

ТЕСТИРОВАНИЕ КАЧЕСТВА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ В ПРОЦЕССЕ ЕГО СЕРТИФИКАЦИИ

Бурый А.С.*

Ключевые слова: программное обеспечение, стандартизация, признаки качества, сертификация программных средств, таксономия, тестирование, меры сходства, кластеры, гранулярные вычисления.

Аннотация.

Цель работы: совершенствование научно-методической базы сертификации программной продукции, как элемента нормативно-правового регулирования в области стандартизации.

Метод: информационный анализ, моделирование, статистический анализ, функционально-логическая классификация.

Результаты: на основе существующих тенденций в области развития методов интеллектуального анализа данных, показана актуальность разработки методов информационной поддержки принятия решений в ходе тестирования программного обеспечения при их сертификации, за счет автоматизации процессов испытаний при обработке экспертной информации по оценке признаков качества программ; показано, что процесс сертификации программного обеспечения основывается на нормативно-правовых актах в области стандартизации, являясь одним из основных элементов защиты интеллектуальной собственности применительно в программным средствам; в формализованном виде представлена задача формирования кластеров программных средств в зависимости от выявленного уровня показателей качества; достоверность сделанных выводов подтверждается результатами моделирования.

DOI:10.21681/1994-1404-2019-1-46-55

Введение

В современном обществе информация и информационные технологии по существу определяют основные направления развития большинства областей науки и техники. Роль информационных коммуникаций активно используется во многих технологических процессах, а выполнение и реализация ряда проектов связаны: с использованием сетевых технологий и распределенных баз данных; с оптимизацией информационных ресурсов и необходимостью постоянного информационного обмена между исполнителями отдельных работ. При этом применяемые на любом рабочем месте программные средства (ПС) должны обеспечивать надежное получение данных, их переработку и дальнейшее использование в коммуникационной среде. Сертификация программной продукции является для пользователей дополнительным фактором – гарантом качества, обеспечивающего заданный уровень совместимости (по операционной среде, применяемым программным пакетам и комплексам, а также входным или выходным форматам данных).

Существующий методический инструментарий взаимодействия организационно-технических структур в ходе подготовки и проведения тестирования про-

граммных средств требует применения новых подходов к процессу испытаний. Это связано как с новым классом самих объектов тестирования, их большого разнообразия и уникальности практически каждого образца, так и с появлением новых методов интеллектуального анализа экспериментальных данных. Опираясь на комплексный «ИКС»-подход («информационно-кибернетически-синергетический») к анализу и оптимизации сложных систем, существующие формы структурирования измерительной информации и ее метрическое представление [9, 10], появляется возможность применить известные методы интеллектуального анализа данных к оценке качественных признаков ПС.

С другой стороны, оценка качества программной продукции требует разработки индивидуальных методик и подходов к проведению сертификационных испытаний.

Правовое регулирование в области стандартизации осуществляется на основании Федерального закона №162-ФЗ «О стандартизации в Российской Федерации»¹, а порядок сертификации и её основные виды регламентируются Федеральным законом №184-ФЗ «О техническом регулировании»².

¹ О стандартизации в Российской Федерации [Текст]: федер. закон от 29 июня 2015 №162-ФЗ. – М.: ФГУП «СТАНДАРТИНФОРМ», 2015. – 72 с.

² О техническом регулировании [Текст]: Федер. закон от 27 декабря 2002 №184-ФЗ. – М.: Издательство «Омега-Л», 2014. – 52 с.

* **Бурый Алексей Сергеевич**, доктор технических наук, эксперт РАН, директор департамента ФГУП «Российский научно-технический центр информации по стандартизации, метрологии и оценке соответствия», Российская Федерация, г. Москва
E-mail: a.s.burij@gostinfo.ru



Рис. 1. Основные направления развития методов тестирования программного обеспечения³

Организационные особенности тестирования программного обеспечения

Процесс тестирования или испытания программных продуктов, с одной стороны, носит консервативный характер, так как строится на методологии начала 90-х годов (смотри, например, ГОСТ Р ИСО/МЭК 9126–93⁴), а с другой, должен учитывать тенденции современного рынка компьютерных технологий, модельно-алгоритмических средств для их оценивания и моделирования процессов функционирования [1 11]. Многовариантный процесс выбора признаков качества, способных характеризовать свойства сложных программных комплексов, объединяющих разнообразие процедуры (подпрограммы, встроенный интерфейс и средства безопасности) с соответствующими связями и отношениями на них делает тестирование в ходе сертификации ПС уникальным для каждого объекта и требует разработки отдельных методик для их проведения⁵.

На рис. 1 показаны основные тенденции в области тестирования программных средств. Так *адаптивность* предполагает постоянную координацию целевых и функциональных задач [7]. Машинное обучение означает отслеживание и учет уникальных, в том числе и нестандартных ситуаций, например, при выявлении сбоев и ошибок, для корректировки применяемых тестов. Тестирование в процессе разработки ПС предполагает параллельный процесс тестирования по завершению определенных этапов разработки программ. Контроль больших данных [12]⁶, как известно, имеет свои особен

ности и связан с большим комплексом задач: от алгоритмов защиты данных до вопросов структурирования, хранения, обеспечения полноты и достоверности данных и заканчивая дублированием данных.

По мере появления новых технологий, в частности Интернет вещей (Internet of Things – IoT технологий), возникают задачи комплексного тестирования формируемых аппаратно-программных средств для решения ряда бизнес-задач, в том числе аутентификации пользователей сервисов, конфиденциальности данных и др. Одной из платформ реализации данной технологии является технология Hadoop, ключевая для Интернета вещей и ряда социальных сетей.

Автоматизация тестирования и испытания (AI) обеспечивает ускорение выполнения ряда рутинных операций, фиксирование и анализ результатов тестирования [6] в сочетании с еще существующими ручными операциями, от которых практически трудно уйти, например, при разработке дизайна, структурировании процессов и алгоритмов. Наконец сокращение цикла модификации программ при разработке новых версий и приложений должно максимально базироваться на существующих методиках для ускорения внедрения новых ПС, за счет применения шаблонов (паттернов) проектирования в рамках конкретного контента. Интеграция существующего и разрабатываемого инструментария обеспечивает выполнение требований по качеству проводимого тестирования ПС и управления решаемыми технологическими и информационными задачами.

Мировая индустрия в области информационных технологий характеризуется постоянным ростом, так в 2019 г. её оборот может достигнуть 5 трлн. долларов. И хотя доля программного обеспечения составляет (см. рис. 2) от данной суммы только 11%, согласно отчету ComptIA⁷, в абсолютном выражении она также значительна и во многом определяет вектор развития

³ 10 software testing trends to watch out for in 2019 // ReQtest. – URL <https://reqtest.com/testing-blog/software-testing-trends-2019/> (дата обращения 23.02.2019).

⁴ ГОСТ Р ИСО/МЭК 9126–93. Информационная технология. Оценка программной продукции. Характеристики качества и руководство по применению. – М.: ИПК Издательство стандартов, 2003. – 12 с.

⁵ Бурый А.С., Морин Е.В. Когнитивная модель оценки качества информационных технологий // Информационно-экономические аспекты стандартизации и технического регулирования. – 2018. – № 1(41). – URL: http://iea.gostinfo.ru/files/2018_01/2018_01_02.pdf (дата обращения 23.02.2019)

⁶ См. Также: Oracle Data Technologies. – URL <https://www.oracle.com/database/technologies/> (дата обращения 23.02.2019); Hadoop.

What it is and why it matters? // SAS. – URL https://www.sas.com/en_us/insights/big-data/hadoop.html (дата обращения 23.02.2019).

⁷ IT industry outlook 2019. – URL <https://www.comptia.org/resources/it-industry-trends-analysis> (дата обращения 22.02.2019).

Информационные и электронные технологии в правовой сфере

информатизации современного общества, наряду с проведением фундаментальных и прикладных исследований по информатике, с разработкой новых информационных технологий, развитием индустрии переработки информации и корректировкой соответствующих правовых и хозяйственных норм [8].

Задача сертификации программного обеспечения заключается в подтверждении соответствия показателей качества (см. ГОСТ Р ИСО/МЭК 9126–93) действующим нормативным документам.

Качество программного обеспечения структурируется на:

фундаментальные основы, включающие методический аппарат, модели характеристик качества, разработки программ и сценариев и моделей безопасности ПС [11];

менеджмент качества, включающий процедуры оценивания качества; практический инструментарий, в том числе требования к программному обеспечению, характеристики возможных дефектов и методы измерения качества⁸.



Рис. 2. Структура мировой индустрии информационных технологий

Основные характеристики или признаки качества программных средств показаны на рис. 3⁹, причем для каждой характеристики определены свои подхарактеристики, которые тестируются в ходе разработки и сер-

⁸ Laporte C.Y., April A. Software quality assurance. Wiley-IEEE Computer Society, Inc., 2018. – URL https://www.researchgate.net/publication/320853147_Software_Quality_Assurance (дата обращения 22.02.2019).

⁹ ГОСТ Р ИСО/МЭК 25010-2015. Информационные технологии. Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов. – М.: Стандартинформ. – 2015. – 36 с.

тификации программных продуктов, а с точки зрения точности измерения их показателей они делятся на:

- *категорийные* или описательные, например, по типу терминологических данных, реализуемые в номинальных шкалах;
- *качественные*, например, факторы удобства применения или использования (понятность, простота использования, обучаемость), оцениваемые экспертами в порядковых шкалах;
- *количественные*, представляемые числовыми упорядоченными точками на интервальных или относительных шкалах.

Большинство из указанных на рис. 3 признаков качества свойственны, как аппаратным, так и программным средствам целевых компьютерных систем различного назначения. Так, *уровень производительности*¹⁰ в качестве подхарактеристик включает временные характеристики, использование ресурсов, потенциальные возможности, что подтверждает предыдущий тезис об общности аппаратных и программных средств.

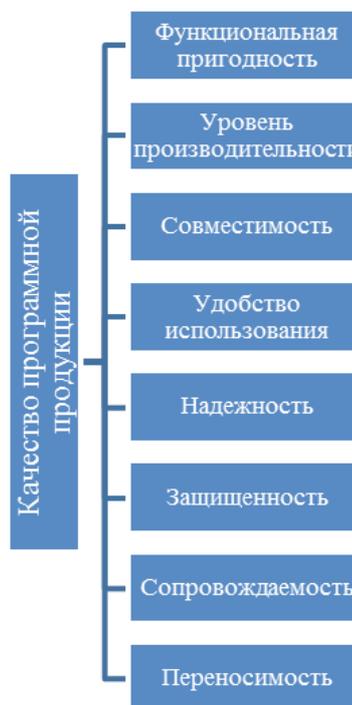


Рис. 3. Модель качества программных средств

Применительно к характеристикам качества требуется сопоставлять формы понятия «мера»: показатель – признак – параметр, свойственные испытательному процессу при разработке¹¹ программных средств, которые проиллюстрированы на рис. 4.

¹⁰ ГОСТ Р ИСО/МЭК 25010-2015. С. 11.

¹¹ Boehm B. A view of 20th and 21st century software Engineering // Proceedings of the 28th international conference on Software engineering, ICSE'06, May 20-28, 2006. – Shanghai, 2006. – P. 12-29. DOI: 10.1145/1134285.1134288

Все представленные термины составлены на основе действующих ГОСТов в области информационных технологий. Разнообразие связей и терминологических отношений подтверждает одно из определений *информации* как свойство объектов (процессов) окружающего материального порождать разнообразие состояний,

которые посредством отражения передаются от одного объекта к другому в процессе их взаимодействия [8], одновременно со структурной и содержательной трансформацией, характерной для коммуникационных процессов [9, 13].



Рис. 4. Таксономия основных понятий элементов качества ПС

Оценка признаков качества программного обеспечения методом кластерного анализа

Основная задача кластеризации – формирование подмножеств объектов по определенному кластерному признаку или свойству, для чего часто используют различные меры (отношения) сходства или различия. Данный подход широко используется в задачах поиска аналогов (прототипов)¹², распознавания образов [5] и диагностирования [14].

Представим задачу формирования оценок признаков качества ПС в следующем виде. Обозначим набор тестируемых объектов, как $O_1, O_2, \dots, O_h \in O$, каждый из которых описывается набором признаков (качеств) $Q_1, Q_2, \dots, Q_q \in Q$, которые оцениваются экспертами, общее число которых определяется мощностью множества $E = \{1, 2, \dots, e\}$. Соответственно

для измерительного пространства L , образованного указанными множествами, т.е.

$$L = O \times Q \times E, \quad (1)$$

общий объем измерений составляет $card(L) = q \times h \times e$. Следовательно, каждый объект представляется в виде матрицы

$$L = \left\| l_{ij} \right\|_{q \times h'}$$

где l_{ij} – есть значение признака -го качества для j -го объекта тестирования (программного средства), при этом (i – единичный наблюдатель), т.е. увеличение количества наблюдателей (экспертов) приведет к увеличению строк матрицы L .

С другой стороны, совокупность таблиц «объект – признак» представляет собой некоторую структуру отношений. При этом в виду разнородности отдельных признаков, представляемых в качественных и количественных шкалах, можно говорить о задаче *многомерного шкалирования* [4].

Любой признак характеризуется множеством допустимых значений, что соответствует отображению:

$$Q_i \xrightarrow{f_i} D_{Q_i}, \quad i = \overline{1, q}, \quad (2)$$

¹² Iglesias F., Kastner W. Analysis of similarity measures in times series clustering for the discovery of building energy patterns // Energies, – 2013.– No. 6. – Pp. 579-597. DOI: 10.3390/en6020579

где D_{Q_i} – множество допустимых значений i -го признака качества, которое в зависимости от типа включенных элементов может быть бинарным, т.е. содержать два элемента $\{0, 1\}$, интервальным, количественным и др. и определяет соответствующую шкалу при переходе, благодаря отображению f_i , к решению измерительной задачи.

Декартово произведение допустимых значений признаков из (2) вида:

$$D_Q = D_{Q_1} \times \dots \times D_{Q_q} \quad (3)$$

называют *признаковым пространством*, размерность которого напрямую зависит от размерности шкал признаков, объединяемых множеством D_Q .

Для каждого тестируемого объекта $O_j \in \mathbf{O}$ существует свой эталон (стандарт) или образ с требуемыми параметрами признаков качества. То есть по результатам тестирования j -го объекта формируется разностный вектор δ_j между оценками, полученными по результатам измеренных признаков качества Q_j^n , и соответствующими их эталонными значениями (Q_j^a) вида:

$$\delta_j = Q_j^n - Q_j^a. \quad (4)$$

Первичная информация, получаемая по результатам экспертного оценивания, структурируется в виде таблицы «объект – признак», где объекты это анализируемые (тестируемые) продукты, а признаки – оценки, выставленные экспертами в соответствии с определенной измерительной шкалой.

На предварительном этапе анализа информации, полученной экспертным путем, необходимо выявить

такие объекты, которые имеют хотя бы один признак, несоответствующий эталонным параметрам.

Для примера проведем анализ двух объектов O_1 и O_2 по двум признакам, измеряемым относительно признаков j -го эталона. На рис. 5 в системе координат двух признаков l_1 и l_2 показан вектор эталонных признаков $Q_j^a = \{l_{1j}^a, l_{2j}^a\}$. Будем считать, что для любого j -го объекта тестирования факт соответствия требованиям к программному обеспечению есть попадание значений контролируемых признаков в область \mathbf{I} , ограниченную линиями, проходящими через координаты эталонного объекта, что можно представить, как

$$l_{ij} \geq l_{ij}^a; i, j = 1, 2. \quad (5)$$

Ограничения области \mathbf{I} являются предельные значения признаков. Часто это единичные значения (при максимизации показателей надежности или других вероятностных характеристик, стремящихся к единичному пределу при повышении параметра качества).

Таким образом, в результате рассмотренной процедуры сравнения для примера, представленного на рис. 5, делается вывод, что объект O_2 не соответствует требованиям, так как для признака l_1 условие (5) не выполняется и, как видно из рисунка, $l_{12} < l_{1j}^a$.

По завершению формирования области \mathbf{I} возникает задача кластеризации объектов, выявляя такие из них, которые по определенным признакам качества могут использоваться для выполнения конкретных задач, например, по степени безопасности, надежности и др.

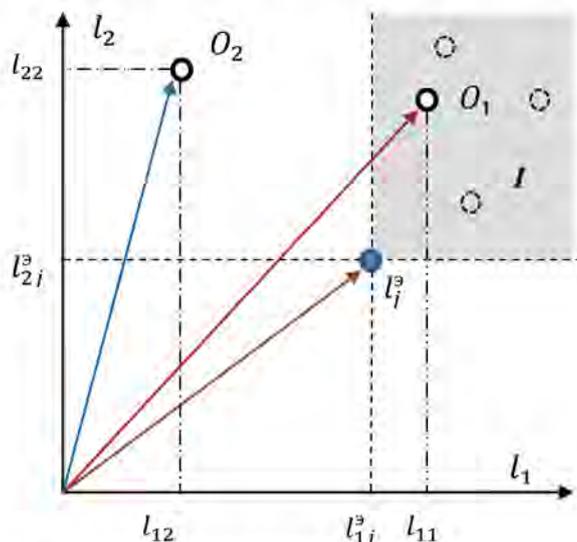


Рис. 5. Сравнение объектов по двум признакам

Для этого из существующих подходов [4, 5] к вычислению коэффициентов сходства (шкал, мер) методами расчета корреляции, расстояния ассоциаций и вероятностных показателей остановимся на разностном векторе вида (4), т.е. расстоянии между объектами (пунктирные кружки на рис. 5) в области \mathbf{I} . В зависимости от величины расстояния осуществляется группирование объектов, соответствующих определенным требованиям. Это расстояние должно удовлетворять таким акси-

омам метрики, как аксиомы тождественности, симметрии и треугольника [4]¹³:

$$a) d(x, y) = 0 \text{ соответствует равенству } x = y;$$

¹³См. Также: Shirkorshidi A.S., Aghabozorgi S., Wah T.Y. A comparison study on similarity and dissimilarity measures in clustering continuous data // PloS one. – 2015. – № 10(12). – P. 1-12. DOI: 10.1371/journal.pone.0144059; Meng L., Huang R., Gu J. A review of semantic similarity measures in WordNet // International Journal of Hybrid Information Technology. – 2013. – № 6(1). – P. 1-12.

Таблица 1

Определение типовых метрик расстояния

Метрики	Формализация	Применение
Евклидово расстояние	$d_{euc}(x, y) = \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2}$	Является геометрическим расстоянием в многомерном пространстве, для однородных объектов
Расстояние Чебышева	$d_{ch}(x, y) = \max_i x_i - y_i $	Для объектов, различающихся хотя бы по одной координате
Взвешенное Евклидово расстояние	$d_{we}(x, y) = \left(\sum_{i=1}^n w_i (x_i - y_i)^2 \right)^{1/2}$	Для случая, когда у каждого признака существует степень важности (по экспертным оценкам)
Манхэттенское расстояние	$d_{mh}(x, y) = \sum_{i=1}^n (x_i - y_i)$	Уменьшает влияние отдельных больших выбросов

b) $d(x, y) = d(y, x)$;

c) $d(x, y) \leq d(x, c) + d(c, y)$, где x, y, c – любые элементы метрического пространства, в котором также выполняется условие $d(x, y) \geq 0$.

Для любого класса мер характерны определенные математические свойства, которые соотносятся с общепринятыми понятиями и которым даны содержательные названия [6], как возможность, вероятность, правдоподобность, сходство, связность, доверие и др.

На примере Евклидова пространства и многообразия метрик¹⁴, частично представленных в табл. 1 (для объектов и с соответствующими координатами), формируются таблицы «объект – признак», которые наиболее близки по своим свойствам стандартным (эталонным) объектам из области I , т.е. $d_{ij}(I) \leq \Delta_{ij}$, где

$$d_{ij}(L) = \left(\sum_{i=1}^q (l_{ij} - l_{ij}^g)^2 \right)^{1/2}, \quad (6)$$

а Δ_{ij} – уровни требований для каждого свойства у соответствующего -го объекта; i и j – соответствуют описанию элементов выражения (1).

Для определения объекта, наиболее близкого по своим свойствам (признакам качества) стандартному образцу воспользуемся минимизацией функционала, называемого стрессом¹⁵:

$$S(L) = \sum_{i < j} w_{ij} (\Delta_{ij} - d_{ij}(L))^2 \rightarrow \min. \quad (7)$$

Здесь $W = \|w_{ij}\|$ – известная симметричная неотрицательная весовая матрица, причем веса выбираются, исходя из целей шкалирования, а также исходя из физического смысла решаемой задачи. Обычно ис-

пользуют $w_{ij} = \Delta_{ij}^\gamma$, тогда при $\gamma < 0$ предпочтительнее меньшие расстояния, в противном случае $\gamma > 0$.

Функционал стресса $S(L)$ приобретает физический смысл потенциальной энергии при $\gamma = -2$ для системы из n связанных точек, тогда уравнение (7) соответствует поиску равновесного состояния системы, в котором потенциальная энергия минимальна.

Наиболее простым и физически понятным свойством кластера является его компактность, которую можно определить как среднюю длину расстояний между h объектами, используя следующее выражение:

$$\bar{D} = \frac{1}{h} \sum_{i=1}^{h-1} \sum_{j>i}^h d_{ij}.$$

Данный показатель представляет, насколько далеки друг от друга любые два узла. Чем меньше \bar{D} , тем ближе друг к другу объекты и кластер компактнее.

Задача разбиения множества исходных данных на кластеры может решаться методом последовательного разбиения данных, для чего используются иерархические алгоритмы, либо агломеративным методом, путем объединения кластеров¹⁶.

Существующие системы хранения больших данных [12] характеризуются многомерностью, масштабно-стью, распределённостью (временной и территориальной). Поиск и переработка хранимых данных с целью получения новой информации требует активного использования неструктурированной информации, основываясь на когнитивных свойствах знаний, на комплексном анализе скрытых компонентов по результатам оценивания, распознавания, кластерного анализа и других методов обработки в ходе обобщения и систематизации данных.

К такому направлению можно отнести теорию нечеткого гранулирования информации Л.А. Заде [16], опирающуюся на теории нечетких множеств и лингви-

¹⁴ Choi, S-S., Cha, S-H., Tappert, C.C. A survey of binary similarity and distance measures // Journal of Systemics, Cybernetics and Informatics. – 2010. – № 8(1). – P. 43-48.

¹⁵ De Leeuw, J., Mair, P. Multidimensional scaling using majorization: SMACOF in R // Journal of Statistical Software. – 2009. – Vol. 31(3). – P. 1-30. DOI: 10.18637/jss.v031.i03

¹⁶ Dalton L., Ballarin V., Brun M. Clustering algorithms: on learning, validation, performance, and applications to genomics // Current Genomics. – 2009. – № 10(6). – P. 430-445. DOI: 10.2174/138920209789177601

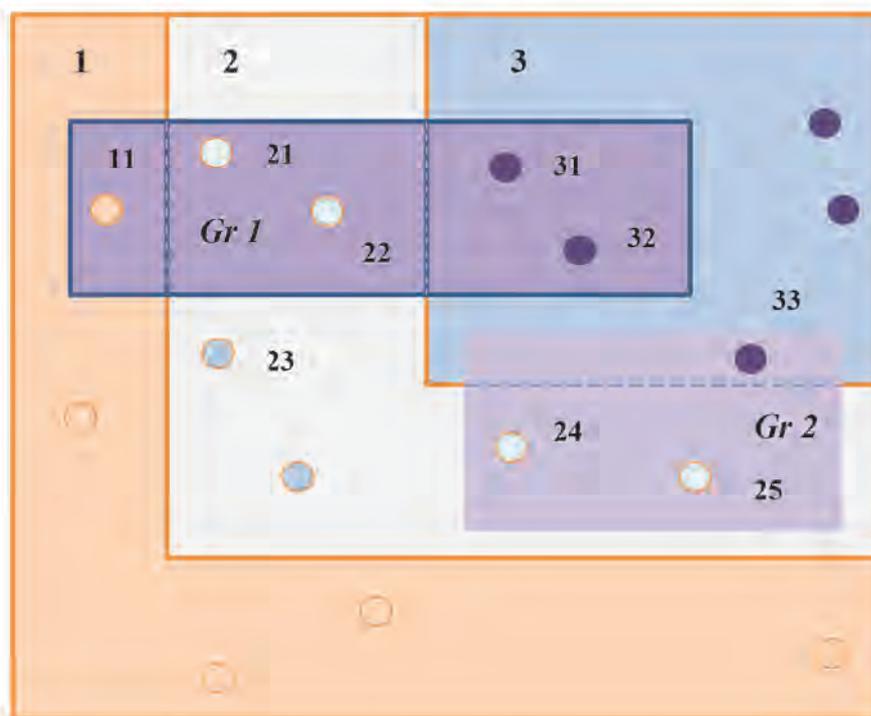


Рис. 6. Распределение объектов тестирования по гранулам

стических переменных и направленную на сокращение вычислительной сложности, а также на идею сближения системы построения машинного вывода к логике рассуждений человека.

Под *гранулой* будем понимать набор элементов, объединенных отношениями сходства со свойствами симметричности и рефлексивности, и обладающий структурным и целевым содержанием¹⁷. Гранулярные вычисления используются для описания неопределенности объектов в условиях неполной или нечеткой информации [2]. Идея гранул очень близка к алгоритмам кластеризации, так как основывается на структуризации как признакового пространства, так и структуризации области решения задачи, делая процесс формирования решения более динамичным и адаптивным к широкому кругу задач¹⁸.

На рис. 6 показан подход к формированию кластерных гранул. Продолжая идею рис. 5, показана двух признаковая плоскость, на которой размещены объекты тестирования, условно обозначенные кружками. Цифрами 1-3 выделены подобласти, с целью показать, что принцип формирования кластерных гранул может быть любой, например, исходя из задач категорирования объекта тестирования. Так, области 2 и 3 соответствуют области *I* для объектов, признаки качества которых удовлетворяют заданным требованиям.

В качестве примера на рис. 6 показаны гранулы, объединяющие объекты в соответствии с принятой системой отношений:

$$Gr1 = \{11, 21, 22, 31, 32\}$$

$$Gr2 = \{24, 25, 33\}$$

При этом если в гранулу *Gr2* вошли объекты, полностью соответствующие требованиям (области 2 и 3), то в гранулу *Gr1* – объект из области 1, один из признаков качества которого не в норме. На практике это соответствует ситуации, когда тестируемое ПС не удовлетворяет жестким требованиям к программному обеспечению, например, для систем управления движением (по точности или устойчивости к сбоям), но может использоваться в учебных целях, в комплексах моделирования и др. с понижением заявленной категории.

Эксперимент

Для оценки влияния отдельных факторов на формирование кластеров объектов тестирования на этапе категорирования программной продукции проведено машинное моделирование. Исходными данными, представленными в табл. 2 являются:

- значения двух характеристик качества – l_1 и l_2 ;
- общее число анализируемых объектов $q = 6; 12; 18$;
- число рассматриваемых кластеров – 2 и 3;
- методы поиска решения – ближайшего соседа (МБС) и дальнего соседа (МДС).

На начальном этапе составляется матрица расстояний между анализируемыми объектами. В качестве меры сходства между признаками объектов использовалось Евклидово расстояние (см. табл.1). В ходе

¹⁷ Тарасов В.Б. Универсальная логика, грануляция информации и искусственный интеллект. – URL <http://www.raai.org/news/pii/ppt/2015/tarasov2015.ppt> (дата обращения 24.02.2019).

¹⁸ Ding S., Du M., Zhu H. Survey on granularity clustering // Cognitive Neurodynamics. – 2015. – Vol. 9(6). – P. 561-572. DOI: 10.1007/s11571-015-9351-3

Таблица 2

Исходные данные для проведения эксперимента

№	1	2	3	4	5	6	7	8	9	10	11	12
l_1	0,1	0,1	0,15	0,2	0,2	0,15	0,3	0,25	0,35	0,3	0,65	0,7
l_2	0,85	0,7	0,65	0,75	0,9	0,7	0,5	0,8	0,95	0,85	0,7	0,2
№	13	14	15	16	17	18	19	20	21	22	23	24
l_1	0,75	0,85	0,85	0,8	0,8	0,9	0,95	0,98	0,1	0,27	0,17	0,35
l_2	0,25	0,3	0,35	0,1	0,15	0,25	0,3	0,4	0,75	0,9	0,7	0,5

Таблица 3

Результаты моделирования

Кластеры	2		3					
	МБС	МДС	МБС			МДС		
6	0,14	0,255	0,141	0,112	0,15	0,255	0,15	0,112
12	0,5	0,885	0,381	0,5	0,502	0,57	0,885	0,502
18	0,403	1,03	0,381	0,5	0,403	0,57	1,03	0,65
24	0,403	1,03	0,361	0,461	0,403	0,57	1,03	0,687
			K1-K2	K1-K3	K2-K3	K1-K2	K1-K3	K2-K3

моделирования на каждом шаге в алгоритме МБС ищется пара с минимальным расстоянием, которая исключается с заменой на объединенный элемент. В итоге получаем 2 кластера с группой объектов в каждом из них и условным расстоянием между кластерами, которое фиксируется в табл. 3. Для трех кластеров та-

ких расстояний 3: между кластерами 1 и 2 обозначено, как (K1-K2), и соответственно, для других вариантов – (K1-K3); (K2-K3).

Результаты моделирования представлены в табл. 3 в виде расстояний между сгруппированными кластерами.

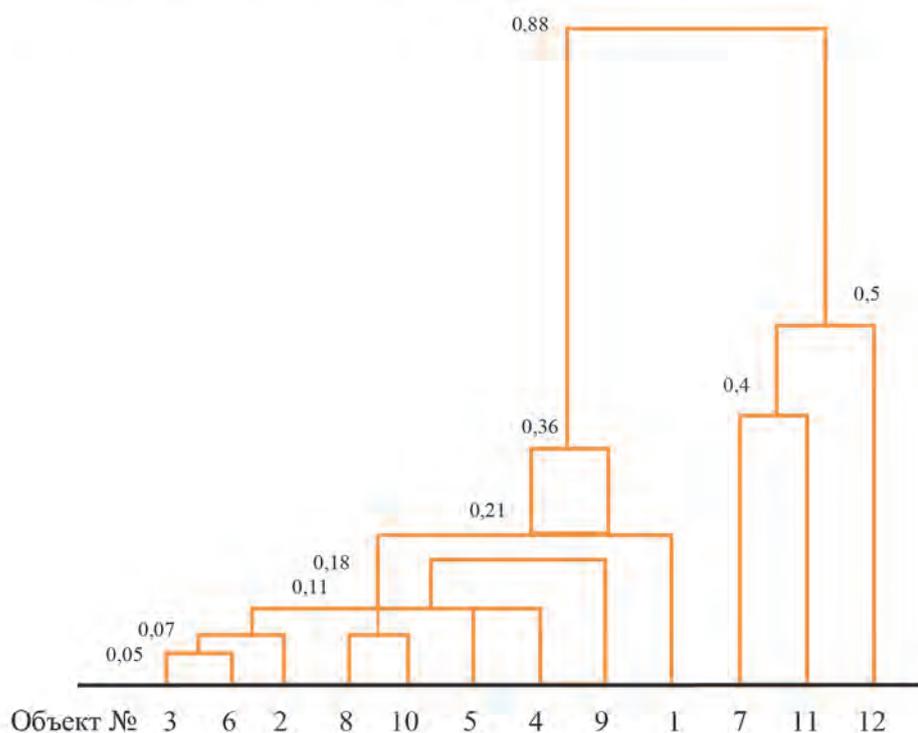


Рис. 7. Дендрограмма распределения 12 объектов по двум кластерам

На рис. 7 представлены результаты распределения 12-и двух признаков объектов по двум кластерам в виде дендрограммы, полученной методом дальнего соседа, которая пошагово иллюстрирует процедуру кластеризации. Номера объектов тестирования расположены под горизонтальной прямой диаграммы, где они записываются в порядке их попадания в формируемые кластеры. Числа над линиями соединения пар объектов, либо их промежуточных объединений, соответствуют расстояниям между ними. В итоге объекты по классам распределились следующим образом:

$K_2 = \{7, 11, 12\};$

$K_1 = \{3, 6, 2, 8, 10, 5, 4, 9, 1\}.$

Заключение

Таким образом, кластерный подход к задачам тестирования программного обеспечения и анализа его результатов с целью категорирования ПС и интеллектуального

анализа данных позволяет повысить обоснованность принятия решений в ходе сертификации ПС. Наличие сертификата является дополнительной правовой нормой подтверждения качества продукции и защиты потребителей, а также элементом реализации государственной политики информатизации средствами стандартизации.

Дальнейшим направлением развития настоящей работы видится в части применений *информационно-энтропийных мер* [15] для формирования информационных кластеров, исходя из целевого предназначения тестируемых объектов и исследования идей гранулирования информации при формировании информационных шаблонов объекта.

Обеспечение необходимого уровня качества программных продуктов должно постоянно отслеживаться в ходе разработки систем высокого уровня безопасности, таких, в частности, как телекоммуникационные инфраструктуры, которые присущи как социосфере, так и техносфере.

Рецензент: **Марков Алексей Сергеевич**, доктор технических наук, профессор, Главный редактор журнала «Вопросы кибербезопасности», г. Москва, Россия.

Литература

1. Бурый А.С. Отказоустойчивые распределенные системы переработки информации. – М.: Горячая линия – Телеком, 2016. – 128 с.
2. Бутенков С.А., Кривша Н.С., Кривша В.В. Топологические пространственные отношения в моделях гранулирования многомерных данных // Известия ЮФУ. Технические науки. – 2010. – № 6 (107). – С. 200 – 211.
3. Дастин Э., Рэшка Д., Пол Д. Автоматизированное тестирование программного обеспечения: Внедрение, управление и эксплуатация. – М.: Изд-во «ЛОРИ», 2003. – 567 с.
4. Дюран Б., Одделл П. Кластерный анализ. – М.: Статистика, 1997. – 128 с.
5. Загоруйко Н.Г., Борисова И.А. и др. Количественная мера компактности и сходства в конкурентном пространстве // Сибирский журнал индустриальной математики. – 2010. – Т. XIII. – № 1. – С. 59-71.
6. Клир Дж. Системология. Автоматизация решения системных задач. – М.: Радио и связь, 1990. – 544 с.
7. Стеллман Э., Грин Дж. Постигая Agile. Ценности, принципы, методологии. – 3-е изд. – М.: Манн, Иванов и Фербер, 2019. – 448 с.
8. Юсупов Р.М., Заболотский В.П. Научно-методические основы информатизации. – СПб.: Наука, 2000. – 455 с.
9. Ловцов Д. А. Концепция комплексного «ИКС»-подхода к исследованию сложных правозначимых явлений как систем // Философия права. – 2009. – № 5. – С. 40 – 45.
10. Ловцов Д. А. Системология правового регулирования информационных отношений в инфосфере: Монография. – М.: РГУП, 2016. – 316 с..
11. Марков А.С. Модели оценки и планирования испытаний программных средств по требованиям безопасности информации // Вестник МГТУ им. Н.Э. Баумана. Сер.: Приборостроение. – 2011. – № 51. – С. 90 – 103.
12. Федосеев С. В. Применение современных технологий больших данных в правовой сфере // Правовая информатика. – 2018. – № 4. – С. 50 – 58.
13. Buryi, A.S., Loban, A.V., Lovtsov, D.A. Compression models for arrays of measurement data in an automatic control systems // Automation and Remote Control. – 1998. – Vol. 59(5). – Part 1. – P. 613 – 631.
14. Buryi, A.S., Polous, A.I., Shlyakonov, V.A. A functional diagnostic method for program-controlled objects // Automation and Remote Control. – 1998. – Vol. 59(4). – Part 2. – P. 599 – 602.
15. Lovtsov D. A. Models for Measuring the Information Resource of a Computerized Control System // Automation and Remote Control. – 1996. – Vol. 57. – No 9. – Part 1. – P. 1221 – 1232.
16. Zadeh L.A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic // Fuzzy Sets and Systems. – 1997. – Vol. 90(2). – P. 111-127.

TESTING THE QUALITY OF SOFTWARE IN THE PROCESS OF ITS CERTIFICATION

Aleksei Buryi, Doctor of Science (Technology), expert at the Russian Academy of Sciences, Director of a department of the Russian Scientific and Technical Centre of Information On Standardisation, Metrology and Conformity Assessment, Moscow, Russian Federation.

E-mail: a.s.burij@gostinfo.ru

Keywords: software, standardisation, quality attributes, software certification, taxonomy, testing, similarity measures, clusters, granular computations.

Abstract.

Purpose of the work: improving the scientific and methodological basis for the certification of software products as an element of legal regulation in the field of standardisation.

Method used: information analysis, modelling, statistical analysis, functional and logical classification.

Results obtained: based on the existing trends in the field of development of methods of intelligent data analysis (data mining), the topicality of developing methods of information support for decision-making in the course of software testing during its certification using the automation of testing processes in processing expert information concerning the evaluation of software quality attributes is shown; it is shown that the process of software certification is based on legal regulations in the field of standardisation, being one of the main elements of intellectual property protection in relation to software; the task of forming clusters of software tools depending on the level of their quality indicators is presented in a formalised form; the reliability of conclusions made is confirmed by the results of modelling.

References

1. Buryi A.S. Otkazoustoichivye raspredelennye sistemy pererabotki informatsii, M. : Goriachaia liniia -- Telekom, 2016, 128 pp.
2. Butenkov S.A., Krivsha N.S., Krivsha V.V. Topologicheskie prostranstvennye otnosheniia v modeliakh granulirovaniia mnogomernykh dannykh, Izvestiia IuFU, Tekhnicheskie nauki, 2010, No. 6 (107), pp. 200-211.
3. Dastin E., Reshka D., Pol D. Avtomatizirovannoe testirovanie programmnoho obespecheniia: Vnedrenie, upravlenie i ekspluatatsiia, M. : LORI, 2003, 567 pp.
4. Diuran B., Odell P. Klasternyi analiz. M.: Statistika, 1997, 128 pp.
5. Zagoruiko N.G., Borisova I.A. i dr. Kolichestvennaia mera kompaktnosti i skhodstva v konkurentnom prostranstve, Sibirskii zhurnal industrial'noi matematiki, 2010, t. XIII, No. 1, pp. 59-71.
6. Klir Dzh. Sistemologiya. Avtomatizatsiia resheniia sistemnykh zadach. M.: Radio i sviaz', 1990, 544 pp.
7. Stellman E., Grin Dzh. Postigaia Agile. Tsennosti, printsipy, metodologii, 3-e izd., M. : Mann, Ivanov i Ferber, 2019, 448 pp.
8. Iusupov R.M., Zabolotskii V.P. Nauchno-metodicheskie osnovy informatizatsii, SPb. : Nauka, 2000, 455 pp.
9. Lovtsov D. A. Kontsepsiia kompleksnogo "IKS"-podkhoda k issledovaniiu slozhnykh pravoznachimykh iavlenii kak sistem, Filosofia prava, 2009, No. 5, pp. 40-45.
10. Lovtsov D. A. Sistemologiya pravovogo regulirovaniia informatsionnykh otnoshenii v infosfere : monografiia, M. : RGUP, 2016, 316 pp.
11. Markov A.S. Modeli otsenki i planirovaniia ispytanii programmnykh sredstv po trebovaniiam bezopasnosti informat-sii, Vestnik MGTU im. N.E. Bauman, ser.: Priborostroenie, 2011, No. 51, pp. 90-103.
12. Fedoseev S. V. Primenenie sovremennykh tekhnologii bol'shikh dannykh v pravovoi sfere, Pravovaia informatika, 2018, No. 4, pp. 50-58.
13. Buryi, A.S., Loban, A.V., Lovtsov, D.A. Compression models for arrays of measurement data in an automatic control systems, Automation and Remote Control, 1998, vol. 59(5), part 1, pp. 613-631.
14. Buryi, A.S., Polous, A.I., Shlyakov, V.A. A functional diagnostic method for program-controlled objects, Automation and Remote Control, 1998, vol. 59(4), part 2, pp. 599-602.
15. Lovtsov D. A. Models for Measuring the Information Resource of a Computerized Control System, Automation and Remote Control, 1996, vol. 57, No. 9, part 1, pp. 1221-1232.
16. Zadeh L.A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems, 1997, vol. 90(2), pp. 111-127.