

ВСЕГДА ЛИ ПРИМЕНИМО «МАГИЧЕСКОЕ ЧИСЛО 5» ПРИ ПРОВЕДЕНИИ ЮЗАБИЛИТИ-ТЕСТОВ?

Якоб Нильсен и Томас Ландауэр, Роберт Вирзи, Джеймс Льюис в разные годы опубликовали статьи, посвященные определению размера выборки при проведении юзабилити-тестов. В своих статьях они представили математическую модель определения вероятности обнаружения юзабилити-проблем. С помощью этой вероятности они доказали возможность определения размера выборки, необходимой для выявления необходимой части проблем в интерфейсе. Автор предоставил эмпирические данные для своих моделей и сделали несколько важных выводов.

Ключевые слова: юзабилити, тест, тестирование, проблема, выборка, вероятность, размер выборки.

Юзабилити-тестирование – довольно дорогое удовольствие. И у разработчиков программного обеспечения, владельцев сайтов возникает естественное желание обнаружить как можно большее количество проблем при минимальном количестве тестов.

Работы в этом направлении ведутся уже давно. Один из главных вопросов – каково оптимальное количество участников теста? Споры по этому вопросу длятся уже более 20 лет.

Роберт Вирзи [2] в 1992 г. опубликовал эмпирические данные, доказавшие возможность использования вероятности биномиального распределения для оценки размера выборки. Он опубликовал три эксперимента, в каждом из которых вычислил вероятность обнаружения проблемы. Была рассчитана вероятность обнаружения проблемы для каждого участника экспертной группы, вычисленная как отношение количества проблем, обнаруженных данным участником, к общему количеству проблем, обнаруженных при проведении тестирования. Итоговая вероятность обнаружения проблемы была вычислена путем усреднения всех вероятностей.

Пример проведения теста приведен в табл. 1 [1]. Здесь «1» означает, что участник обнаружил данную проблему. В этом гипотетическом примере все участники обнаружили проблему № 1, а проблему № 10 нашли только первый и десятый участник. В приведенном примере вероятность обнаружения проблемы составила 0,5.

* Боженкова Елизавета Николаевна – магистрант кафедры информатики и кибернетики Байкальского государственного университета экономики и права, belizan@mail.ru.

Таблица 1

Пример расчета коэффициента обнаружения проблемы

Участник	Пробле- ма 1	Пробле- ма 2	Пробле- ма 3	Пробле- ма 4	Пробле- ма 5	Пробле- ма 6	Пробле- ма 7	Пробле- ма 8	Пробле- ма 9	Пробле- ма 10	Итого	Коэффи- циент
1	1	1		1		1		1		1	6	0,6
2	1	1		1		1		1			5	0,5
3	1	1		1	1	1					5	0,5
4	1	1		1			1				4	0,4
5	1	1	1	1		1			1		6	0,6
6	1	1	1					1			4	0,4
7	1	1	1		1						4	0,4
8	1	1	1		1		1				5	0,5
9	1		1		1		1		1		5	0,5
10	1		1		1		1		1	1	6	0,6
Итого	10	8	6	5	5	4	4	3	3	2	50	0,5

Предположим, что тест проводился только тремя участниками. В этом случае таблица обнаружения проблемы приобретет вид, приведенный в табл. 2. Вроде бы, вероятность обнаружения проблемы составила 0,533, что несильно отличается от начального расчета. Но, если посмотреть внимательно, то проблемы № 3, 7 и 9 не были обнаружены ни одним из участников, т.е. общее число обнаруженных проблем составило не 10, а 7. И, таким образом, меняется весь расчет. Корректировочный расчет приведен в табл. 3. Как видно из него, вероятность составляет уже 0,762, что является завышением к исходной более, чем на 50 %. Это является серьезной проблемой и требует перерасчета, чтоб вернуться обратно к первоначально рассчитанной вероятности.

Таблица 2

Пример расчета коэффициента обнаружения проблемы тремя участниками

Уча- стник	Про- бле- ма 1	Про- бле- ма 2	Про- бле- ма 3	Про- бле- ма 4	Про- бле- ма 5	Про- бле- ма 6	Про- бле- ма 7	Про- бле- ма 8	Про- бле- ма 9	Про- бле- ма 10	Ито- го	Коеф- фи- циент
1	1	1		1		1		1		1	6	0,6
2	1	1		1		1		1			5	0,5
3	1	1		1	1	1					5	0,5
Итого	3	3	0	3	1	3	0	2	0	1	16	0,533

Таблица 3

Пример корректировочного расчета коэффициента обнаружения проблемы тремя участниками

Участ- ник	Про- блема 1	Про- блема 2	Про- блема 4	Про- блема 5	Про- блема 6	Про- блема 8	Про- блема 10	Итого	Коеф- фициент
1	1	1	1		1	1	1	6	0,857
2	1	1	1		1	1		5	0,714
3	1	1	1	1	1			5	0,714
Итого	3	3	3	1	3	2	1	16	0,762

Для верного расчета вероятности необходима другая выборка участников теста. Число сочетаний таких выборок можно рассчитать по формуле:

$$C = \frac{n!}{r! * (n - r)!} \quad (1)$$

где n – общее число участников теста, r – количество участников выборки.

Например, если нас интересуют выборки из двух участников при десяти участвовавших, то количество таких выборок составит

$$C = \frac{10!}{2! * 8!} = \frac{9 * 10}{2!} = 45.$$

Число комбинаций для трех, четырех и пяти участников, выбранных из десяти, соответственно составит 120, 210 и 252. Как видим, точный расчет пу-

тем перебора всех возможных выборок очень трудоемко. И это при том, что рассматриваемый нами тест включал всего десять участников. Например, данные Якоба Нильсена и Рольфа Молиха содержали данные по проведению тестирования из 76 участников и 30 обнаруженных проблем. Число комбинаций выборок из шести участников составило бы уже 218 618 940.

Роберт Вирзи, Джеймс Льюис в своих статьях доказали возможность использования метода Монте-Карло для расчета вероятности обнаружения проблем. Вирзи, применив этот метод к трем тестированиям, доказал асимптотическое отношение между числом участников теста и долей обнаруженных проблем. На основании проведенных исследований Вирзи сделал следующие выводы:

1. Четыре или пять участников теста способны обнаружить 80% имеющихся юзабилити-проблем.

2. Привлечение дополнительных участников теста приводит к выявлению меньшего количества новых проблем.

3. Наблюдающие за первыми участниками теста обнаруживают более серьезные юзабилити-проблемы.

В 1992 г. Вирзи предложил следующую формулу для вычисления доли обнаруженных проблем:

$$D = 1 - (1 - p)^n, (2)$$

где p – вероятность обнаружения проблемы, n – количество участников теста.

В 1993 г. Якоб Нильсен и Томас Ландауэр [3], используя распределение Пуассона предложили похожую формулу для вычисления числа обнаруженных проблем:

$$T = N * (1 - (1 - p)^n) (3)$$

где N – общее количество проблем, p – вероятность обнаружения проблемы, n – количество участников теста.

Отсюда и возникло «магическое число 5», так как из вышеприведенных формул следовало, что пять участников тестирования способны обнаружить 80 % юзабилити-проблем.

Далее эта тема была поддержана в ряде работ. В частности, на основании формулы (1) в [5] была предложена интересная математическая модель, описывающая поведение юзабилити-проблем. Но данная математическая модель не прошла проверку по критериям χ^2 и Колмогорова-Смирнова.

Споры же по количеству участников тестирования не утихают до сих пор. Так, Джитт Линдгаард и Джарини Чаттратичарт [4] в 2008 г., используя вышеупомянутые данные Якоба Нильсена и Рольфа Молиха, провели более сложный анализ. В качестве параметров этого анализа они, использовали, в частности, уровень знаний участников теста, проранжировали вопросы, учитывали, какие именно вопросы задавались и в каком виде. В результате такого исследования они не нашли корреляции между числом участников теста и количеством обнаруженных проблем.

Поэтому тема оптимального числа участников до сих пор не раскрыта. Требуется дальнейшее исследование. И это хороший повод, например, доработать математическую модель, предложенную в [5].

Список использованной литературы

1. Lewis J. R. Sample sizes for usability studies: Additional considerations. / Lewis, J. R. // Human Factors – 1994. – № 36. – С. 368–378.
2. Virzi R. A. Refining the test phase of usability evaluation: How many subjects is enough? / Virzi R. A. // Human Factors. – 1992. – № 34. – С. 443–451.
3. Nielsen J., Landauer T.K. A mathematical model of the finding of usability problems. / Nielsen J., Landauer T.K. // Proceedings of ACM INTERCHI'93 Conference (Amsterdam, Netherlands: ACM Press). – 1993. – С. 206–213.
4. Lindgaard G., Chattratchart J. Usability Testing: What Have We Overlooked? / Lindgaard G., Chattratchart J. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '07. – 2008 – С. 1415–1424.
5. Krapivin O. Anpassungstests für das Binomialmodell zur Erklärung von Usability-Test-Daten: Bachelorarbeit zur Erlangung des akademischen Grades/ Krapivin O. – Köthen – 2012. – 48 с.