http://dx.doi.org/10.35596/1729-7648-2024-30-4-62-72

Оригинальная статья
Original paper

УДК 004.33.054

ПОСТРОЕНИЕ УПРАВЛЯЕМЫХ ВЕРОЯТНОСТНЫХ ТЕСТОВ С ЗАДАННЫМ РАССТОЯНИЕМ ХЭММИНГА

В. Н. ЯРМОЛИК, Д. В. ДЕМЕНКОВЕЦ, В. А. ЛЕВАНЦЕВИЧ, В. В. ПЕТРОВСКАЯ

Белорусский государственный университет информатики и радиоэлектроники
(г. Минск, Республика Беларусь)

Поступила в редакцию 18.07.2024

Аннотация. Рассмотрены вопросы тестирования вычислительных систем и их составных компонентов. Показана невысокая эффективность вероятностного тестирования из-за неиспользования имеющейся информации как об объекте тестирования, так и о ранее сгенерированных тестовых наборах. Отмечена перспективность применения информации о предыдущих тестовых наборах при построении управляемых вероятностных тестов. Выделен и исследован класс управляемых вероятностных тестов с малым числом тестовых наборов. Проанализированы известные стандартные управляемые вероятностные тесты с малым числом тестовых наборов. Рассмотрен метод генерирования управляемых вероятностных тестов с заданным расстоянием Хэмминга, основой которого является представление двоичных тестовых наборов в виде совокупности символов иных систем счисления. Обеспечение заданного значения расстояния Хэмминга для управляемого вероятностного теста в недвоичной системе счисления гарантирует такое же значение и для двоичной интерпретации формируемых тестовых наборов. Эффективность предложенного метода построения тестов экспериментально проанализирована и подтверждена для случая тестирования запоминающих устройств на наличие сложных неисправностей взаимного влияния.

Ключевые слова: тест, тестовый набор, тестирование вычислительных систем, вероятностное тестирование, расстояние Хэмминга, граница Плоткина, неисправности запоминающих устройств, тестирование запоминающих устройств.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Построение управляемых вероятностных тестов с заданным расстоянием Хэмминга / В. Н. Ярмолик [и др.] // Цифровая трансформация. 2024. Т. 30, № 4. С. 62–72. http://dx.doi.org/10.35596/1729-7648-2024-30-4-62-72.

DESIGN OF CONTROLLED RANDOM TESTS WITH THE GIVEN HAMMING DISTANCE

VYACHESLAV N. YARMOLIK, DENIS V. DEMENKOVETS,
VLADIMER A. LEVANTSEVICH, VITA V. PETROVSKAYA

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Submitted 18.07.2024

Abstract. The article considers the issues of testing computing systems and their components. It shows the low efficiency of probabilistic testing due to the failure to use the available information about both the test object and previously generated test sets. It is noted that it is promising to use information about previous test sets when constructing controlled probabilistic tests. A class of controlled probabilistic tests with a small number of test sets is identified and studied. Known standard controlled probabilistic tests with a small number of test sets are analyzed. A method for generating controlled probabilistic tests with a given Hamming distance is considered, the basis of which is the representation of binary test sets as a set of symbols of other number systems. Providing a given value of the Hamming distance for a controlled probabilistic test in a non-binary number system guarantees the same value for the binary interpretation of the generated test sets. The efficiency of the proposed test construction method is experimentally analyzed and confirmed for the case of testing memory devices for the presence of complex faults of mutual influence.

Keywords: test, test pattern, computer system testing, random testing, Hamming distance, Plotkin bound, memory faults, memory testing.

Conflict of interests. The authors declare no conflict of interests.

For citation. Yarmolik V. N., Demenkovets D. V., Levantsevich V. A., Petrovskaya V. V. (2024) Design of Controlled Random Tests with the Given Hamming Distance. *Digital Transformation.* 30 (4), 62–72. http://dx.doi.org/10.35596/1729-7648-2024-30-4-62-72 (in Russian).

Введение

Несмотря на десятилетия усилий по разработке альтернативных технологий, тестирование вычислительных систем – основной способ проверки качества программных и аппаратных средств вычислительных систем и их приложений. Однако тестирование остается трудоемким, медленным и несовершенным процессом. Поэтому важным является использование систематических автоматизированных методов построения тестовых процедур. Попытки автоматизировать создание тестовых процедур и данных с помощью различных форм вероятностного формирования и отбора были предприняты с начала 1960-х годов и стали постоянной практикой как в научной, так и в прикладной сферах [1–4]. Вероятностное тестирование (random testing) просто по своей концепции, легко реализуется и эффективно выявляет неисправности вычислительных систем и их компонентов, а именно: программного обеспечения, аппаратных средств и запоминающих устройств [4, 5]. Данный вид тестирования широко использовался и используется сам по себе как метод тестирования, более того, он составляет основную часть многих других методов тестирования [5].

Основным недостатком вероятностного тестирования является его невысокая эффективность из-за неиспользования имеющейся информации об объекте тестирования при построении тестовых данных. Поэтому в настоящее время активно развиваются различные методы построения тестов, в которых случайный фактор играет не доминирующую роль [5–7]. Лидирующую позицию среди них занимает адаптивное вероятностное тестирование (adaptive random testing) [8–10]. В русскоязычных публикациях подобное тестирование часто называют управляемым вероятностным тестированием (controlled random testing). Данный вид тестирования и его многочисленные модификациии основаны на вычислении некоторых характеристик для управляемого формирования очередного случайного тестового набора. Большинство известных подходов генерирования адаптивных вероятностных тестов базируется на применении расстояния Хэмминга в качестве характеристики, определяющей выбор очередного набора. Поиск тестового набора из потенциальных кандидатов в тестовые наборы состоит в нахождении такого кандидата в тесты, который удовлетворяет заданным критериям. В случае расстояния Хэмминга этим критерием является само расстояние, пороговое значение которого влияет как на процедуру выбора тестового набора, так и на длину самого теста. Чем выше значения критериев выбора, в частности, расстояния Хэмминга, тем сложнее процедура выбора и заметнее уменьшение длины вероятностного теста. Необходимость перебора потенциальных кандидатов в тестовые наборы и вычисления для них характеристик увеличивает сложность формирования управляемых вероятностных тестов [5, 9].

Анализ управляемых вероятностных тестов

Как показано в [5–13], очередной тестовый набор $T_i = t_{i,0}, t_{i,1}, \dots, t_{i,n-1}$, где $t_{i,l} \in \{0, 1\}$, $l \in \{0, 1, 2, \dots, n-1\}$, управляемого вероятностного теста формируется максимально отличающимся от ранее сгенерированных наборов $T_0, T_1, T_2, \dots, T_{i-1}$. Таким образом, принимается гипотеза, что для двух тестовых наборов, имеющих максимальное отличие, количество обнаруживаемых неисправностей (ошибок) вторым набором будет максимальным. В качестве критерия отличия тестового набор T_i от предыдущих наборов $T_0, T_1, T_2, \dots, T_{i-1}$ часто применяется расстояние Хэмминга

(Hamming distance) $HD(T_i, T_j)$ для $j \in \{0, 1, \ldots, i-1\}$. Соотношение очередного тестового набора к предыдущим наборам определяется минимальным расстоянием Хэмминга $\min HD(T_i, T_j)$, величина которого соответствует определению согласно [14, 15].

Определение 1. Значение $\min HD(T_i, T_j)$ равняется минимальному расстоянию Хэмминга между двумя произвольными тестовыми наборами T_i и T_j, $i \neq j \in \{0, 1, 2, \ldots, q-1\}$, из множества наборов теста $T = \{T_0, T_1, T_2, \ldots, T_{q-1}\}$.

В терминах теории помехоустойчивого кодирования характеристику $\min HD(T_i, T_j)$ можно рассматривать как кодовое расстояние d кода T, которое равняется наименьшему расстоянию Хэмминга между различными парами кодовых слов $T_0, T_1, T_2, \ldots, T_{q-1}$. Поэтому, исходя из основополагающих принципов теории кодирования, можно сформулировать ряд полезных выводов, которые необходимо учитывать при генерировании управляемых вероятностных тестов [5].

В частности, характерной особенностью управляемых вероятностных тестов является ограниченность их длины, причем, чем больше минимальное расстояние $\min HD(T_i, T_j)$, используемое как критерий включения T_i в тест, тем меньше количество q наборов, удовлетворяющих данному критерию. Это следует из предельной оценки Хэмминга (Hamming bound) [15]. Для $\min HD(T_i, T_j) = 2r + 1$, r – целое, эта оценка может быть представлена как неравенство

$$q \leq 2^n / \sum_{l=0}^{r} \binom{n}{l}. \tag{1}$$

Например, в случае, когда $n = 8$ и $d = \min HD(T_i, T_j) = 5 = 2 \cdot 2 + 1$, согласно (1), оказывается возможным построение теста, состоящего из $q \leq 2^8 / \sum_{l=0}^{2} \binom{8}{l} = 2^8 / (1 + 8 + 28) \leq 6$ тестовых наборов с расстоянием Хэмминга $HD(T_i, T_j)$ больше или равным 5. Увеличение расстояния Хэмминга $\min HD(T_i, T_j)$, например, до значения 7, уменьшает оценку q до величины 2, откуда следует, что управляемый вероятностный тест для $n = 8$ и $d = \min HD(T_i, T_j) = 7$ будет состоять не более чем из двух наборов $T = \{T_0, T_1\}$. Важно отметить, что набор T_0 формируется случайным образом и может принимать любое из $2^n = 2^8$ двоичных значений, а второй набор выбирается согласно критерию $\min HD(T_0, T_1) \geq 7$. Таким образом, существует большое множество управляемых вероятностных тестов T с $\min HD(T_i, T_j) = 7$, однако каждый из них состоит только из двух наборов T_0 и T_1.

Для синтеза вероятностных тестов с малым количеством наборов q в [12] рассмотрены классические коды с максимальным минимальным расстоянием Хэмминга $\min HD(T_i, T_j) > n/2$. Показано, что теорема Плоткина позволяет определить максимально возможное количество q кодовых слов среди двоичных кодов длины n для минимального кодового расстояния, а граница Плоткина дает верхний предел этого количества [11].

Определение 2 (граница Плоткина). Если $d = \min HD(T_i, T_j) \geq n/2$, то для q выполняется неравенство

$$q \leq \begin{cases} \dfrac{2d}{2d-n}, & \text{для} \quad 2d - n > 0; \\ 4d, & \text{для} \quad 2d - n = 0. \end{cases} \tag{2}$$

Для случая, когда $2d - n > 0$ (2), справедливо следующее соотношение:

$$d \leq qn/(2(q-1)). \tag{3}$$

Основополагающая идея построения управляемых вероятностных тестов $MMHD(q) = \{T_0, T_1, T_2, \ldots, T_{q-1}\}$ с малым количеством тестовых наборов q заключается в обеспечении максимально возможного значения $\max\_\min HD(T_i, T_j)$ для $d = \min HD(T_i, T_j) > n/2$ [12]. Соотношение (3) позволяет получить максимальное минимальное расстояние $\max\_\min HD(T_i, T_j)$ для заданных значений q. Как показано в [5, 12], для случая $q = 3$, в соответствии с (3), $\max\_\min HD(T_i, T_j) \leq 3n/4$. Однако для теста $MMHD(3)$, так же, как и для $MMHD(4)$, как показано в [5], $\max\_\min\{MMHD(3)\} = \max\_\min\{MMHD(4)\} = 2n/3$. При увеличении числа наборов q значение $\max\_\min\{MMHD(q)\}$ быстро сходится к величине, равной $n/2$. Поэтому данный алгоритм целесообразно применять для синтеза вероятностных тестов с малым числом наборов $q < 10$ [12].

В [13] рассмотрен метод синтеза оптимальных управляемых вероятностных тестов (Optimal Controlled Random Tests (OCRT)) с использованием расстояния Хэмминга $HD(T_i, T_j)$ для тестовых наборов T_i и T_j и их декартова расстояния $CD(T_i, T_j)$. Подобные тесты характеризуются тем, что для них $d = \min HD(T_i, T_j) \geq n/2$. Для общего случая количество наборов OCRT определяется как $q = 2(\lceil \log_2 n \rceil + 1)$, а конструктивный алгоритм для формирования тестовых наборов представлен в [5, 13]. В случае, когда $n = 2^m$, где m – целое, количество q наборов OCRT = $\{T_0, T_1, T_2, \ldots, T_{q-1}\}$ равняется $2(m + 1)$. Например, для $n = 4$ количество тестовых наборов OCRT $q = 6$, а для $n = 8$ – соответственно, 8. Как несложно заметить (см. Определение 2), граница Плоткина в первом случае дает верхнюю оценку количества наборов, равную 8, а во втором – 16. С ростом разрядности наборов n сложность OCRT-теста еще более отстоит от максимальных граничных оценок (2).

Суммируя проведенный анализ, можно сделать следующие выводы. Во-первых, увеличение пороговых значений $\min HD(T_i, T_j)$ при синтезе управляемых вероятностных n-разрядных тестовых наборов уменьшает их количество q, и наоборот, уменьшение задаваемых $\min HD(T_i, T_j)$ увеличивает длину теста. Во-вторых, известные формальные методы генерирования управляемых вероятностных тестов с фиксированной длиной характеризуются лишь приближением к их максимальным теоретическим оценкам (1) и (2). И, наконец, третьим, главным выводом по результатам приведенного анализа является высокая вычислительная сложность генерирования управляемых вероятностных тестов, требующая выбора тестового набора из большого числа кандидатов в тесты. Только для $\min HD(T_i, T_j) \geq n/2$ и при других ограничениях удается избежать рутинной процедуры поиска тестовых наборов, удовлетворяющих заданным критериям, из множества потенциальных кандидатов в тесты [9, 12, 13].

Задача построения управляемых вероятностных тестов с произвольным фиксированным расстоянием Хэмминга как критерием включения кандидата в генерируемый тест и невысокой вычислительной сложностью по-прежнему является актуальной.

Управляемые вероятностные тесты с заданным расстоянием Хэмминга

Для общего случая определение расстояния Хэмминга основано на сравнении двух последовательностей $T_i = t_{i,0}, t_{i,1}, \ldots, t_{i,n-1}$ и $T_j = t_{j,0}, t_{j,1}, \ldots, t_{j,n-1}$, включающих по n символов $t_{i,l}$ и $t_{j,l}$ из произвольного алфавита [11]. Расстояние Хэмминга $HD(T_i, T_j)$ между T_i и T_j как количество позиций, в которых $t_{i,l}$ и $t_{j,l}$ различаются, описывается соотношением

$$HD(T_i, T_j) = \sum_{l=0}^{n-1} \mathrm{I}(t_{i,l} \neq t_{j,l}). \tag{4}$$

Выражение $\mathrm{I}(t_{i,l} \neq t_{j,l})$ представляет собой индикаторную функцию, равную единице при $t_{i,l} \neq t_{j,l}$ и нулю в противном случае. В результате сравнения одноименных символов наборов T_i и T_j минимальное значение $\min HD(T_i, T_j)$ равняется 0 при совпадении всех символов, а $\max HD(T_i, T_j) = n$ при несовпадении всех n символов. Чаще всего рассматривается случай двоичных тестовых наборов, которые можно интерпретировать как наборы символов из иных алфавитов [14, 15]. Например, четверичного, восьмеричного, шестнадцатеричного и других алфавитов, для которых фиксированное количество последовательных бит исходного двоичного набора представляет собой двоичный код символа соответствующего алфавита. Двоичный набор $T_i = T_{i(2)} = 01111011_{(2)}$, рассматривая каждые два последовательных бита в отдельности, можно представить в четверичной системе счисления, в которой используется алфавит из четырех символов 0, 1, 2 и 3, как $T_{i(4)} = 1323_{(4)}$. В шестнадцатеричной системе этот же набор $T_{i(2)}$ примет вид $T_{i(16)} = 7B_{(16)}$. Принимая разрядность n исходного двоичного тестового набора как величину, равную 2^m, следует отметить, что при изменении алфавита представления набора T_i изменяется и его размерность: от n двоичных символов до одного символа 2^{2^m}-й системы счисления.

Предложенная интерпретация последовательностей двоичных символов T_i и T_j, состоящих из $n = 2^m$ бит, в виде последовательностей $T_{i(2)}, T_{i(4)}, T_{i(16)}, \ldots, T_{i(2^{2^m})}$ и $T_{j(2)}, T_{j(4)}, T_{j(16)}, \ldots, T_{j(2^{2^m})}$ позволяет применять для их сравнения расстояние Хэмминга (4). При этом анализируются пары символов из двух наборов, состоящих из одинакового количества символов одного и того же алфавита. Основываясь на предыдущем примере набора $T_i = 01111011_{(2)}$ и используя $T_j = 01011011_{(2)}$, на основании (4) получим $HD_{(2)}(01111011, 01011011) = 1$, $HD_{(4)}(1323, 1123) = 1$, $HD_{(16)}(7B, 5B) = 1$ и $HD_{(256)}(\{, [) = 1$.

Приведенный пример определения расстояния Хэмминга показывает возможность получения на основании (4) нескольких численных оценок соотношения исходных двоичных наборов T_i и T_j. Обобщая приведенные рассуждения, определим расстояние Хэмминга для двоичных $n = 2^m$-разрядных тестовых наборов T_i и T_j, представленных в $w = 2^{2^r}$-м алфавите, где $r \in \{0, 1, 2, \ldots, m\}$.

Определение 3. Расстояние Хэмминга $HD_{(w)}(T_i, T_j)$ для двоичных тестовых наборов $T_i = t_{i,0}, t_{i,1}, \ldots, t_{i,2^m-1}$ и $T_j = t_{j,0}, t_{j,1}, \ldots, t_{j,2^m-1}$, где $t_{i,l}, t_{j,l} \in \{0, 1\}$; $l \in \{0, 1, 2, \ldots, 2^m-1\}$; $w = 2^{2^r}$; $r \in \{0, 1, 2, \ldots, m\}$, определяется согласно (4), когда 2^r последовательных бит T_i и T_j интерпретируются как символы 2^{2^r}-го алфавита.

Расстояния Хэмминга $HD_{(w)}(T_i, T_j)$, где $w = 2^{2^r}$ и $r \in \{0, 1, 2, \ldots, m\}$, для двоичных тестовых наборов T_i и T_j имеют следующие основные свойства и соотношения [15].

Свойство 1. Минимальное численное значение $\min HD_{(w)}(T_i, T_j)$ равняется 0, а максимальное $\max HD_{(w)}(T_i, T_j) = \max HD_{(2^{2^r})}(T_i, T_j) = 2^{m-r}$ – определяется количеством символов соответствующего алфавита w в сравниваемых наборах T_i и T_j.

В соответствии со свойством 1 имеем: $\max HD_{(2)}(T_i, T_j) = 2^m$, $\max HD_{(4)}(T_i, T_j) = 2^{m-1}$, $\max HD_{(16)}(T_i, T_j) = 2^{m-2}, \ldots, \max HD_{(2^{2^m})}(T_i, T_j) = 1$.

Свойство 2. Численные значения расстояний Хэмминга $HD_{(w)}(T_i, T_j)$ связаны следующими соотношениями: $HD_{(2)}(T_i, T_j) \geq HD_{(4)}(T_i, T_j) \geq HD_{(16)}(T_i, T_j) \geq \ldots \geq HD_{(2^{2^m})}(T_i, T_j)$.

Выполнение данного свойства объясняется тем, что при вычислении $HD_{(2^{2^r})}(T_i, T_j)$ используются символы наборов $T_{i,(2^{2^r})}$ и $T_{j,(2^{2^r})}$, каждый из которых состоит их двух символов наборов $T_{i,(2^{2^{r-1}})}$ и $T_{j,(2^{2^{r-1}})}$, т. е. $t_{i,l}(2^{2^r}) = t_{i,2l}(2^{2^{r-1}}), t_{i,2l+1}(2^{2^{r-1}})$ и $t_{j,l}(2^{2^r}) = t_{j,2l}(2^{2^{r-1}}), t_{j,2l+1}(2^{2^{r-1}})$, где $l = 0, 1, 2, \ldots, 2^{m-r} - 1$. Таким образом, результатом сравнения $\mathrm{I}(t_{i,l}(2^{2^r}) \neq t_{j,l}(2^{2^r}))$ символов наборов $T_{i,(2^{2^r})}$ и $T_{j,(2^{2^r})}$ могут быть только два значения – 0 или 1, используемые для получения $HD_{(2^{2^r})}(T_i, T_j)$ (4). Результатом сравнения двух пар символов последовательностей $T_{i,(2^{2^{r-1}})}$ и $T_{j,(2^{2^{r-1}})}$ могут быть 0, 1 и 2. При этом при выполнении неравенства $t_{i,l}(2^{2^r}) \neq t_{k,l}(2^{2^r})$ значение $HD_{(2^{2^r})}(T_i, T_j)$ (4) увеличится только на 1, а $HD_{(2^{2^{r-1}})}(T_i, T_j)$ может быть увеличено как на 1, так и на 2, но как минимум на 1.

Важным следствием свойства 2 является выполнение неравенства $HD_{(2)}(T_i, T_j) \geq HD_{(2^{2^r})}(T_{i,(2^{2^r})}, T_{j,(2^{2^r})})$, которое можно интерпретировать следующим образом. Обеспечение заданного расстояния Хэмминга $\min HD_{(w)}(T_i, T_j)$ для наборов T_i и T_j, представленных в $w = 2^{2^r}$-м алфавите, гарантирует значение величины $HD_{(2)}(T_i, T_j)$ для этих же наборов, представленных в двоичном алфавите, не менее чем $\min HD_{(w)}(T_i, T_j)$.

Поясним данное следствие на примере двоичного набора $T_i = 01111011_{(2)}$, который в четверичном алфавите имеет вид $T_i = 1323_{(4)}$. Если в качестве T_j выбрать набор $2021_{(4)}$, то $HD_{(4)}(1323, 2021) = 3$. В то же время для двоичного случая величина $HD_{(2)}(01111011, 10001001) = 5$ принимает значение не менее чем $HD_{(4)}(1323, 2021) = 3$, что соответствует свойству 2.

Продолжая рассматривать этот пример, далее можно отметить, что для формирования четверичного набора T_j, максимально отличного от набора $T_i = 1323_{(4)}$, т. е. для которого $HD_{(4)}(T_i, T_j) = 4$, необходимо сформировать каждый четверичный символ этого набора T_j отличающимся от соответствующего символа набора T_i. Таким образом, в качестве первого символа набора T_j могут быть 0, 2 и 3, в качестве второго – 0, 1 и 2, третьего – 0, 1 и 3, а четвертого – 0, 1 и 2. Один из возможных вариантов набора T_j имеет вид $0210_{(4)}$, а общее их количество равняется 3^4. В результате приведенных рассуждений наборы $T_i = 1323_{(4)}$ и $T_j = 0210_{(4)}$ можно интерпретировать как управляемый вероятностный тест, состоящий из двух наборов с $HD_{(2)}(T_i, T_j) \geq HD_{(4)}(T_i, T_j) = 4$. Действительно, $HD_{(2)}(T_i, T_j) = HD_{(2)}(01111011, 00100100) = 6$.

Третий набор T_k управляемого вероятностного теста, состоящего уже из трех четверичных наборов, а именно T_i, T_j и T_k, каждый из которых максимально удален друг от друга, формируется аналогичным образом. Соответственно, первым четверичным символом набора T_k может быть 2 либо 3 в силу того, что первый символ набора T_i равняется 1, а набора T_j – 0. Аналогичным образом определяются кандидаты в последующие символы набора T_k. В результате третьим набором T_k может быть, например, набор $2132_{(4)}$, максимально удаленный от T_i и T_j, так как $HD_{(4)}(1323, 2132) = HD_{(4)}(0210, 2132) = 4$. Общее количество возможных значений набора T_k равняется 2^4.

Следующим, четвертым набором T_o управляемого вероятностного теста, максимально удаленным от T_i, T_j и T_k, может быть только набор, имеющий вид $3001_{(4)}$, так как в противном случае были бы использованы повторяющиеся символы в одноименных разрядах наборов T_i, T_j, T_k и T_o. Наличие одинаковых символов приводит к тому, что $HD_{(4)}$ становится меньше 4. Сформированный таким образом управляемый вероятностный тест, включающий четыре тестовых набора, а также численные значения расстояний Хэмминга, приведен в табл. 1.

Таблица 1. Значения расстояний Хэмминга $HD_{(4)}$ и $HD_{(2)}$ для тестовых наборов T_i, T_j, T_k и T_o
Table 1. Values of Hamming distances $HD_{(4)}$ and $HD_{(2)}$ for test patterns T_i, T_j, T_k and T_o

T_i, T_j, T_k, T_o		$HD_{(4)}$	$HD_{(2)}$
$1323_{(4)}$ $0210_{(4)}$ $2132_{(4)}$ $3001_{(4)}$	$01111011_{(2)}$ $00100100_{(2)}$ $10011110_{(2)}$ $11000001_{(2)}$	$HD_{(4)}(T_i, T_j) = 4$ $HD_{(4)}(T_i, T_k) = 4$ $HD_{(4)}(T_j, T_k) = 4$ $HD_{(4)}(T_i, T_o) = 4$ $HD_{(4)}(T_j, T_o) = 4$ $HD_{(4)}(T_k, T_o) = 4$	$HD_{(2)}(T_i, T_j) = 6$ $HD_{(2)}(T_i, T_k) = 5$ $HD_{(2)}(T_j, T_k) = 5$ $HD_{(2)}(T_i, T_o) = 5$ $HD_{(2)}(T_j, T_o) = 5$ $HD_{(2)}(T_k, T_o) = 6$

Анализ приведенного примера формирования управляемого вероятностного теста, состоящего из четырех тестовых наборов на основе четверичного алфавита, позволил сформировать тест, для которого $\max\_\min HD(T_x, T_y) = 5$, $x \neq y \in \{i, j, k, o\}$ (табл. 1). При построении данного теста разрядность тестовых наборов должна принимать значение, кратное двум, и может быть произвольной. Важным является использование в каждом разряде наборов различных значений четверичных символов, так, как это было получено в случае $n = 8$ (табл. 1). С помощью подобной процедуры можно сгенерировать аналогичные тесты, состоящие из произвольного четного количества бит n. При этом единственным критерием при формировании тестовых наборов является использование в одноименных разрядах наборов всех четверичных символов 0, 1, 2 и 3. Примеры таких управляемых вероятностных тестов, состоящих из четырех наборов, приведены в табл. 2.

Таблица 2. Управляемые вероятностные тесты, состоящие из четырех наборов
Table 2. Controlled random tests consisting of four patterns

$n = 2$	$n = 4$	$n = 8$	$n = 16$
$1_{(4)} = 01_{(2)}$ $3_{(4)} = 11_{(2)}$ $0_{(4)} = 00_{(2)}$ $2_{(4)} = 10_{(2)}$	$01_{(4)} = 0001_{(2)}$ $32_{(4)} = 1110_{(2)}$ $10_{(4)} = 0100_{(2)}$ $23_{(4)} = 1011_{(2)}$	$0000_{(4)} = 00000000_{(2)}$ $1111_{(4)} = 01010101_{(2)}$ $2222_{(4)} = 10101010_{(2)}$ $3333_{(4)} = 11111111_{(2)}$	$30201000_{(4)} = 1100100001000000_{(2)}$ $21312131_{(4)} = 1001110110011101_{(2)}$ $12023222_{(4)} = 0110001011101010_{(2)}$ $03130313_{(4)} = 0011011100110111_{(2)}$

Как видно из табл. 2, для случая $n = 8$ тест имеет регулярную структуру, когда каждый разряд тестовых наборов представляет собой одинаковую последовательность четверичных символов 0, 1, 2 и 3. Главной особенностью представленных в табл. 2 тестов является обеспечение для их тестовых наборов значения $\max\_\min HD(T_x, T_y) \geq n/2$.

Основываясь на рассмотренных примерах построения управляемого вероятностного теста, состоящего из четырех тестовых наборов и обеспечивающего фиксированное $\max\_\min HD(T_x, T_y) \geq n/2$, рассмотрим формальную процедуру построения подобных тестов для различных значений $\max\_\min HD(T_x, T_y)$. Основой указанной процедуры является следующее утверждение, которое сформулируем для случая тестовых наборов, состоящих из $n = 2^m$ бит.

Утверждение. Управляемый вероятностный тест, сформированный из 2^{2^r}, $r \in \{0, 1, 2, \ldots, m\}$, тестовых наборов, представленных в $w = 2^{2^r}$-м алфавите, в каждом разряде которых используются неповторяющиеся символы данного алфавита, обеспечивает $\max\_\min HD(T_x, T_y) \geq n/2^r$ для двоичного случая представления теста. Например, для $w = 2^{2^1} = 4$-го алфавита имеем $\max\_\min HD(T_x, T_y) \geq$

$\geq n/2^1 = n/2$, что подтверждается ранее приведенными примерами (табл. 1, 2). Если предположить, что $w = 2^n = 2^{2^m}$, получим max_min$HD(T_x, T_y) \geq n/2^w = 1$, а управляемый вероятностный тест будет состоять из $w = 2^n$ наборов. В случае двоичного алфавита ($w = 2^{2^0}$) max_min$HD(T_x, T_y) \geq$ $\geq n/2^0 = n$ и соответствующий тест состоит только из двух наборов ($w = 2$).

В общем случае управляемый вероятностный тест, для которого max_min$HD(T_x, T_y) \geq n/2^r$, будет состоять из 2^{2^r}, $r \in \{0, 1, 2, \ldots, m\}$, тестовых наборов размерностью $n = 2^m$ бит. Процедура их построения состоит в формировании в каждых последовательных 2^r разрядах для всех наборов всевозможных символов 2^{2^r} – го алфавита в произвольном порядке. Результаты построения подобных тестов для $r = 1$ (четверичного алфавита) представлены в табл. 1, 2.

Предложенная процедура построения управляемых вероятностных тестов с заданным расстоянием Хэмминга max_min$HD(T_x, T_y)$ легко обобщается для случая, когда рассматриваются последовательные не только 2^r разряды тестовых наборов, но и произвольное их количество $z \leq n$. Тогда аналогично ранее рассмотренному случаю управляемый вероятностный тест, состоящий из 2^z тестовых наборов, будет обеспечивать фиксированные значения расстояний Хэмминга, равные величинам, не менее чем $\lfloor n/z \rfloor$. Данный тест формируется по ранее описанной процедуре с использованием 2^z символов 2^z-го алфавита.

При рассмотрении трех последовательных разрядов тестовых наборов как символов, представленных в восьмеричном алфавите, при условии, что n кратно трем, управляемый вероятностный тест, состоящий из восьми тестовых наборов, обеспечивает фиксированные значения расстояния Хэмминга, величина которого будет не менее чем $n/3$. В табл. 3 приведены два варианта управляемого вероятностного теста для $n = 9$ с заданным max_min$HD(T_x, T_y) \geq n/3 = 9/3 = 3$, построенные согласно рассмотренному методу.

Таблица 3. Управляемые вероятностные тесты, состоящие из восьми наборов
Table 3. Controlled random tests consisting of eight patterns

$000_{(8)} = 000000000_{(2)}$	$705_{(8)} = 111000101_{(2)}$
$111_{(8)} = 001001001_{(2)}$	$611_{(8)} = 110001001_{(2)}$
$222_{(8)} = 010010010_{(2)}$	$533_{(8)} = 101011011_{(2)}$
$333_{(8)} = 011011011_{(2)}$	$422_{(8)} = 100010010_{(2)}$
$444_{(8)} = 100100100_{(2)}$	$350_{(8)} = 011101000_{(2)}$
$555_{(8)} = 101101101_{(2)}$	$247_{(8)} = 010100111_{(2)}$
$666_{(8)} = 110110110_{(2)}$	$174_{(8)} = 001111100_{(2)}$
$777_{(8)} = 111111111_{(2)}$	$066_{(8)} = 000110110_{(2)}$

Как видно из табл. 3, расстояние Хэмминга в обоих случаях принимает фиксированные значения в диапазоне от 3 до 9.

Оценка эффективности тестов с заданным расстоянием Хэмминга

Управляемые вероятностные тесты находят широкое применение при тестировании вычислительных систем и их составных компонентов, таких как запоминающие устройства и в первую очередь оперативные запоминающие устройства (ОЗУ) [5–9]. Во всех случаях эффективность управляемых вероятностных тестов сравнивается с эффективностью вероятностных тестов, состоящих из такого же количества тестовых наборов. В случае запоминающих устройств применение вероятностных тестов, состоящих из p наборов, позволяет достичь полноты покрытия $FC_{Test}(Fault, p)$ их сложных неисправностей, оцениваемой выражением [5]:

$$FC_{Test}(Fault, p) = \left(1 - \left(1 - \frac{FC_{Test}(Fault)}{100\%}\right)^p\right) \cdot 100\%. \quad (5)$$

Значение $FC_{Test}(Fault)$ показывает процентное отношение количества обнаруженных сложных неисправностей при использовании одного вероятностного набора и заданной неразрушающей маршевой процедуры тестирования (Test), используемых для тестирования ОЗУ. В случае запоминающих устройств случайный тестовый набор применяется в качестве начального состояния его ячеек, а под сложными их неисправностями понимаются кодочувствительные неисправности и неисправности взаимного влияния [5].

С целью оценки эффективности предлагаемых управляемых вероятностных тестов был реализован следующий эксперимент. С использованием языка Object Pascal и среды разработки Delphi 7.0 спроектировали программное средство, в котором смоделировали ОЗУ с общим количеством ячеек памяти n.

Для реализации процедуры тестирования Test применялся классический неразрушающий маршевый тест Николаидиса под названием TMATS [5]. Данный тест состоит из двух фаз. Начальная фаза, предназначенная для вычисления эталонных значений сигнатуры SF, имеет вид $\{\Uparrow(ra); \Uparrow(r\bar{a})\}$. Вторая фаза, представляющая собой базовый тест, имеет вид $\{\Uparrow(ra,\ w\bar{a});$ $\Uparrow(r\bar{a})\}$ и предназначена для вычисления реальной сигнатуры при наличии неисправности в запоминающем устройстве. Для получения эталонной сигнатуры SF, соответствующей содержимому ОЗУ без внесенных неисправностей, перед выполнением базового теста запускается начальная фаза теста. Для обнаружения неисправностей результаты всех операций чтения второй фазы теста сжимаются с помощью сигнатурного анализатора в реальную сигнатуру SR. Далее сигнатуры сравниваются и их неравенство SF ≠ SR устанавливает обнаружение внесенной в ОЗУ неисправности. В противном случае, когда SF = SR, фиксируется факт необнаружения неисправности. В качестве анализатора состояния памяти использовался сигнатурный анализатор, описываемый порождающим полиномом $\varphi(x) = 1 + x + x^3 + x^{12} + x^{16}$.

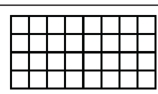
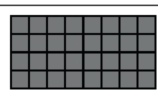
Эксперимент проводился с многократным применением вероятностного теста, состоящего из четырех наборов начальных состояний ОЗУ, сгенерированных случайным образом. Состояние каждой ячейки ОЗУ 0 либо 1 формировалось равновероятно и независимо от состояния других ячеек памяти. Подобный тест многократно повторялся, и для каждой его итерации оценивалось обнаружение либо необнаружение смоделированной неисправности ОЗУ. В качестве вносимых сложных неисправностей ОЗУ на каждой итерации эксперимента случайным образом выбиралась одна из восьми неисправностей взаимного влияния прямого действия (Idempotent Coupling Faults (*CFid*)) – $\wedge\langle\uparrow, 0\rangle$, $\wedge\langle\uparrow, 1\rangle$, $\wedge\langle\downarrow, 0\rangle$, $\wedge\langle\downarrow, 1\rangle$, $\vee\langle\uparrow, 0\rangle$, $\vee\langle\uparrow, 1\rangle$, $\vee\langle\downarrow, 0\rangle$, $\vee\langle\downarrow, 1\rangle$ – и определялся адрес ячейки агрессора (Aggressor Cell) и ячейки жертвы (Victim Cell) [5].

Далее выполнялся запуск процедуры тестирования Test (TMATS). Первоначально проводилось тестирование на первом случайном состоянии ячеек ОЗУ, т. е. первом тестовом наборе вероятностного теста. В случае несовпадения сигнатур, реальной и эталонной, фиксировался факт обнаружения неисправности и выполнялся переход к следующей итерации вероятностного теста. При совпадении сигнатур осуществлялся переход к тестированию на втором тестовом наборе вероятностного теста текущей итерации. В случае совпадения сигнатур для второго набора проводилось тестирование на третьем наборе. Таким образом, устанавливался факт обнаружения либо необнаружения внесенной неисправности взаимного влияния для каждой итерации вероятностного теста. По окончании всех заданных итераций вероятностного теста вычислялась процентная полнота покрытия $FC_{Test}(Fault, 4)$ неисправностей взаимного влияния.

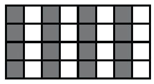
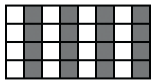
Аналогичным образом последовательно выполнялось тестирование и на фиксированных начальных состояниях, соответствующих управляемым вероятностным тестам с заданным расстоянием Хэмминга. Применялись тесты, состоящие из четырех тестовых наборов, обеспечивающих $\max\_\min HD(T_x, T_y) \geq n/2$. Пример подобного управляемого вероятностного теста T графически представлен в табл. 4 для ОЗУ, состоящего из матрицы 4Ч8 запоминающих ячеек ($n = 32$), обеспечивающего $\max\_\min HD(T_x, T_y) \geq 16$.

Таблица 4. Управляемый вероятностный тест для оперативного запоминающего устройства, состоящий из четырех наборов

Table 4. A controlled probabilistic test for random access memory consisting of four sets

T	ОЗУ	SF
T_1		0000111100100000
T_2		1111010011100111

Окончание табл. 4
Ending of Tab. 4

T	ОЗУ	SF
T_3		0101100110011101
T_4		1010001001011010

Как видно из табл. 4, тест состоит из четырех тестовых наборов, для каждого из которых приведено эталонное значение сигнатуры SF. Результаты экспериментальных исследований для различного количества итераций (100, 1000, 10000, – для вероятностных тестов (Set 1) и управляемых вероятностных тестов (Set 2)) приведены на рис. 1.

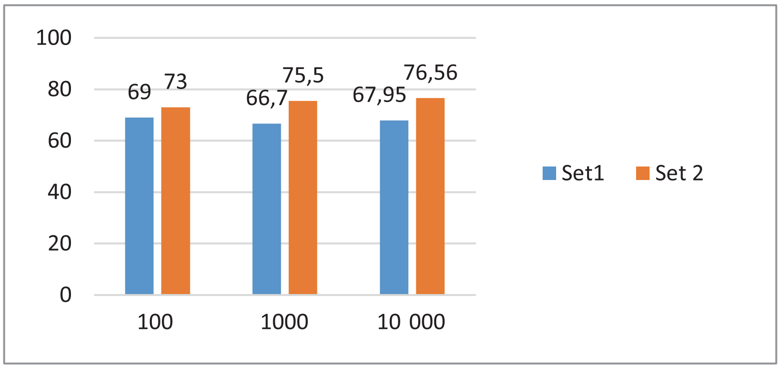


Рис. 1. Полнота покрытия неисправностей оперативного запоминающего устройства
Fig. 1. Complete coverage of operational memory device faults

На рис. 1 показаны экспериментальные данные о полноте покрытия $FC_{Test}(Fault, 4)$ для случая неисправностей взаимного влияния ОЗУ. Как видно из приведенных значений, в случае управляемых вероятностных тестов полнота покрытия заметно выше полноты покрытия таких же неисправностей вероятностными тестами.

Заключение

1. Представлен метод построения управляемых вероятностных тестов с заданным расстоянием Хэмминга. В отличие от вероятностных тестов, предложенная процедура формирования тестовых наборов управляемых вероятностных тестов характеризуется невысокой вычислительной сложностью, определяемой сложностью генерирования всех символов рассматриваемого алфавита в произвольном порядке.
2. В простейшем случае построение тестовых наборов заключается в последовательном переборе всего алфавита используемой системы счисления. На примере оперативного запоминающего устройства показана эффективность управляемых вероятностных тестов с заданным расстоянием Хэмминга по сравнению с вероятностными тестами.
3. Дальнейшие исследования целесообразно расширить в части свойств предложенного метода и его использования для различных задач тестового диагностирования вычислительных систем и их компонентов. Наиболее интересным представляется применение метода формирования управляемых вероятностных тестов для тестирования программного обеспечения.

Список литературы

1. Renfer, G. F. Automatic Program Testing / G. F. Renfer // Proceedings of 3rd Conference of the Computing and Data Processing Society of Canada. Canada: University of Toronto Press, 1962.
2. Duran, J. An Evaluation of Random Testing / J. Duran, S. Ntafos // IEEE Transactions on Software Engineering. 1984. Vol. SE-10, No 4. P. 438–444.
3. Arcuri, A. Random Testing: Theoretical Results and Practical Implications / A. Arcuri, M. Z. Iqbal, L. Briand // IEEE Transactions on Software Engineering. 2011. Vol. 38, No 2. P. 258–277.

4. Malaiya, Y. K. The Coverage Problem for Random Testing / Y. K. Malaiya, S. Yang // Proceedings of ITC. USA, 1984. P. 237–245.
5. Ярмолик, В. Н. Контроль и диагностика вычислительных систем / В. Н. Ярмолик. Минск: Бестпринт, 2019.
6. An Orchestrated Survey on Automated Software Test Case Generation / S. Anand [et al.] // Journal of Systems and Software. Elsevier. 2014. Vol. C-39, No 4. P. 582–586.
7. Grindal, M. Combination Testing Strategies / M. Grindal, J. Offutt, S. F. Andler // GMU Technical Report ISE-TR-04-05. USA: George Mason University, 2004.
8. Malaiya, Y. K. Antirandom Testing: Getting the Most Out of Black-Box Testing / Y. K. Malaiya // International Symposium on Software Reliability Engineering. France: Toulouse, 1995. P. 86–95.
9. A Survey on Adaptive Random Testing / R. Huang [et al.] // IEEE Transactions on Software Engineering. 2021. Vol. 47, No 10. P. 2052–2083.
10. Ярмолик, С. В. Управляемое случайное тестирование / С. В. Ярмолик, В. Н. Ярмолик // Информатика. 2011. Т. 29, № 1. С. 79–88.
11. Peterson, W. W. Error-Correction Codes / W. W. Peterson, E. J. Weldon. London: The MIT Press, 1972.
12. Ярмолик, С. В. Синтез вероятностных тестов с малым числом наборов / С. В. Ярмолик, В. Н. Ярмолик // Автоматика и вычислительная техника. 2011. Т. 45, № 3. С. 19–30.
13. Yarmolik, S. V. Controlled Random Tests / S. V. Yarmolik, V. N. Yarmolik // Automation and Remote Control. Springer. 2012. Vol. 73, No 10. P. 1704–1714.
14. Садовский, М. Г. О сравнении символьных последовательностей / М. Г. Садовский // Вычислительные технологии. 2005. Т. 10, № 3. С. 106–116.
15. Ярмолик, В. Н. Модификации способов определения расстояния Хэмминга для их применения в качестве мер различия при генерировании управляемых вероятностных тестов / В. Н. Ярмолик, В. В. Петровская, Н. А. Шевченко // Информатика. 2024. Т. 21, № 2. С. 54–72.

References

1. Renfer G. F. (1962) Automatic Program Testing. *Proceedings of 3rd Conference of the Computing and Data Processing Society of Canada*. Canada, University of Toronto Press.
2. Duran J., Ntafos S. (1984) An Evaluation of Random Testing. *IEEE Transactions on Software Engineering*. SE-10 (4), 438–444.
3. Arcuri A., Iqbal M. Z., Briand L. (2011) Random Testing: Theoretical Results and Practical Implications. *IEEE Transactions on Software Engineering*. 38 (2), 258–277.
4. Malaiya Y., Yang S. (1984) The Coverage Problem for Random Testing. *Proceedings of ITC*, USA. 237–242.
5. Yarmolik V. N. (2019) *Control and Diagnostics of Computing Systems*. Minsk, Bestprint Publ. (in Russian).
6. Anand S., Burke E., Chen T., Clark J., Cohen M., Grieskamp W., et al. (2014) An Orchestrated Survey on Automated Software Test Case Generation. *Journal of Systems and Software. Elsevier*. C-39 (4), 582–586.
7. Grindal M., Offutt J., Andler S. F. (2004) Combination Testing Strategies. *GMU Technical Report ISE-TR-04-05*. USA, George Mason University.
8. Malaiya Y. K. (1995) Antirandom Testing: Getting the Most Out of Black-Box Testing. *International Symposium on Software Reliability Engineering*. France, Toulouse. 86–95.
9. Huang R., Sun W., Xu Y., Chen H., Towey D., Xia X. (2021) A Survey on Adaptive Random Testing. *IEEE Transactions on Software Engineering*. 47 (10), 2052–2083.
10. Yarmolik S. V., Yarmolik V. N. (2011) Controlled Random Testing. *Informatika*. 29 (1), 79–88 (in Russian).
11. Peterson W. W., Weldon E. J. (1972) *Error-Correction Codes. Cambridge*. London, The MIT Press.
12. Yarmolik S. V., Yarmolik V. N. (2011) Synthesis of Probabilistic Tests with a Small Number of Sets. *Automatic Control and Computer Sciences*. 45 (3), 19–30 (in Russian).
13. Yarmolik S. V., Yarmolik V. N. (2012) Controlled Random Tests. *Journal of Automation and Remote Control. Springer*. 73 (10), 1704–1714.
14. Sadovsky M. G. (2005) Compare Symbol Sequences. *Journal of Computational Technologies*. 10 (3), 106–116 (in Russian).
15. Yarmolik V. N., Petrovskaya V. V., Shevchenko N. A. (2024) Modifications to Hamming Distance Calculations for Use as Dissimilarity Measures to Generate Controlled Random Test. *Informatika*. 21 (2), 54–72 (in Russian).

Вклад авторов

Ярмолик В. Н. предложил идею построения управляемых вероятностных тестов с заданным значением расстояния Хэмминга.

Деменковец Д. В. участвовал в постановке и реализации эксперимента, обсуждении полученных результатов и оформлении рукописи статьи.

Леванцевич В. А., Петровская В. В. участвовали в постановке эксперимента, обсуждении полученных результатов и оформлении рукописи статьи.

Authors' contribution

Yarmolik V. N. proposed the idea of constructing controlled probabilistic tests with a given value of the Hamming distance.

Demenkovets D. V. participated in setting up and implementation of the experiment, discussing the results obtained, and preparing the manuscript.

Levantsevich V. A., Petrovskaya V. V. participated in setting up the experiment, discussing the results obtained, and preparing the manuscript.

Сведения об авторах

Ярмолик В. Н., д-р техн. наук, проф., проф. каф. программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники (БГУИР)

Деменковец Д. В., магистр техн. наук, ст. преп. каф. программного обеспечения информационных технологий, БГУИР

Леванцевич В. А., магистр техн. наук, ст. преп. каф. программного обеспечения информационных технологий, БГУИР

Петровская В. В., магистр техн. наук, асс. каф. программного обеспечения информационных технологий, БГУИР

Адрес для корреспонденции

220013, Республика Беларусь,
г. Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 44 700-88-96
E-mail: demenkovets@bsuir.by
Деменковец Денис Викторович

Information about the authors

Yarmolik V. N., Dr. of Sci. (Tech.), Professor, Professor at the Department of Software for Information Technologies, Belarusian State University of Informatics and Radioelectronics (BSUIR)

Demenkovets D. V., M. of Sci., Senior Lecture at the Department of Software for Information Technologies, BSUIR

Levantsevich V. A., M. of Sci., Senior Lecture at the Department of Software for Information Technologies, BSUIR

Petrovskaya V. V., M. of Sci., Assistant Professor at the Department of Software for Information Technologies, BSUIR

Address for correspondence

220013, Republic of Belarus,
Minsk, Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 44 700-88-96
E-mail: demenkovets@bsuir.by
Demenkovets Denis Victorovich