

ИНФОРМАТИКА, КИБЕРНЕТИКА,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Научная статья

УДК 004.052.4+004.94

DOI: 10.17072/1993-0550-2022-4-43-53

**Исследование дефектов моделей функциональных требований,
созданных студентами при объектно-ориентированном анализе****Наталья Николаевна Дацун**Пермский государственный национальный исследовательский университет, Пермь, Россия
nndatsun@inbox.ru, <https://orcid.org/0000-0001-8560-7036>

Аннотация. Оценка степени сформированности профессиональных компетенций ИТ-специалистов при выполнении заданий по практико-ориентированным дисциплинам является актуальной проблемой. Дефекты моделей функциональных требований, созданных студентами при объектно-ориентированном анализе, оказывают влияние на остальные этапы моделирования информационной системы. Целью является исследование дефектов в UCD. Задачами являются исследование распределения количества типов дефектов и их количества в диаграммах прецедентов с низким уровнем детализации, созданных студентами ИТ-направлений подготовки на этапе объектно-ориентированного анализа; изучение зависимости между дефектами различных типов в течение нескольких учебных годов с целью обнаружения влияния педагогического воздействия на плотность дефектов. Применены методы статистического анализа, использующие шкалы отношений. В результате исследования корпуса из 247 моделей за период наблюдений с 2016–2017 по 2021–2022 учебный год выявлено 2585 дефектов. Использована их классификация: лексические, синтаксические и семантические. В каждом классе выделены типы дефектов. В исследовании использованы показатели: количество типов дефектов и количество этих дефектов. В результате выявлено, что распределения показателей дефектов UCD не являются нормальными. Использование в учебном процессе учебно-методического пособия с примерами создания UCD, начиная с 2019–2020 учебного года, привело к устойчивому снижению количества дефектов, что подтверждено статистически значимым на уровне 95 % по критерию Вилкоксона–Манна–Уитни отличием показателей для пары учебных годов 2018–2019 и 2019–2020. Выявлена положительная средняя связь между количеством дефектов лексического и семантического классов в 2016–2017, 2017–2018 и 2021–2022 учебных годах. Положительная сильная связь выявлена среди лексических дефектов между типами, которые относятся к атрибутам одного элемента модели. Во всех остальных случаях не выявлено связи, которая может быть принята во внимание. Использованный подход может быть применен для анализа других видов UML-моделей студентов.

Ключевые слова: сформированность компетенций; объектно-ориентированный анализ; UCD с низким уровнем детализации; дефект модели; количество дефектов; закон распределения; корреляция; педагогический эффект

Для цитирования: Дацун Н. Н. Исследование дефектов моделей функциональных требований, созданных студентами при объектно-ориентированном анализе // Вестник Пермского университета. Математика. Механика. Информатика. 2022. Вып. 4(59). С. 43–53. DOI: 10.17072/1993-0550-2022-4-43-53.

Статья поступила в редакцию 10.10.2022; одобрена после рецензирования 20.10.2022; принята к публикации 01.11.2022.



Эта работа © 2022 Дацун Н. Н. лицензируется под CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <http://creativecommons.org/licenses/by/4.0/>

Research article

Investigation of Defects in Models of Functional Requirements Created by Students in Object-oriented Analysis

Nataliya N. Datsun

Perm State University, Perm, Russia

nmdatsun@inbox.ru, <https://orcid.org/0000-0001-8560-7036>

Abstract. Assessment of the degree of level of IT specialists professional competencies when performing tasks in practice-oriented disciplines is an actual problem. Defects in functional requirements models created by students in object-oriented analysis affect the rest of the information system modeling stages. The aim is to study defects in UCDs. The tasks are study of the distribution of the number of types of defects and their number in UCD with a low level of detail, created by IT-students at the stage of object-oriented analysis; study of the relationship between defects of various types over several academic years in order to detect the influence of pedagogical influence on the density of defects. Statistical analysis methods using relationship scales are applied. As a result of the study of a corpus of 247 models for the period from 2016–2017 to 2021–2022 academic year, 2585 defects were identified. Their classification is used: lexical, syntactic and semantic. Types of defects are identified in each class. The study used indicators: the number of defects types and the number of these defects. As a result, it was revealed that the distributions of UCD defects indicators are not normal distributions. The use of a teaching aid with examples of creating UCD in the educational process, starting from the 2019–2020 academic year, has led to a steady decrease in the number of defects. This is confirmed by a statistically significant 95 % difference in indicators for a pair of academic years 2018–2019 and 2019–2020 according to the Wilcoxon–Mann–Whitney criterion. A positive average relationship was found between the number of defects in the lexical and semantic classes in 2016–2017, 2017–2018 and 2021–2022 academic years. A positive strong relationship was found in the class of lexical defects between types that refer to the attributes of one element of the model. In all other cases, no connection has been identified that can be taken into account. The approach used can be applied to analyze other types of student UML models.

Keywords: *level competencies; object-oriented analysis; UCD with low level of detail; model defect; number of defects; distribution; correlation; pedagogical effect*

For citation: Datsun N. N. Investigation of Defects in Models of Functional Requirements Created by Students in Object-oriented Analysis // Bulletin of Perm University. Mathematics. Mechanics. Computer Science. 2022;4(59):43–53. (In Russ.). DOI: 10.17072/1993-0550-2022-4-43-53.

The article was submitted 10.10.2022; approved after reviewing 20.10.2022; accepted for publication 01.11.2022.

Введение

Унифицированный язык моделирования UML (Unified Modeling Language, UML) [1] предназначен для визуализации, спецификации и документирования систем, в которых большая роль принадлежит программному обеспечению [2]. Его используют в индустрии программного обеспечения для моделирования бизнесов и при объектно-ориентированном анализе и проектировании (Object Oriented Analysis and Design, OOAD) сложных информационных систем (ИС). Поэтому ИТ-индустрия ожидает наличие навыков объектно-ориентированного моделирования у будущих специалистов [3].

Однако различные исследования восприятия студентами моделирования на языке UML свидетельствуют, что этот процесс воспринимается новичками в OOAD как довольно сложный [4, 5].

Поэтому существует необходимость изучения дефектов, которые возникают в моделях, созданных студентами в процессе освоения OOAD, снижая их качество.

В исследованиях и ИТ-практики [6], и студенты [4] высоко оценивают полезность такого вида UML диаграмм как диаграммы прецедентов (Use Case Diagram, UCD) [1] для моделирования существенных аспектов системы. С другой стороны, студенты указывают, что моделирование вариантов использования было для них непросто [7].

При рассмотрении дефектов в моделях на языке UML используют различные классификации [5, 8, 9, 10, 11]. Их объединяет выделение в отдельные категории таких нарушений в моделях, как лексика, синтаксис и семантика.

Формулирование функциональных требований к ИС является одной из задач этапа

анализа, который выполняют студенты при изучении дисциплин по ООАД. Результатом является оформление функциональных требований в виде диаграмм прецедентов с низким уровнем детализации (Low Level of Detail, Low LoD) [12]. Объектом нашего исследования являются модели студентов в виде UCD с Low LoD, созданные за период с 2016–2017 по 2021–2022 учебный год, а предметом исследования – дефекты в этих моделях.

Чтобы понять и оценить влияние дефектов в моделях студентов, были сформулированы исследовательские вопросы.

ИВ1: Какой закон имеют распределения количества типов дефектов в UCD студентов и количество самих дефектов?

Для распределения количества типов дефектов была выдвинута нулевая и альтернативная гипотезы H_{10} и H_{11} :

– H_{10} : Распределение количества типов дефектов имеет нормальный закон распределения.

– H_{11} : Предположение о нормальности распределения количества типов дефектов противоречит имеющимся данным.

Аналогично выглядят нулевая и альтернативная гипотезы для распределения количества дефектов в UCD (общего и по классам дефектов).

ИВ2: Распределение дефектов моделей (количество их типов и количество самих дефектов) не зависит от года обучения студентов?

Для ответа на этот вопрос сформулированы статистические гипотезы:

– H_{20} : гипотеза об отсутствии различий между количеством типов дефектов в моделях 2016–2017 и 2017–2018 учебного года.

– H_{21} : гипотеза о значимости различий между количеством типов дефектов в моделях 2016–2017 и 2017–2018 учебного года.

Аналогично выглядят нулевая и альтернативная гипотезы для количества типов дефектов в UCD для всех пар учебных годов за период с 2017–2018 по 2021–2022 год.

Аналогично выглядят нулевая и альтернативная гипотезы для количества дефектов в UCD для всех пар учебных годов за период с 2016–2017 по 2021–2022 год.

ИВ3: Существует связь между количеством дефектов одного класса и этим показателем остальных классов?

ИВ4: Дефекты каких типов оказывают влияние на количество дефектов других типов моделей?

1. Методология

Используемая методология представляет собой тематическое исследование. Оно изучает специфичность возникновения дефектов в моделях с низким уровнем детализации, созданных студентами с 2016–2017 по 2021–2022 учебный год при описании функциональных требований к ИС.

1.1. Процесс

В UCD с Low LoD допустимы только акторы, прецеденты и ассоциации между акторами и прецедентами [13]. Проверка правильности моделей выполняется преподавателем, студентами-оценщиками при выполнении парного оценивания разработанного проекта или студентами/преподавателем с помощью системы [8].

1.2. Участники

Были использованы модели студентов:

1) трех направлений подготовки бакалавриата (третий и четвертый годы обучения), с 2016–2017 по 2021–2022 учебный год, очная и заочная формы обучения,

2) специалитета (третий год обучения), с 2016–2017 по 2020–2021 учебный год, очная форма обучения.

Создание моделей выполнялось индивидуально всеми студентами заочной формы обучения (с 2017–2018 по 2021–2022 учебные годы) и студентами очной формы одного направления подготовки (с 2016–2017 по 2019–2020 учебные годы). Остальные модели были созданы студентами при работе командами в составе из 2–3 человек. В 2019 г. по дисциплинам ООАД было подготовлено и применяется учебном процессе учебно-методическое пособие для проведения практических занятий и выполнения лабораторных работ. Это пособие содержит примеры создания моделей на языке UML.

1.3. Сбор данных

Корпус моделей получен из отчетов студентов, которые были загружены на проверку через элемент "Задание" в соответствующих курсах на платформе moodle. Всего рассмотрена 261 модель. 14 моделей не содержали ошибок, поэтому в дальнейшем исследовании участвовали 247 моделей.

Информация о корпусе моделей приведена в табл. 1.

Таблица 1. Распределение количества моделей по годам

Учебный год создания модели	Общее количество моделей	Количество моделей с дефектами
2016–2017	49	48
2017–2018	57	57
2018–2019	46	45
2019–2020	51	48
2020–2021	26	26
2021–2022	32	23
Всего:	261	247

Дефекты моделей идентифицировались согласно классификации, принятой в работах [8, 9, 13]: лексические, синтаксические и семантические.

Информация о дефектах была собрана для анализа по двум параметрам:

- 1) количество типов дефектов;
- 2) количество дефектов.

Поэтому в нашем исследовании применяются методы статистического анализа, использующие шкалы отношений.

В корпусе моделей функциональных требований с низким уровнем детализации обнаружен 21 тип дефектов, информация о которых приведена в табл. 2.

Этот список использован как чек-лист при проверке моделей на основе выбранного подхода CBR (чтения, основанного на списке требований, Checklist-Based Reading) [9, 14].

1.4. Анализ данных

На этапе описания данных корпуса моделей применен метод описательной статистики. Дополнительно для ответа ИВ1 выполнена оценка распределений количества типов дефектов и количества дефектов по асимметрии и эксцессу.

При изучении сходств и различий плотности дефектов за последовательные учебные периоды использованы статистические критерии Крамера–Уэлча и Вилкоксона–Манна–Уитни.

Корреляционный анализ применен на этапе исследования зависимостей между классами дефектов, а также зависимостей между типами дефектов внутри класса и между классами.

Обработка данных выполнена средствами табличного процессора, в том числе с использованием его пакета анализа.

2. Результаты

Для ответа на ИВ1 была получена описательная статистика исследуемого корпуса моделей, которая представлена в табл. 3.

Результаты оценки распределений по асимметрии и эксцессу приведены в табл. 4.

Результаты определения достоверности совпадений и различий для ответа на ИВ2 с помощью критериев Крамера–Уэлча и Вилкоксона–Манна–Уитни приведены в табл. 5.

Значения корреляции между количеством дефектов исследуемых классов для ответа на ИВ3 приведены табл. 6.

Для ответа на ИВ4 получены значения коэффициентов корреляции между количеством дефектов различных типов в внутри исследуемых классов, которые представлены табл. 7–9.

Полученные значения коэффициентов корреляции между количеством дефектов различных классов не представлены. Модуль значения всех этих коэффициентов меньше 0.28. Дополнительно были определены значения коэффициентов корреляции между количеством дефектов различных типов в внутри исследуемых классов и между количеством дефектов различных классов за каждый учебный год из исследуемого временного диапазона.

3. Обсуждение

94.6 % диаграмм прецедентов с LoD, созданных студентами за период с 2016–2017 по 2021–2022 учебный год, содержат дефекты. Из них 78.45 % составляют дефекты лексические, 18.26 % – семантические, 3.29 % – синтаксические.

С 95%-ным уровнем надежности можно утверждать, что среднее количество типов дефектов лексических в корпусе моделей составляет 2.23 ± 0.01 , синтаксических – 0.21 ± 0.004 , семантических – 1.18 ± 0.006 , всех типов – 3.63 ± 0.015 . Среднее количество дефектов составляет 8.21 ± 0.07 , 0.34 ± 0.007 , 1.94 ± 0.013 и 10.50 ± 0.074 соответственно.

Общее количество дефектов за исследуемый период уменьшилось с 798 до 100. Распределение количества дефектов каждого класса, обнаруженных в моделях за отдельный учебный год, изменилось следующим образом:

- 1) класс А – уменьшение с 87 % до 62 %,
- 2) класс В – увеличение с 1.5 % до 10 %,
- 3) класс С – увеличение с 11.5% до 28 %.

Таблица 2. Описание типов дефектов, их классификация и информация количестве дефектов соответствующего типа

Идентификатор типа дефекта	Описание дефекта	2016–2017 уч. г.	2017–2018 уч. г.	2018–2019 уч. г.	2019–2020 уч. г.	2020–2021 уч. г.	2021–2022 уч. г.
Класс А. Лексические ошибки							
"01.01"	Использование элемента не из UCD	0	0	0	0	6	8
"01.02"	Отсутствие границы системы	48	55	9	7	3	0
"01.03"	Название системы неверное	0	0	1	1	0	0
"01.04"	Название системы отсутствует	48	56	22	19	4	0
"01.05"	Неверный символ актора	16	17	4	0	0	0
"01.06"	Неверное имя актора	19	16	7	5	13	13
"01.07"	Неверный символ прецедента	92	37	0	1	2	0
"01.08"	Неверное название прецедента	28	21	65	66	26	26
"01.09"	Названия прецедентов перемешаны (глаголы и имена существительные)	32	45	36	8	9	0
"01.10"	Неверный символ ассоциации	411	426	150	116	19	15
Всего:		694	673	294	223	82	62
Класс В. Синтаксические ошибки							
"01.11"	Название системы неверно размещено	0	0	0	6	0	3
"01.12"	Актеры внутри системы	0	0	1	0	0	1
"01.13"	Система как актер	8	4	5	1	2	0
"01.14"	Прецедент за границей системы	0	0	0	1	0	0
"01.15"	Имена акторов не уникальные	0	2	0	2	0	0
"01.16"	Названия прецедентов не уникальные	0	0	0	2	4	0
"01.17"	Прецедент не имеет отношений с другими элементами диаграммы	1	9	8	7	0	6
"01.18"	Актер не имеет ни одного отношения ассоциации с прецедентами	3	7	1	0	1	0
Всего:		12	22	15	19	7	10
Класс С. Семантические ошибки							
"01.19"	Лишние акторы	57	63	61	68	23	2
"01.20"	Лишние прецеденты	14	17	13	28	2	4
"01.21"	Два и более акторов с одинаковыми наборами ассоциированных прецедентов	21	36	15	11	15	22
Всего:		92	116	89	107	40	28
Всего по классам А-С:		798	811	398	349	129	100

Таблица 3. Описательная статистика корпуса исследуемых моделей

Название показателя	Общее количество		Количество лексических		Количество синтаксических		Количество семантических	
	типов дефектов	дефектов	типов дефектов	дефектов	типов дефектов	дефектов	типов дефектов	дефектов
Среднее	3.63	10.50	2.23	8.21	0.21	0.34	1.18	1.94
Стандартная ошибка	0.12	0.59	0.10	0.55	0.03	0.06	0.05	0.10
Медиана	4	8	2	6	0	0	1	2
Мода	4	3	3	0	0	0	1	1
Стандартное отклонение (σ)	1.89	9.27	1.61	8.63	0.45	0.94	0.74	1.60
Дисперсия выборки	3.58	85.91	2.60	74.48	0.20	0.89	0.55	2.56
Экссесс (E)	-0.14	1.86	-0.88	1.59	2.92	25.51	-0.07	1.24
Асимметричность (A)	0.51	1.43	0.13	1.31	1.93	4.46	0.30	1.07
Минимум	1	1	0	0	0	0	0	0
Максимум	9	46	6	43	2	8	3	9
Сумма	897	2593	552	2028	53	85	292	480
Уровень надежности (95.0 %)	0.24	1.16	0.20	1.08	0.06	0.12	0.09	0.20
Межквартильный размах	2	3	1	1	0	0	1	1
1,33 σ	2.51	12.33	2.14	11.48	0.60	1.25	0.98	2.13
Размах	8	45	6	43	2	8	3	9
6 σ	11.34	55.61	9.67	51.77	2.69	5.64	4.44	9.60
Доля наблюдений в окрестности $\pm\sigma$	66.80 %	78.14 %	57.89 %	85.43 %	80.16 %	94.33 %	54.25 %	68.02 %
Доля наблюдений в окрестности $\pm 2\sigma$	96.76 %	93.12 %	97.98 %	94.33 %	98.38 %	96.36 %	95.95 %	97.17 %

Таблица 4. Оценки распределений дефектов по асимметрии и эксцессу

Название показателя распределения	Общее количество		Количество лексических		Количество синтаксических		Количество семантических	
	типов дефектов	дефектов	типов дефектов	дефектов	типов дефектов	дефектов	типов дефектов	дефектов
$ A \leq 3 \sqrt{D(A)}$ *	-	+	+	-	-	-	+	-
$ E \leq 5 \sqrt{D(E)}$ **	+	-	+	-	-	-	+	+

*D(A) – дисперсия асимметрии; **D(E) – дисперсия эксцесса

Таблица 5. Значения критериев Крамера–Уэлча и Вилкоксона–Манна–Уитни

Учебный год	Количество типов дефектов		Количество дефектов	
	T _{эмп}	W _{эмп}	T _{эмп}	W _{эмп}
2016–2017 и 2017–2018	2.39	3.83	1.23	2.17
2017–2018 и 2018–2019	5.81	3.78	2.92	2.74
2018–2019 и 2019–2020	1.05	2.66	1.06	1.92
2019–2020 и 2020–2021	1.25	0.01	1.47	1.01
2020–2021 и 2021–2022	0.65	1.10	0.17	1.14

Таблица 6. Корреляция между количеством дефектов различных классов

Учебный год	Значение коэффициента корреляции между количеством дефектов классов		
	A и B	A и C	B и C
2016–2017	0.28	0.38	-0.09
2017–2018	0.02	0.45	0.10
2018–2019	0.18	0.04	0.05
2019–2020	0.06	0.15	0.11
2020–2021	0.29	-0.20	-0.08
2021–2022	-0.16	0.31	0.01
За весь период	0.05	0.25	0.06

Таблица 7. Матрица корреляции количества дефектов в классе А

	"01.01"	"01.02"	"01.03"	"01.04"	"01.05"	"01.06"	"01.07"	"01.08"	"01.09"	"01.10"
"01.01"	1									
"01.02"	-0.24	1								
"01.03"	-0.02	-0.09	1							
"01.04"	-0.30	+0.80	-0.11	1						
"01.05"	-0.05	+0.16	-0.02	+0.12	1					
"01.06"	+0.02	+0.04	-0.03	+0.01	+0.18	1				
"01.07"	-0.07	+0.31	-0.03	+0.25	+0.15	-0.00	1			
"01.08"	+0.13	-0.18	-0.04	-0.10	-0.08	+0.07	-0.08	1		
"01.09"	+0.02	+0.20	-0.03	+0.20	+0.19	-0.05	-0.02	+0.19	1	
"01.10"	-0.13	+0.48	-0.06	+0.44	+0.22	-0.00	+0.28	-0.04	+0.09	1

Таблица 8. Матрица корреляции количества дефектов в классе В

	"01.11"	"01.12"	"01.13"	"01.14"	"01.15"	"01.16"	"01.17"	"01.18"
"01.11"	1							
"01.12"	-0.02	1						
"01.13"	-0.06	-0.03	1					
"01.14"	-0.01	-0.01	-0.02	1				
"01.15"	-0.02	-0.01	-0.03	+0.71	1			
"01.16"	-0.02	-0.01	-0.03	-0.01	-0.01	1		
"01.17"	-0.04	-0.02	-0.05	-0.01	-0.02	-0.02	1	
"01.18"	-0.03	-0.01	+0.04	-0.01	-0.01	-0.01	+0.28	1

Таблица 9. Матрица корреляции количества дефектов в классе С

	"01.19"	"01.20"	"01.21"
"01.19"	1		
"01.20"	+0.17	1	
"01.21"	-0.03	-0.05	1

Доля количества дефектов класса А среди всех ошибок за исследуемый период снижалась монотонно с 26.8 % до 2.4 %. Процент дефектов классов В и С также снизился с 0.5 % и 3.6 % в начале периода до 0.4 % и 1.1% соответственно в конце периода, однако монотонность снижения отсутствует.

Таким образом, в корпусе моделей UCD с низким уровнем детализации за период с 2016–2017 по 2021–2022 учебный год количество дефектов имеет тенденцию к снижению с одновременным изменением распределения количества дефектов в соответствующих классах.

Ответ на ИВ1. Значения показателей описательной статистики в табл. 3 свидетельствуют о том, что предположение о нормальности распределения всех вышеперечисленных распределений противоречит имеющимся данным. Нулевая гипотеза H_{10} отвергается.

Аналогично отвергаются нулевые гипотезы о распределении общего количества де-

фектов в UCD и по трем классам (лексических, синтаксических и семантических) дефектов.

Только два распределения (типов дефектов класса А и типов дефектов класса С) могут быть признаны близкими к нормальному.

Ответ на ИВ2. Полученные эмпирические значения из табл. 4 для критерия Крамера–Уэлча свидетельствуют о том, что характеристики сравниваемых выборок совпадают с уровнем значимости 0,05:

1) для количества типов дефектов для пар учебных годов, начиная с 2018–2019 учебного года до конца исследуемого периода;

2) для количества дефектов для пар учебных годов исследуемого периода, также начиная с 2018–2019 учебного года до конца исследуемого периода.

Поэтому нулевая гипотеза H_{20} отвергается так же, как и нулевая гипотеза для пары 2017–2018 и 2018–2019 учебного год.

Различия, которые не выявил критерий Крамера–Уэлча, были выявлены критерием Вилкоксона–Манна–Уитни. Полученные $W_{эмп}$ позволили уточнить, что следует принять нулевые гипотезы об отсутствии различий в моделях только:

1) между количеством типов дефектов для пар учебных годов, начиная с 2019–2020 учебного года до конца исследуемого периода;

2) между количеством дефектов для пар учебных годов, начиная с 2018–2019 учебного года до конца исследуемого периода.

Также можно утверждать, что достоверность различий по статистическому критерию Вилкоксона–Манна–Уитни равна 95 %:

1) для количества типов дефектов для всех пар учебных годов от 2016–2017 до 2019–2020;

2) для количества дефектов для всех пар учебных годов от 2016–2017 до 2018–2019.

Поэтому можно сделать вывод о том, что использование в учебном процессе с 2019–2020 учебного года по дисциплинам ООАД учебно-методического пособия с примерами для практического блока приводит к статистически значимым (на уровне 95 % по критерию Вилкоксона–Манна–Уитни) отличиям количества типов дефектов и количества дефектов для каждой пары учебных годов, подтверждая снижение значений этого показателя.

Ответ на ИВ3. Значения коэффициентов корреляции в табл. 6 показывают, что можно утверждать о существовании положительной заметной (средней) связи между количеством дефектов классов А и С в 2016–2017, 2017–2018 и 2021–2022 учебных годах. В остальных случаях связь слабая и очень слабая.

Наличие указанной выше средней положительной связи можно объяснить:

1) преобладанием ошибок классов А и С среди дефектов за весь исследуемый интервал наблюдений;

2) наибольшей долей ошибок этих классов в 2016–2017 и 2018–2019 учебных годах за весь период наблюдений;

3) снижением количества ошибок классов А на порядок за период с 2018–2019 по 2021–2022 учебный год с одновременным изменением распределения ошибок между классами А и С.

На слабую степень связи между классами А и С с классом В оказывает влияние небольшая доля синтаксических дефектов, обнаруженных в исследованных моделях.

Ответ на ИВ4. Из табл. 7 видно, что внутри класса А дефект типа "01.02" имеет положительную сильную связь с дефектом типа "01.04". Оба типа дефектов связаны с одним и тем же элементом модели (системой) и представляют собой отсутствие соответ-

ствующего атрибута этого элемента: границы (дефект типа "01.02") и названия (дефект типа "01.04"). В модели эти атрибуты тесно связаны, так как название системы размещается рядом с ее границей. Поэтому пропуск названия системы в неверных моделях в большинстве случаев сопровождался отсутствием у нее границы. Из табл. 7 также видно, что положительная средняя связь существует у дефекта типа "01.10" с "01.02" и "01.04". Можно утверждать, что одновременно при пропуске атрибутов системы в моделях использовался неверный символ ассоциации. Дополнительно были определены коэффициенты корреляции между количеством дефектов этих типов по учебным годам.

В 2016–2017 и 2017–2018 учебных годах все модели имели дефекты "01.02" и "01.04" и 91.6 % и 64.3 % из них соответственно также и дефект "01.10". И наоборот, в 2021–2022 учебном году ни одна из 23 моделей не имела дефекты "01.02" и "01.04" только одна (4.3 %) имела дефект "01.10".

Таким образом, в моделях правильное указание атрибутов системы и использование верного символа ассоциации связаны. Эти дефекты можно признать значимыми, так как "01.02" встречается в 122 работах (49 % корпуса моделей), "01.04" – в 149 работах (60 % соответственно) и "01.10" – в 108 (44 %). Между остальными типами дефектов класса А сила связи слабая и очень слабая. Это можно объяснить тем, что типы лексических дефектов относятся только к словарю модели, и связь между ними может проявиться только на уровне нескольких атрибутов одной лексемы (как в случае "системы" с дефектами типов "01.02" и "01.04").

Из табл. 8 видно, что внутри класса В дефект типа "01.14" имеет положительную сильную связь с дефектом типа "01.15". Однако признать эту зависимость значимой для исследованного корпуса моделей нельзя, так как "01.14" встречается в одной работе (0.40 % корпуса моделей), а "01.15" – в двух (0.81 % соответственно).

Также в этой таблице выявлена очень слабая положительная корреляция (+0,28) между близкими типами дефектов "01.17" и "01.18". Однако эти дефекты следует принять во внимание, так как "01.17" встречается в 10 работах (11,8 % моделей класса В) и "01.18" – в 7 работах (8.2 % соответственно). Между остальными типами дефектов класса В сила связи очень слабая.

Последнее утверждение относится ко всем типам дефектов в классе С, о чем свидетельствуют значения коэффициентов корреляции в табл. 9.

Заключение

Был исследован корпус из 261 модели функциональных требований, созданных студентами при объектно-ориентированном анализе в виде UCD с низким уровнем детализации за учебный период с 2016–2017 по 2021–2022 учебный год. 247 моделей имели дефекты, анализ которых проводился по трем классам (лексические, синтаксические, семантические) и типам дефектов внутри указанных классов. Для оценки использованы показатели: количество типов дефектов и количество дефектов. При исследовании применены метод описательной статистики и методы статистического анализа для величин, измеренных в шкалах отношений. Исследование привело к следующим результатам:

1. Количество дефектов за период наблюдений имеет тенденцию к снижению (почти в 8 раз). Наибольшее сокращение количества дефектов произошло среди лексических (с 87 % до 62 %), которые являлись преобладающими весь период.

2. Законы распределения количества типов дефектов и количества дефектов не являются нормальными. Близкими к нормальному могут быть признаны только распределение типов лексических и семантических дефектов.

3. Использование в учебном процессе в 2019–2020 учебном году учебно-методического пособия с примерами подготовки моделей на языке UML привело к статистически значимым на уровне 95 % по критерию Вилкоксона–Манна–Уитни отличиям количества типов дефектов и количества дефектов для пары учебных годов 2018–2019 и 2019–2020. Начиная с 2019–2020 учебного года до конца исследуемого периода отличия не наблюдаются. Поэтому можно утверждать, что использование пособия приносит значимый педагогический эффект с момента начала его применения.

4. Существует положительная средняя связь между количеством дефектов лексического и семантического классов в 2016–2017, 2017–2018 и 2021–2022 учебных годах. В остальных случаях связь не может быть принята во внимание. Степень влияния синтакси-

ческих дефектов невелика, так как их доля составляет около 3 %.

5. Среди лексических дефектов выявлена положительная сильная связь только между теми типами, которые ассоциированы с атрибутами одного и того же элемента модели. Сильной связи среди синтаксических и семантических дефектов не выявлено так же, как и между дефектами различных классов.

Подход, использованный в данном исследовании, может быть применен для анализа других видов моделей, созданных студентами на этапе объектно-ориентированного анализа для снижения вероятности внедрения дефектов в модели и/или повышения вероятности раннего обнаружения дефектов при моделировании.

В качестве направлений дальнейших исследований необходимо изучить:

1) зависимости показателей дефектов модели от сложности последних;

2) зависимости изменения показателей дефектов модели при переходе к UCD с высоким уровнем детализации.

Список литературы

1. *OMG® Unified Modeling Language® (OMG UML®) Version 2.5.1*. URL: <https://www.omg.org/spec/UML/2.5.1/PDF> (дата обращения: 22.09.2022).
2. Буч Г., Рамбо Д., Якобсон И. Язык UML. Руководство пользователя. М.: ДМК Пресс, 2006. 496 с.
3. Bourque P., Fairley R.E. Guide to the software engineering body of knowledge (SWEBOK®): Version 3.0. Washington: IEEE Computer Society Press. 2014. URL: <https://www.computer.org/education/bodies-of-knowledge/software-engineering> (дата обращения: 22.09.2022).
4. Boberić-Krstićev D., Tešendić D. Experience in teaching OOAD to various students. *Informatics in Education*. 2013. 12 (1). P. 43–58. URL: <https://eric.ed.gov/?id=EJ1064307> (дата обращения: 22.09.2022).
5. *Insights in students' problems during UML modeling* / R. Reuter [et al.] // *Proceeding of the IEEE Global Engineering Education Conference*. 2020. art. no. 9125110 P. 592–600. URL: <https://ieeexplore.ieee.org/document/9125110> (дата обращения: 22.09.2022).
6. Anke J., Bente S. UML in der Hochschullehre: Eine kritische Reflexion // *CEUR Workshop*

- Proceedings. München: 2019. Vol. 2358, art. № 01. P. 8–20. URL: <http://ceur-ws.org/Vol.2358/paper-01.pdf> (дата обращения: 22.09.2022).
7. *Lopes A., Steinmacher I., Conte T.* UML Acceptance: Analyzing the Students' Perception of UML Diagrams // Proceedings of the XXXIII Brazilian Symposium on Software Engineering. 2019. P. 264–272. URL: <https://dl.acm.org/doi/10.1145/3350768.3352575> (дата обращения: 22.09.2022).
 8. *Validation Automation of UML Diagrams Created by Students / T.S. Gasheva [et al.] // Trudy ISP RAN/Proc. ISP RAS.* 2021. Vol. 33, № 4. P. 7–18. DOI: 10.15514/ISPRAS-2021-33(4)-1.
 9. *Отинов А.В., Дацун Н.Н.* Автоматизация проверки UCD студентов // Актуальные проблемы математики, механики и информатики: сб. статей по материалам студ. конф. / Перм. гос. нац. исслед. ун-т. Пермь, 2021. С. 107–114.
 10. *Bolloju N., Leung F.S.K.* Assisting novice analysts in developing quality conceptual models with UML. Communications of the ACM. 2006. Vol. 49, № 7. P. 108–112. doi: 10.1145/1139922.1139926.
 11. *Mistakes in UML Diagrams: Analysis of Student Projects in a Software Engineering Course / S. Chren [et al.] // Software Engineering Education and Training: proc. of the IEEE/ACM 41st Int. Conf. on Software Engineering.* 2019. P. 100–109. doi: 10.1109/ICSE-SEET.2019.00019.
 12. *Does the level of detail of UML diagrams affect the maintainability of source code?: a family of experiments / A.M. Fernández-Sáez [et al.] // Empirical Software Engineering.* 2016. Vol. 21, № 1. P. 212–259. URL: <https://dl.acm.org/doi/10.1007/s10664-014-9354-4> (дата обращения: 22.09.2022).
 13. *Дацун Н.Н.* Моделирование информационных систем. Практикум [Электронный ресурс]: учеб.-метод. пособие. Пермь: Перм. гос. нац. исслед. ун-т, 2022. URL: <http://www.psu.ru/files/docs/science/books/uchebnie-posobiya/Dacun-Modelirovanie-Informacionnyh-Sistem-Praktikum.pdf> (дата обращения: 22.09.2022).
 14. *Naveed A., Ikram N.* A Novel Checklist: Comparison of CBR and PBR to Inspect Use Case Specification // Requirements Engineering in the Big Data Era. Communications in Computer and Information Science. Vol. 558. Berlin, Heidelberg: Springer, 2015. URL: https://doi.org/10.1007/978-3-662-48634-4_8 (дата обращения: 22.09.2022).
- ## References
1. *OMG® Unified Modeling Language® (OMG UML®) Version 2.5.1.* URL: <https://www.omg.org/spec/UML/2.5.1/PDF> (accessed: 22.09.2022).
 2. *Buch G., Rambo D., Yakobson I.* Yazyk UML. Rukovodstvo pol'zovatelya. M.: DMK Press; 2006. 496 s. (In Russ.).
 3. *Bourque P., Fairley R.E.* Guide to the software engineering body of knowledge (SWEBOK®): Version 3.0. Washington: IEEE Computer Society Press; 2014. URL: <https://www.computer.org/education/bodies-of-knowledge/software-engineering> (accessed: 22.09.2022).
 4. *Boberić-Krstićev D., Tešendić D.* Experience in teaching OOAD to various students. Informatics in Education. 2013;12(1):43–58. URL: <https://eric.ed.gov/?id=EJ1064307> (accessed: 22.09.2022).
 5. *Reuter R. [et al.]* Insights in students' problems during UML modeling. Proceeding of the IEEE Global Engineering Education Conference. 2020. (9125110):592–600. URL: <https://ieeexplore.ieee.org/document/9125110> (accessed: 22.09.2022).
 6. *Anke J., Bente S.* UML in der Hochschullehre: Eine kritische Reflexion. CEUR Workshop Proceedings. München. 2019. (2358:01):8–20. URL: <http://ceur-ws.org/Vol-2358/paper-01.pdf> (accessed: 22.09.2022).
 7. *Lopes A., Steinmacher I., Conte T.* UML Acceptance: Analyzing the Students' Perception of UML Diagrams. Proceedings of the XXXIII Brazilian Symposium on Software Engineering. 2019. 264–272. URL: <https://dl.acm.org/doi/10.1145/3350768.3352575> (accessed: 22.09.2022).
 8. *Gasheva T.S. [et al.]* Validation Automation of UML Diagrams Created by Students. Trudy ISP RAN/Proc. ISP RAS. 2021. (33:4):7–18. DOI: 10.15514/ISPRAS-2021-33(4)-1.
 9. *Otinov A.V., Datsun N.N.* Avtomatizaciya proverki UCD studentov // Aktual'nye problemy ma-tematiki, mekhaniki i informatiki: sb. statej po materialam stud. konf. Perm. gos. nac. issled. un-t. Perm', 2021. P. 107–114. (In Russ.).
 10. *Bolloju N., Leung F.S.K.* Assisting novice analysts in developing quality conceptual models with UML. Communications of the ACM.

- 2006;(49:7):108–112.
doi:10.1145/1139922.1139926.
11. *Chren S.* [et al.] Mistakes in UML Diagrams: Analysis of Student Projects in a Software Engineering Course. Software Engineering Education and Training: proc. of the IEEE/ACM 41st Int. Conf. on Software Engineering. 2019.100–109. doi: 10.1109/ICSE-SEET.2019.00019.
 12. *Fernández-Sáez A.M.* [et al.] Does the level of detail of UML diagrams affect the maintainability of source code?: a family of experiments. Empirical Software Engineering. 2016;(21:1):212–259. URL: <https://dl.acm.org/doi/10.1007/s10664-014-9354-4> (accessed: 22.09.2022).
 13. *Datsun N.N.* Modelirovanie informacionnyh sistem. Praktikum [Elektronnyj resurs]: uchebn.-metod. posobie. Perm': Perm. gos. nac. issled. un-t, 2022. URL: <http://www.psu.ru/files/docs/science/books/uchebnie-posobiya/Dacun-Modelirovanie-Informacionnyh-Sistem-Praktikum.pdf> (accessed: 22.09.2022). (In Russ.).
 14. *Naveed A., Ikram N.* A Novel Checklist: Comparison of CBR and PBR to Inspect Use Case Specification. Requirements Engineering in the Big Data Era. Communications in Computer and Information Science. Vol. 558. Berlin, Heidelberg: Springer, 2015. URL: https://doi.org/10.1007/978-3-662-48634-4_8 (accessed: 22.09.2022).

Информация об авторе:

Н. Н. Дацун – кандидат физико-математических наук, доцент, доцент кафедры математического обеспечения вычислительных систем механико-математического факультета Пермского государственного национального исследовательского университета (614068, г. Пермь, ул. Букирева, 15), AuthorID 734216.

Information about the author:

N. N. Datsun – Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Department of Computing Systems Mathematical Support of the Mechanics and Mathematics Faculty, Perm State University (Bukireva street, 15, Perm, 614068), AuthorID 734216.