

ИСПОЛЬЗОВАНИЕ ТЕОРИИ ВРЕМЕННЫХ РЯДОВ ДЛЯ ВЫДЕЛЕНИЯ ВТОРИЧНЫХ ОШИБОК НА ЭТАПЕ ТЕСТИРОВАНИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Для виділення вторинних помилок із загального потоку помилок при тестуванні програмного забезпечення запропоновано використовувати теорію часових рядів. Показано, що математичний апарат цієї теорії дає змогу відокремлювати вторинні помилки із загального потоку.

Для выделения вторичных ошибок из общего потока ошибок при тестировании программного обеспечения предложено использовать теорию временных рядов. Показано, что математический аппарат этой теории позволяет выделять вторичные ошибки из общего потока.

For the selection of secondary errors from the general thread of errors at testing of software it is suggested to use the theory of temporal rows. It is shown that the mathematical vehicle of this theory allows to distinguish secondary errors from a general thread.

Введение. В настоящее время компьютерные информационные системы (ИС) получили широкое применение во всех областях человеческой деятельности. Вместе с несомненными выгодами от их использования растут и негативные эффекты, связанные с тем, что человечество становится заложником их надежности и безопасности.

К сожалению, гарантировать отсутствие ошибок нельзя ни для одной программы, потому вопрос прогнозирования их количества и интенсивности является актуальным. В процессе исправления ошибок в программном обеспечении (ПО) ИС могут быть внесены дополнительные ошибки, которых не было в ПО в начале тестирования. Такие ошибки получили название вторичных.

Негативное влияние вторичных ошибок приводит к невозможности точного моделирования процесса выявления из-за систематического прогрессирующего отклонения реальной кривой интенсивности проявления ошибок от прогнозной. Проблема моделирования надежности программного обеспечения с учетом возможного появления вторичных ошибок на настоящее время является не решенной, сложной и неоднозначной.

Под надежностью ПО понимается его способность безотказно выполнять опреде-

ленные функции при заданных условиях в течение заданного периода времени с достаточно большой вероятностью достоверности [4].

Для прогнозирования изменения показателей надежности с течением времени используются модели надежности программного обеспечения (МНПО) [7]. На сегодняшний день насчитывается около тридцати различных моделей и их модификаций, которые отличаются подходами к процессу моделирования и начальными предположениями. Но ни одна из них не позволяет выделять и прогнозировать вторичные ошибки. Поэтому задача их выделения на этапе тестирования ПО ИС является актуальной.

Процесс выявления вторичных ошибок развивается во времени, поэтому для его анализа можно использовать теорию временных рядов.

Под временным рядом понимается множество равноудаленных (или с переменным периодом) наблюдений, которые характеризуют поведение во времени процесса или объекта на выбранном (заданном) временном интервале. Модель временного ряда представляет собой уравнение, связывающее наблюдение, полученное в некоторый конкретный момент времени, с наблюдениями, полученными ранее по той же и (или) другим характеристикам изучаемой переменной [5].

Анализ проблем выявления вторичных ошибок. Все модели надежности создаются с единственной целью – прогнозирования показателей надежности, т.е. с помощью определенной модели получим возможность предсказывать количество обнаруженных ошибок в будущем. Но из того, что вторичные ошибки являются непредсказуемыми, следует, что верно их учесть не сможет ни одна МНПО. При внесении в ПО значительных изменений, количество дополнительно внесенных дефектов будет больше, чем при простом исправлении существующих. Кроме того, за счет изменения параметров ПО изменится сам закон изменения распределения выявления ошибок во времени. Для использования МНПО в этом случае нужно заново определять эти параметры и выполнять прогнозирование с уже новыми их значениями.

Если допустить, что при исправлении существующих ошибок вторичные ошибки не вносятся, то кривая распределения интенсивности их проявления во времени будет подчиняться одному и тому же закону [1] (образует тренд). Это дает возможность использовать аппарат временных рядов для выделения вторичных ошибок.

Тренд – это долговременная тенденция изменения исследуемого временного ряда. Тренды могут быть описаны различными уравнениями – линейными, логарифмическими, степенными и т. д. Фактический тип тренда устанавливают на основе подбора его функциональной модели статистическими методами либо сглаживанием исходного временного ряда [2].

Пример использования. В качестве примера рассмотрим использование аппарата временных рядов в случае появления вторичных дефектов в ПО учетной ИС (УИС) «Агрокомплекс» [3].

Динамика выявления ошибок в УИС «Агрокомплекс» приведена в табл. 1. Далее на рис. 1 приведен график изменения количества ошибок по результатам эксперимента.

Из рис. 1 видно, что закон изменения количества ошибок изменялся трижды. Эти изменения отвечают значительным переработкам системы, о которых говорилось выше.

1. Динамика выявления ошибок

| Кол-во дней с начала эксплуатации | Число ошибок | Кол-во дней с начала эксплуатации | Число ошибок |
|-----------------------------------|--------------|-----------------------------------|--------------|
| 0 | 0 | 1185 | 456 |
| 30 | 72 | 1215 | 473 |
| 58 | 133 | 1307 | 480 |
| 90 | 156 | 1491 | 487 |
| 211 | 176 | 1519 | 503 |
| 243 | 226 | 1551 | 518 |
| 365 | 275 | 1581 | 520 |
| 396 | 308 | 1612 | 527 |
| 485 | 313 | 1642 | 537 |
| 516 | 321 | 1673 | 544 |
| 546 | 335 | 1704 | 548 |
| 608 | 350 | 1734 | 561 |
| 638 | 351 | 1765 | 565 |
| 730 | 355 | 1795 | 568 |
| 789 | 368 | 1826 | 571 |
| 881 | 392 | 1857 | 574 |
| 911 | 398 | 1916 | 577 |
| 973 | 402 | 1946 | 578 |
| 1003 | 410 | 2099 | 580 |
| 1034 | 420 | 2130 | 583 |
| 1064 | 428 | 2160 | 584 |
| 1126 | 447 | 2191 | 585 |

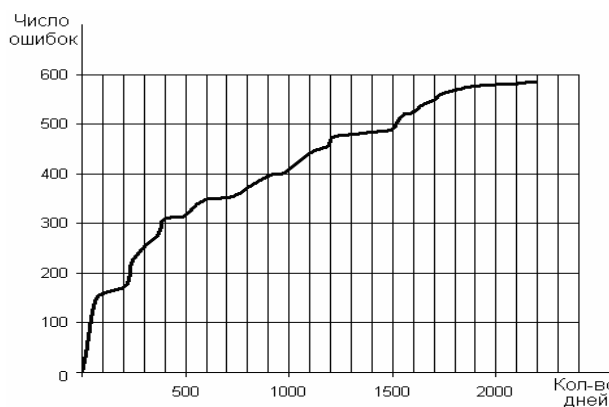


Рис. 1. Экспериментальная кривая

В результате этих изменений корректировались структурные параметры системы и, соответственно, параметры модели. Поэтому моделирование проводилось отдельно для каждого периода с постоянным законом изменения. Для удобства сравнения результатов для моделирования на втором и третьем интервалах осуществлен перенос начала координат в начальную точку интервала.

Для выделения тренда первичных ошибок производилась аппроксимация экспери-

ментальной зависимости по методу наименьших квадратов (МНК) [6], полученный временной ряд обозначен как $a(t)$. Для выделения вторичных ошибок (ряд $F(t)$) из экспериментального временного ряда $x(t)$ используется формула $F(t) = x(t) - a(t)$.

Результаты расчетов для трех последовательных временных интервалов приведены в таблицах 2, 3 и 4. Соответствующие им графики $F(t)$ показаны на рис. 2, 3 и 4.

2. Интервал 1

| t | x(t) | a(t) | F(t) |
|-----|------|----------|----------|
| 0 | 0 | 0 | 0 |
| 30 | 72 | 53,54915 | 18,45085 |
| 58 | 133 | 96,77705 | 36,22295 |
| 90 | 156 | 139,3095 | 16,6905 |
| 211 | 176 | 250,5796 | -74,5796 |
| 243 | 226 | 270,3714 | -44,3714 |
| 365 | 275 | 322,4598 | -47,4598 |
| 396 | 308 | 331,3589 | -23,3589 |
| 485 | 313 | 350,4236 | -37,4236 |
| 516 | 321 | 355,3076 | -34,3076 |
| 546 | 335 | 359,3652 | -24,3652 |
| 608 | 350 | 366,0485 | -16,0485 |
| 638 | 351 | 368,61 | -17,61 |

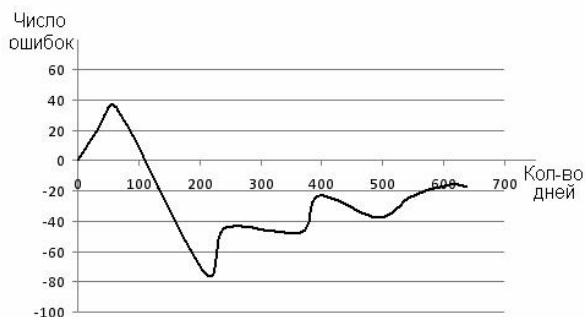


Рис.2. График отклонений на участке 1

Анализ полученных результатов. График отклонений на первом участке не показывает систематической зависимости от времени. Отклонения носят случайный характер и могут объясняться погрешностями, возникшими из-за недостаточно организованного процесса фиксации выявленных ошибок.

Отклонения на втором участке целиком лежат в отрицательной области. Это свидетельствует о том, что в соответствующем временном интервале фиксировались не все возникающие ошибки.

3. Интервал 2

| t | x(t) | a(t) | F(t) |
|-----|------|----------|----------|
| 0 | 0 | 0 | 0 |
| 59 | 13 | 34,09198 | -21,092 |
| 151 | 37 | 74,1131 | -37,1131 |
| 181 | 43 | 84,41646 | -41,4165 |
| 243 | 47 | 102,3268 | -55,3268 |
| 273 | 55 | 109,5903 | -54,5903 |
| 304 | 65 | 116,2749 | -51,2749 |
| 334 | 73 | 122,0356 | -49,0356 |
| 396 | 92 | 132,0496 | -40,0496 |
| 455 | 101 | 139,6201 | -38,6201 |
| 485 | 118 | 142,8656 | -24,8656 |
| 577 | 125 | 150,7952 | -25,7952 |

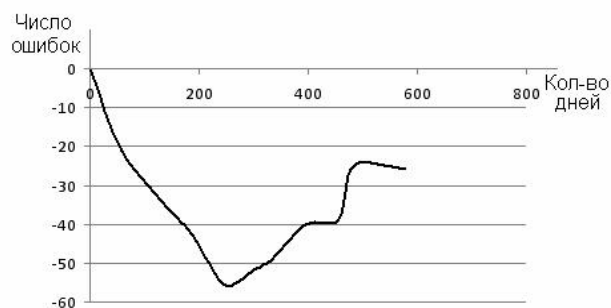


Рис.3. График отклонений на участке 2

4. Интервал 3

| t | x(t) | a(t) | F(t) |
|-----|------|----------|----------|
| 0 | 0 | 0 | 0 |
| 28 | 16 | 10,26943 | 5,73057 |
| 60 | 31 | 20,74293 | 10,25707 |
| 90 | 33 | 29,47004 | 3,52996 |
| 121 | 40 | 37,50146 | 2,49854 |
| 151 | 50 | 44,42298 | 5,57702 |
| 182 | 57 | 50,79273 | 6,20727 |
| 213 | 61 | 56,45464 | 4,54536 |
| 243 | 74 | 61,3341 | 12,6659 |
| 274 | 78 | 65,82459 | 12,17541 |
| 304 | 81 | 69,69451 | 11,30549 |
| 335 | 84 | 73,25594 | 10,74406 |
| 366 | 87 | 76,4216 | 10,5784 |
| 425 | 90 | 81,50727 | 8,49273 |
| 455 | 91 | 83,68751 | 7,31249 |
| 608 | 93 | 91,6479 | 1,3521 |
| 639 | 96 | 92,76973 | 3,23027 |
| 669 | 97 | 93,73653 | 3,26347 |

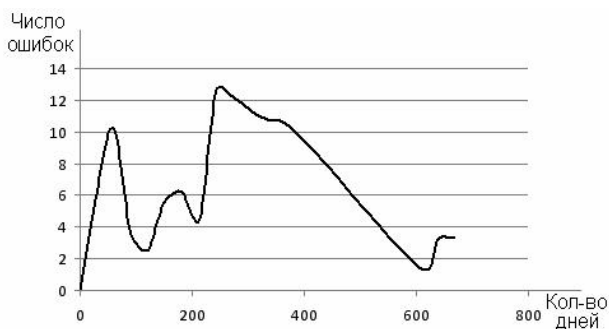


Рис.4. График отклонений на участке 3

На третьем интервале отклонения положительны, что свидетельствует о внесении в программное обеспечение вторичных ошибок. На этом этапе, как следует из таблицы 4, в ПО было внесено 24 вторичные ошибки и выявлено 94 первичные. Следовательно, количество внесенных вторичных ошибок составляет около 24 % от их общего числа ошибок. Этот результат хорошо соотносится с данными, полученными в [2] другими методами.

Выводы. Применение теории временных рядов позволяет разделить общий временной ряд на две составляющие – первичные ошибки, закон изменения которых подчиняется строгим законам, и вторичные, которые вносятся в ПО случайным образом.

Имеется возможность оценки общего количества внесенных в ПО вторичных ошибок. Эти данные могут использоваться для оценки качества тестирования и квалификации разработчиков ПО.

Кроме того, использование временных рядов позволяет оценить точность выполненного тестирования и достоверность полученных результатов.

Список использованной литературы

1. Бокс Дж. Анализ временных рядов / Бокс Дж., Дженкинс Л (в 2-х томах). – М.: Мир, 1972. – С.
2. Маевский Д.А. Структурная динамика программных систем и прогнозирования их надежности при наличии вторичных дефектов. / Д.А. Маевский // Радиоэлектронные и компьютерные системы, НТЖ. – Харьков: ХАИ. – 2010. – № 3. – С.197-203.
3. Маевский Д.А. Информационная система «Агрокомплекс» для бухгалтерского и оперативного учета в сельском хозяйстве / Д.А. Маевский, Т.Я. Тинтулова, С.Г. Анто-

щук // Аграрный вестник Причерноморья. Технич.науки. – 2009. – Вип. 48. – С. 151 - 156.

4. Оценка надежности программного обеспечения на различных этапах жизненного цикла различных программ / В.Г. Промыслов, А.В. Антонов, С.И. Масолкин, А.С. Степанянец // Труды V международной конференции «Идентификация систем и задачи управления» SCIPRO'06. Москва, 30 января – 2.02. 2006. – М.: Ин-т проблем управления им.В.А. Трапезникова, РАН. – 2006. – С. 1300-1304.

5. Харченко В.С. Методы моделирования и оценки качества и надежности программного обеспечения / В.С. Харченко, В.В. Скляр, О.М. Тарасюк. //Уч. пособие. – Харьков: Нац. аэрокосм. ун-т «ХАИ», 2004. – 159 с.

6. Шор. Я. Б. Статистические методы анализа и контроля качества и надежности / Я.Б.Шор. – М.: Госэнергоиздат. – 1962. – С. 92-98.

7. Lyu M.R. Handbook of Software Reliability Engineering / Lyu M.R. McGraw-Hill Company, 1996. – 805 p.

Получено 12.01.2011



Маевский
Дмитрий Андреевич,
канд. техн. наук, доцент,
зав. каф. теоретич. основ
и общ. электротехники
Одесск. нац. политехн.
ун-та
тел. (048) 734-84-54



Жеков
Олег Прокопьевич,
ассистент, аспирант той
же кафедры
тел. (048) 734-80-25