавливаться на генерации этих пар. Методы такой генерации «эталонной» пары для вычисления СКО рассмотрены в [7], а метод генерации применительно к методу наименьших квадратов – в [6].

6. Представление и интерпретация исполненных измерений и испытаний.

Применение испытываемого ПО к «эталонным» данным приводит, как это видно из рисунка, к появлению результатов тестирования, которые сравниваются с «эталонными» результатами, из которых был генерирован этот «эталонный» набор. Сравнение, происходит с помощью показателя (метрики) качества, о которых говорилось выше и которые при этом вычисляются. Параметры (метрики) качества представляются как функции управляющего параметра (или параметров) в табличной и/или в графической форме. Удобно также пользоваться статистикой испытаний (численной или графической), включающей в себя значения математического ожидания, СКО, минимума, максимума обнаруженных различий.

Значения вычисленных параметров (метрик) качества сравниваются с результатами испытаний, выраженными в терминах этих метрик. Тем самым происходит выработка объективного суждения о качестве испытываемого ПО и его пригодности для решения конкретной измерительной задачи.

Таков в общих чертах подход к аттестации ПО, предлагаемый специалистами NPL. Нам представляется, что этот подход в основной своей части может служить основой для разработки соответствующих методик аттестации ПО у нас в России.

Для интересующихся вопросами аттестации ПО следует обратить внимание на Web – ресурсы, посвященные вопросам аттестации научного программного обеспечения. Этот ресурс представлен в основном 2-мя сайтами. Это сайт NPL, точнее его подраздел, посвященный программе SSFM-2, а также подраздел сайта NIST, посвященный информационным технологиям.

а) <http://www.npl.co.uk/ssfm>. Данный ресурс посвящен уже упомянутому проекту Software Support for Metrology (SSFM). Кроме того, этот ресурс содержит практически самостоятельный подраздел, а именно, сайт программы METROS (Metrology Software) - <http://www.npl.co.uk/ssfm/metros>. Система

METROS представляет собой библиотеку ключевых математических функций и алгоритмов, используемых в метрологии и разбитых по областям измерений, в которых они находят наибольшее применение. На сайте также находятся описания алгоритмов, реализующих эти функции.

b) <http://www.itl.nist.gov/div898/sird/>. На этом сайте представлены «эталонные» данные и соответствующие им «эталонные» результаты по следующим четырем основным наборам алгоритмов:

- дисперсионный анализ;
- линейная регрессия;
- нелинейная регрессия;
- одномерная суммарная статистика.

На этом сайте, в отличие от сайта NPL, «эталонные» наборы данных представлены гораздо более широко. Например, для каждого набора алгоритмов «эталонные» наборы делятся по уровню сложности, классификации модели (например, экспоненциальная, рациональная или линейная, квадратическая, полиномиальная и др.)

Следует отметить, что, начиная с 2000 года, в ФГУП ВНИИМС достаточно успешно производится испытание ПО методом «черного ящика», при этом реализованы все необходимые программные и аппаратные составляющие для таких испытаний. Речь, правда, идет не о средствах измерений, а об игровых автоматах с денежным выигрышем (ИАДВ). Дело в том, что ключевой частью ИАДВ является игровая программа, являющаяся программной реализацией соответствующего игрового алгоритма, и с этой точки зрения нет никакой разницы между испытанием ПО, входящего в состав средства измерения или в состав ИАДВ.

Ситуация в данной области на настоящий момент характеризуется тем, что законодательно закреплены основные потребительские требования к ИАДВ, в том числе и к его ПО, в виде зарегистрированных Министерством юстиции РФ Правил по метрологии [10], разработана методика статистических испытаний ИАДВ [11, 12], разработаны и аппаратно реализованы соответствующие контроллеры испытаний, аккредитованы несколько соответствующих испытательных центров, в которых по материалам испытаний выдаются соответствующие сертификаты. Следует отметить, что при этих испытаниях для надежного установления, например, такого потребительского

качества игровой программы как зависимость результата игры только от случая, приходится использовать по меньшей мере четыре весьма специфических статистических метода и критерия, таких как метод ранговых корреляций, комбинаторный метод, критерий Колмогорова – Смирнова, критерий Стьюдента. Имеются также наработки по использованию для этих целей критериев “хи – квадрат Пирсона” [13] и “три – сигма” [14].

Опыт, накопленный ВНИИМС в области испытаний ПО ИАДВ, может оказаться полезным при организации и проведении работ по аттестации ПО, используемого в средствах измерений.

На основании описанного опыта испытаний ПО и приведенного обзора международных и зарубежных публикаций можно сформулировать основные положения программы разработки и внедрения в метрологическую практику аттестации программного обеспечения, используемого в средствах измерений.

1. Разработка требований к ПО, используемому в средствах измерений, сначала на уровне методики института (рекомендации), а затем возможно в виде правил по метрологии, зарегистрированных Министерством юстиции РФ. В основу таких требований могут быть положены Рекомендации [9] и наработки NPL и PTB. При этом необходимо иметь в виду, что сфера действия государственного метрологического контроля и надзора в России намного шире, чем в европейских странах. Поэтому в качестве первого шага предполагается требования сформулировать только к такому ПО, которое используется в наиболее распространенных видах измерений (измерения веса, геометрических размеров, электрические измерения, измерения массы и объема газов и жидкостей, расходомерия).
2. Разработка методик испытаний ПО, используемого в разных средствах измерений и измерительных систем, начиная с самых массовых. При этом в каждом конкретном случае необходимо определиться с жесткостью испытаний в зависимости от критичности соответствующего ПО. Трудно представить ситуацию, когда метрологические характеристики средства измерения или измерительной системы полностью бы зависели от соответствующего ПО. Тем не менее, необходимо разработать

количественные оценки критичности ПО и в каждом конкретном случае научиться их оценивать. При этом аттестация ПО должна быть органически встроена в уже существующую систему испытаний средств измерений с целью утверждения их типа без какой-либо радикальной ее перестройки. Предполагается при этом, что основные метрологические характеристики программных средств, о которых уже говорилось, могут быть определены в процессе испытаний традиционными способами, но это можно считать справедливым лишь для неадаптивных программ.

3. Разработка контроллеров испытаний в наиболее распространенных областях измерений.
4. Создание системы центров аттестации ПО, используемого в средствах измерений, прежде всего на базе существующих научных метрологических центров и метрологических институтов, сосредоточив в них соответствующие программные, аппаратные и кадровые ресурсы.
5. Рассмотреть вопрос о целесообразности создания на базе ФГУП ВНИИМС Всероссийского реестра программных средств, используемых в средствах измерений, с целью облегчения идентификации соответствующего ПО и быстрого установления преднамеренных его искажений.

Список литературы

1. H.R. Cook, M.G. Cox, M.P. Dainton, P.H. Harris. Methodology for testing spreadsheets and other packages used in metrology. Report to National Measurement System Policy Unit, September 1999
2. МИ 2174-91 «ГСИ. Аттестация алгоритмов и программ обработки данных при измерениях. Основные положения»
3. МИ 2517-99 «ГСИ. Метрологическая аттестация программного обеспечения средств измерений параметров физических объектов и полей с использованием компьютерных программ генерации цифровых тестовых сигналов»
4. МИ 2518-99 «ГСИ. Метрологическая аттестация алгоритмов и программ генерации цифровых тестовых сигналов»
5. B. Wichman with R. Barker, M. Cox, P. Harris. Software Support for Metrology. Best Practice Guide N 1. Measurement System Validation: Validation of Measurement Software. Report to National Measurement System Policy Unit. April 2000
6. H.R. Cook, M.G. Cox, M.P. Dainton, P.M. Harris. Testing Spreadsheets and Other Packages Used in Metrology. A Case Study. Report to National Measurement System Policy Unit. September 1999
7. M.G. Cox, M.P. Dainton, P.M. Harris. Testing Spreadsheets and Other Packages Used in Metrology. Testing Functions for the Calculation of the Standard Deviation. Report to National Measurement System Policy Unit. October 2000
8. Measuring Instruments Directive (MID/3), Brussels 15.09.2000, COM(2000) 566 final, European Commission – III.D.2
9. Требования к программному обеспечению на основе Директивы по измерительным приборам. WELMEC (European cooperation in legal metrology) 7.1, октябрь, 1999
10. ПР 50.2.023 – 2000 «ГСИ. Правила проведения испытаний игровых автоматов с денежным выигрышем с целью утверждения типа и контроля за их соответствием утвержденному типу»
11. МИ 2562 – 2001 «ГСИ. Игровые автоматы с денежным выигрышем. Методика обработки статистической информации при испытаниях с целью утверждения типа»

- 12.Акимов А.А., Ю.А. Кудеяров, Кузнецов В.П. Статистическая обработка результатов испытаний игровых автоматов с денежным выигрышем. Законодательная и прикладная метрология, № 5, 2001, 35 – 43
- 13.Ю.А. Кудеяров, А.Г. Спроге. Критерии согласия, используемые при статистической обработке результатов испытаний игровых автоматов с денежным выигрышем. Законодательная и прикладная метрология, № 6,2001, 25 –29
14. Ю.А. Кудеяров. Использование критерия «3 – сигма» при испытаниях игровых автоматов с денежным выигрышем с целью утверждения их типа. Законодательная и прикладная метрология, № 1, 2002

АННОТАЦИЯ

В статье дан обзор существующих на настоящий момент в зарубежной и отечественной практике подходов к метрологической аттестации программного обеспечения, используемого в средствах измерений. Предложена программа разработки такой системы аттестации, рассчитанная прежде всего на средства измерения и измерительные системы в сфере действия государственного метрологического контроля и надзора.