

РАЗРАБОТКА АЛГОРИТМИЧЕСКИХ И ПРОГРАММНЫХ СРЕДСТВ ДЛЯ ОЦЕНКИ ДОВЕРИЯ К РЕЗУЛЬТАТАМ НАГРУЗОСНОГО ТЕСТИРОВАНИЯ ШЛЮЗОВ БЕЗОПАСНОСТИ

*К.С. Сунцов, студент, гр. 8Б71
Томский политехнический университет
E-mail: kss20@tpu.ru*

Введение

При разработке программного обеспечения для шлюзов безопасности проводится несколько видов тестирования программного обеспечения (ПО), в том числе нагрузочное тестирование. При нагрузочном тестировании объект тестирования подвергается нагрузке генератором трафика и снимаются показатели производительности, главным образом - пропускная способность, количество запросов в секунду, задержка.

Результаты тестирования отражают качество разработанного программного обеспечения и позволяют руководителю проекта принять решение о возможности публикации очередной версии ПО. Соответственно, требуется обеспечить достоверность результатов тестирования. Целью данной работы является анализ достоверности результатов нагрузочного тестирования.

Описание алгоритма

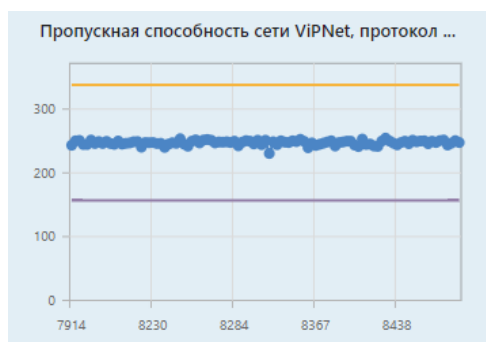
Для решения задачи анализа достоверности результатов тестирования было разработано веб-приложение ASP.NET в среде JetBrains Rider на языке C#. Достоверность результатов определяется методами математической статистики [1].

Данные результатов изменений извлекают из базы данных [2], после чего они группируются по наборам параметров тестирования: в одну группу попадают результаты, полученные при одинаковой версии ПО, одинаковой методике тестирования, одинаковых настройках объекта тестирования и тестового окружения.

Из полученных групп выбирается исследуемая группа и исследуемый показатель производительности, после чего определяются статистические характеристики распределения, а именно функция и плотность распределения, а также числовые параметры распределения: математическое ожидание и среднеквадратическое отклонение. Затем строятся четыре графика:

- графическое представление значений показателя (график «значение-дата измерения», рис.1,а);
- функция распределения (рис.1,б);
- плотность вероятности (рис.1,в);
- результаты сравнения эмпирической кривой распределения с нормальным законом (рис.1,г).

Как видно из рисунка 1,г, он содержит эмпирическую плотность распределения и соответствующую ей теоретическую плотность нормального распределения. Параметры нормального распределения совпадают с параметрами выборки значений показателя, а эмпирическое распределение относительных частот формируется из гистограммы выборки.



а)



б)

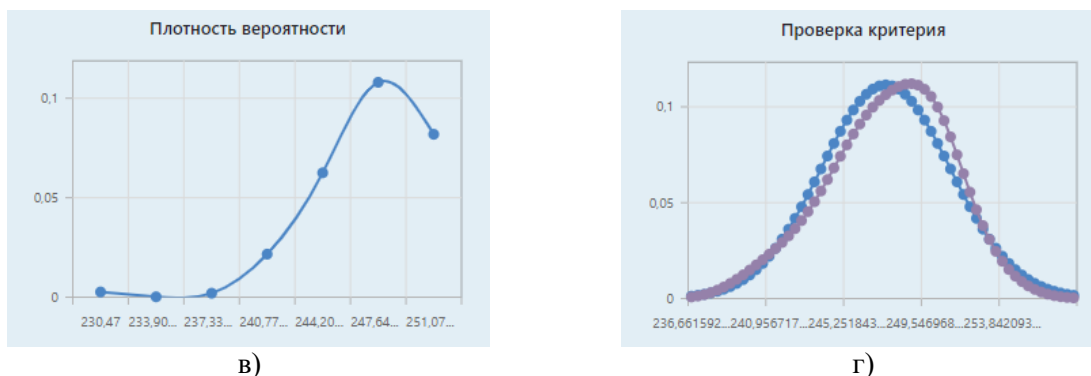


Рис. 1. Пример строящихся графиков: а) – график «значение-дата измерения»; б) – функция распределения; в) – плотность вероятности; г) - график «проверка критерия»

Последний график (рис.1,г) также отражает достаточность и достоверность данных. Так, при малом объеме выборки мы получим несовпадение линий. Также интересной является ситуация, когда при большом объеме выборки возникает несколько экстремумов. В этом случае можно сделать вывод, что объект тестирования принимает несколько состояний, что свидетельствует либо о нестабильности объекта тестирования, либо о неучтенных в системе параметрах тестирования (то есть группы значений были сформированы неверно).

Далее определяется доверительный интервал. Для этого значения эмпирического распределения сравниваются с помощью критерия «хи-квадрат» со значениями нормального распределения, и в случае удовлетворения условиям критерия, доверительный интервал находится для данного нормального распределения на основе квантиля. В противном случае – как $M \pm \sigma$ и в зависимости от оценки эксперта рекомендуется провести дополнительное тестирование.

Также с помощью критерия «хи-квадрат» оценивается совпадение распределения выборки с гамма-распределением, непрерывным одномерным распределением, распределениями Гумбеля и Пуассона. Полученные результаты сравнения используются для исследовательских целей.

Тестирование алгоритма

В базе данных изначально не было достаточного количества результатов однотипных тестирований, поэтому были проведены 200 замеров с помощью специально разработанного тестировочного робота – по 100 замеров для двух наборов параметров, один из которых представлен на рисунке 1. Значения критерия «хи-квадрат» для данного случая был 4,45 при критическом значении 0,05[3]. Поскольку эмпирическая плотность распределения имеет один экстремум, и нет причин сомневаться в правильности тестирования, рекомендуется провести дополнительное тестирование.

Вторая группа значений показала значение критерия «хи-квадрат» равно $7 \cdot 10^{-14}$ при критическом значении 0,05, при этом в доверительный интервал (2,265; 0,457) попали лишь 27% результатов, но они обладают высокой достоверностью, и могут быть приняты во внимание при оценке производительности данной версии ПО при данных параметрах.

Заключение

Результатом данного исследования является алгоритм и его реализация, способные оценить достаточность результатов тестирования, их достоверность, наличие проблем в методике, либо объекте тестирования.

Данные считаются достаточными при следующих условиях:

1. Все параметры объекта тестирования и тестового окружения при измерениях совпадали.
2. Количество значений позволяет сделать вывод о нормальности распределения с помощью критерия «хи-квадрат» с уровнем значимости 0,05.
3. Эмпирическая плотность распределения имеет один экстремум.

Список использованных источников

1. Лагутин М.Б. Наглядная математическая статистика. — М.: Бином. Лаборатория знаний, 2009. — 472 с.
2. Jon P Smith. Entity Framework Core in Action – «Manning Publications», 2018. – 486 с.
3. Выбор уровня значимости при проверке статистических гипотез. [Электронный ресурс]. –URL: <https://www.statmethods.ru/stati/vybor-urovnya-znachimosti-pri-proverke-statisticheskikh-gipotez/> (дата обращения 05.03.2021).