Автоматизация. Современные Технологии

ЕЖЕМЕСЯЧНЫЙ МЕЖОТРАСЛЕВОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

ИЗДАЁТСЯ С 1947 ГОДА

8

2015

УЧРЕДИТЕЛИ:

ОБЩЕСТВО
С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ
"НАУЧНО-ТЕХНИЧЕСКОЕ ИЗДАТЕЛЬСТВО
"ИННОВАЦИОННОЕ МАШИНОСТРОЕНИЕ"

Журнал зарегистрирован 29 мая 2014 г.
за № ФС77-58102 в Роскомнадзоре

Журнал входит в перечень утверждённых ВАК РФ изданий для публикации трудов соискателей учёных степеней, а также в систему Российского индекса научного цитирования (РИНЦ)

ООО "Издательство
"Инновационное машиностроение"

Адрес редакции:
107076, Москва,
Колодезный пер., д. 2А, стр. 2.
Тел.: (495) 661-38-80.
E-mail: ast@mashin.ru; astmashin@yandex.ru;
http://www.mashin.ru

Адрес издательства:
107076, Москва, Колодезный пер., д. 2А, стр. 2.
Тел.: (499) 268-38-58.
Факс: (499) 269-48-97.

СОДЕРЖАНИЕ

АВТОМАТИЗАЦИЯ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ И ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ

Ларченко А.Г. Система автоматизированного управления высокочастотным диагностированием изделий из полимерных материалов . . . 3

Бугров Ю.Н., Курнасов Е.В. АЦП параллельного типа с преобразованием двоично-десятичных чисел в двоичный код для интерфейса управляющих устройств . . . 9

СИСТЕМЫ И ПРИБОРЫ УПРАВЛЕНИЯ

Власов А.И., Зотьева Д.Е., Евдокимов В.С., Ревзин Г.Г., Феоктистов Д.В. Гибридная система управления малыми беспилотными летательными аппаратами . . . 15

Пролетарский А.В., Пронин А.В., Смирнова Е.В., Шпак М.А. Автоматизация процедуры миграции данных из реляционной базы данных в онтологию предметной области на примере медицинской информационной системы . . . 25

Цибизова Т.Ю., Чан Нгок Хыонг. Способы реализации процедуры идентификации на основе фильтра Вольтерра . . . 31

Фролова И.В. Современные вопросы защиты информации в измерительно-информационных системах . . . 34

СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Шибанов Г.П. Оценка эффективности внедрения новых технологий в перспективные авиационные комплексы. . . 37

Головешкин В.А., Выборнов А.Н., Ульянов М.В. Операционная чувствительность алгоритмов . . . 41

ОБЗОР ПЕРИОДИЧЕСКОЙ ПЕЧАТИ

По страницам журналов . . . 47

CONTENTS

AUTOMATION OF SCIENTIFIC-RESEARCH AND PRODUCTION PROCESSES

Larchenko A.G. Automatic control system for high-frequency diagnosis of products from polymeric materials . . . 3

Bugrov Yu.N., Kurnasov E.V. Parallel type АЦП with the transformation of the binary-decimal numbers into binary code for the interface control devices . . . 9

SYSTEMS AND CONTROL DEVICES

Vlasov A.I., Zoteva D.E., Evdokimov V.S., Revzin G.G., Feoktistov D.V. Hybrid control system of the small unmanned ferial vehicles (UAVs) . . . 15

Proletarskiy A.V., Pronin A.V., Smirnova E.V., Shpak M.A. Automation of the data migration procedure from a relational database in the object area ontology for medical information system as an example . . . 25

Tsibizova T.Yu., Chan Nrok Khyong. Identification methods of implementation procedures from a Volterra filter . . . 31

Frolova I.V. Modern questions over the information protection in measuring and information systems. . . 34

MODERN INFORMATION TECHNOLOGY

Shibanov G.P. Implementation efficiency valuation of the new technologies in the perspective aircraft complexes. . . 37

Goloveshkin V.A., Vybornov A.N., Ylyanov M.V. Algorithms operating sensitivity . . . 41

SURVEY OF PERIODICALS

On the pages of magazines . . . 47

Журнал распространяется по подписке, которую можно оформить в любом почтовом отделении (индекс по каталогу "Роспечать" — 70537, по каталогу "Пресса России" — 27838, по каталогу российской прессы "Почта России — 60267) или непосредственно в издательстве по факсу (499) 269 4897, по e-mail: realiz@mashin.ru, на сайте www.mashin.ru (без почтовых наценок, с любого месяца, со своего рабочего места); телефоны для справок: (499) 269 5298

Сдано в набор 02.06.2015.
Подписано в печать 23.07.2015.
Формат 60 × 88 1/8. Бумага офсетная.
Усл. печ. л. 5,88. Цена свободная.

Отпечатано в ООО "Канцлер"
150008, г. Ярославль, ул. Клубная, д. 4, кв. 49.
Оригинал-макет: ООО "Адвансед солюшнз".
119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: www.aov.ru

УДК 004.421

В.А. Головешкин, д-р техн. наук, проф., **А.Н. Выборнов**, канд. физ.-мат. наук, доц.
(Московский государственный университет приборостроения и информатики),
М.В. Ульянов, д-р техн. наук, проф.
(Институт проблем управления им В.А. Трапезникова РАН, г. Москва)
muljanov@mail.ru

ОПЕРАЦИОННАЯ ЧУВСТВИТЕЛЬНОСТЬ АЛГОРИТМОВ

Рассмотрена новая оценка качества алгоритмов — операционная чувствительность алгоритма по функции трудоёмкости и особенности её применения. Предложен вариант функционала операционной чувствительности, т.е. её количественной оценки на основе информационной чувствительности алгоритма и размаха варьирования трудоёмкости. Предложенная оценка проиллюстрирована на модельном примере.

Ключевые слова: алгоритмическое обеспечение; операционная чувствительность; критерий временной устойчивости.

A new valuation of the algorithms quality — the algorithm operating sensitivity on function of labor input and feature of its application is considered. A variant of the functional operating sensitivity, ie, its quantitative valuation from a algorithm information sensitivity and labor input scaling swing is proposed. The proposed valuation is illustrated in the model example.

Key words: algorithmic ensuring; operating sensitivity; temporal stability criterion.

Введение. Решение задачи выбора алгоритма, рационального по комплексному критерию качества, для данной проблемной области применения является актуальной задачей разработки алгоритмического обеспечения программных средств. В таких проблемных областях, как системы реального времени и распределённые вычисления к применяемым алгоритмам, предъявляются жёсткие требования не только по собственно временной эффективности, но и по временной устойчивости [1], что связано с необходимостью обеспечения балансировки загрузки процессоров. В этой связи актуальным является построение такой оценки качества алгоритма, которая объединяла бы эти требования.

Одной из наиболее важных ресурсных функций алгоритма является функция трудоёмкости, отражающая требования алгоритма к операционному ресурсу механизма реализации. Функция трудоёмкости в худшем и лучшем случаях определяется, соответственно, как максимальное и минимальное число элементарных операций в принятой модели вычислений, заданных алгоритмом на входах фиксированной длины [2]. При исследовании временной эффективности программных реализаций алгоритма функция трудоёмкости ассоциирована со временем получения результата на определённых исходных данных. В этом аспекте результаты ресурсного анализа алгоритма должны обеспечивать возможность прогнозирования его трудоёмкости как для разных длин входов, так и для различных входов фиксированной длины. Общепринятым является подход, при котором прогнозирование по изменению длин входов осуществляется на основе трудоёмкости в среднем, а прогнозирование на фиксированной длине входа — по трудоёмкостям в лучшем и худшем случаях [3]. В аспекте решения задачи прогнозирования на фиксированной длине входа количественно-зависимые алгоритмы (класс N [4]) представляют собой наиболее благоприятный класс. Но, к сожалению, подавляющее большинство алгоритмов относятся к количественно-параметрическому классу (класс NPR [4]) и обладают ненулевым, а в ряде случаев, и значительным размахом варьирования трудоёмкости при фиксированной длине входа. Для целого ряда алгоритмов (на актуальных длинах входа) значение этого размаха достаточно велико. Прогнозирование по трудоёмкости в худшем случае (гарантированная оценка сверху) даёт для этого класса почти всегда сильно завышенные результаты, а прогнозирование по трудоёмкости в среднем не учитывает информацию о размахе варьирования. При этом очевидно, что качество прогноза во многом определяется тем, насколько велико влияние различных входов фиксированной длины на трудоёмкость, т. е. насколько велика информационная чувствительность исследуемого алгоритма [5] в рамках выбранной количественной оценки. В связи с этим, в аспекте проблематики прогнозирования временной эффективности для программных реализаций алгоритмов и оценки их стабильности по времени для последующего рационального выбора, представляет интерес задача введения оценки качества алгоритма, отражающей его операционную стабильность.

Далее в статье будет использоваться следующая терминология и обозначения [1, 2], связанные с анализом ресурсной эффективности алгоритмов:

D — вход алгоритма A: конечное множество слов фиксированной длины в бинарном алфавите, задающее конкретную решаемую задачу;

$n = \lambda(D)$ — длина входа алгоритма: $D \underset{\lambda}{\rightarrow} N$, целочисленная функция, в общем случае определяемая как функция от мощности множества D: $\lambda(D) = \lambda(|D|) = n$;

$f_A(D)$ — трудоёмкость алгоритма A на входе D — целочисленная функция, значение которой есть число элементарных операций (в принятой модели вычислений), заданных алгоритмом A на входе D;

$D_n = \{D|\lambda(D) = n\}$ множество всех входов алгоритма A, имеющих длину n;

$f_A^{\wedge}(n)$ — трудоёмкость алгоритма в худшем случае на всех допустимых входах длины n, т. е. максимум $f_A(D)$ на множестве D_n;

$f_A^{\vee}(n)$ — трудоёмкость алгоритма в лучшем случае, — минимум $f_A(D)$ на множестве D_n.

При этом для всех классов алгоритмов всегда выполнено: $f_A^{\vee}(n) \leqslant f_A(D \in D_n) \leqslant f_A^{\wedge}(n)$.

Информационная чувствительность и её статистическая оценка. Впервые понятие информационной чувствительности алгоритма по трудоёмкости введено в работе [5]. Понятие «информационная чувствительность» отражает тот факт, что на разных входах D, имеющих фиксированную длину n, алгоритм задаёт различное число элементарных операций принятой модели вычислений. Ключевым для содержательной интерпретации этого термина является выбор функционала — количественной оценки, обладающей свойством сопоставимости, т. е. дающей возможность решения задач сравнения алгоритмов и их рационального выбора. Разброс значений трудоёмкости алгоритма на различных входах фиксированной длины означает его ненулевую информационную чувствительность. В общем случае причиной такой вариации является влияние значений элементов конкретного входа и/или их порядка, равно как и других содержательных особенностей входов при их фиксированной длине на число задаваемых алгоритмом элементарных операций. Теоретические границы такого разброса на входах длины n задаются значениями функций $f_A^{\wedge}(n)$ и $f_A^{\vee}(n)$. Рассматривая трудоёмкость алгоритма на фиксированной длине входа как дискретную ограниченную случайную величину, можно либо теоретически обосновать задаваемый алгоритмом закон распределения, либо на основе экспериментальных данных вычислить статистические оценки такого распределения и предложить некоторую аппроксимацию, удовлетворяющую одному из критериев согласия. Таким образом, основной идеей построения количественной оценки информационной чувствительности является рассмотрение трудоёмкости алгоритма как дискретной ограниченной случайной величины. Для решения задач прогнозирования и рационального выбора такая оценка должна, очевидно, вводиться как функция длины входа алгоритма.

Статистическая оценка информационной чувствительности предложена в работе [5] на основе следующих рассуждений. На множестве входов фиксированной длины трудоёмкость алгоритма рассматривается как дискретная ограниченная случайная величина. Классической точечной мерой рассеяния случайной величины является σ — среднеквадратическое отклонение, которое оценивается по данным выборки через выборочную исправленную дисперсию [6]. Корректное определение количественной оценки информационной чувствительности должно учитывать также и длину сегмента варьирования. Это связано с тем, что при одинаковом значении дисперсии достоверно более чувствительным должен быть алгоритм с большей длиной сегмента возможных значений трудоёмкости. Для учёта длины этого сегмента используется такое понятие математической статистики, как размах варьирования [6]. Очевидно, что теоретический размах варьирования трудоёмкости алгоритма является функцией длины входа, и, вводя для него обозначение $R(n)$, получаем:

$$R(n) = f_A^{\wedge}(n) - f_A^{\vee}(n). \quad (1)$$

На этой основе в работе [5] вводится понятие нормированного (относительного) размаха варьирования функции трудоёмкости для входов длины n — $R_N(n)$ как отношение половины теоретического размаха варьирования к его середине:

$$R_N(n) = \frac{f_A^{\wedge}(n) - f_A^{\vee}(n)}{f_A^{\wedge}(n) + f_A^{\vee}(n)}. \qquad (2)$$

При этом значение $R_N(n) = 0$ соответствует ситуации, когда $f_A^{\wedge}(n) - f_A^{\vee}(n) = 0$, т. е. идентифицирует принадлежность алгоритма к классу N — классу количественно-зависимых алгоритмов.

Одной из общепринятых точечных характеристик выборки является коэффициент вариации V, определяемый как отношение стандартного отклонения к среднему значению [6]. Для выборки, полученной на основе экспериментального исследования трудоёмкости алгоритма, коэффициент вариации V также зависит от размерности и имеет вид:

$$V(n) = s_{f_A}(n) \big/ \overline{f_A}(n), \; 0 \leqslant V(n) \leqslant 1, \qquad (3)$$

где $s_{f_A}(n)$ — выборочное стандартное отклонение трудоёмкости как дискретной ограниченной случайной величины при фиксированной длине входа n; $\overline{f_A}(n)$ — выборочное среднее, рассчитанное по данным выборки.

С учётом этих рассуждений в работе [5] на основе формул (2) и (3) вводится статистическая количественная оценка информационной чувствительности алгоритма по трудоёмкости:

$$\delta_{IS}(n) = V(n)\, R_N(n), \; 0 \leqslant \delta_{IS}(n) \leqslant 1. \qquad (4)$$

Поскольку оценка $\delta_{IS}(n)$ использует только статистические точечные оценки трудоёмкости как случайной величины, то её применение возможно и при отсутствии знаний о законе распределения значений трудоёмкости или какой-либо его аппроксимации. Таким образом, значения оценки $\delta_{IS}(n)$ могут быть получены на основе экспериментальных исследований алгоритма, по результатам которых вычисляется значение $V(n)$, и его теоретического анализа, необходимого для вычисления нормированного размаха варьирования по формуле (2).

Содержательно важным недостатком статистической количественной меры является то, что она не позволяет указать интервал значений трудоёмкости при заданной вероятности, а следовательно, не может быть адаптирована к требованиям разработчиков алгоритмического обеспечения, в частности, для обеспечения требований стабильности по времени.

Квантильная оценка информационной чувствительности. Задача теоретического обоснования закона распределения, соответствующего данному алгоритму, даже при фиксированной длине входа является достаточно сложной. Идея рассмотрения трудоёмкости алгоритма при фиксированной длине входа как ограниченной случайной величины, аппроксимируемой некоторой известной функцией плотности распределения вероятностей, привела к введению понятия доверительной трудоёмкости [7] на основе вычисления γ-квантиля аппроксимирующего закона распределения. На основе этого подхода в работе [8] была предложена и квантильная оценка информационной чувствительности алгоритмов. Основная идея состоит в определении длины сегмента нормированных значений трудоёмкости, по которому интеграл от функции плотности равен заданной вероятности (надёжности) γ, отражающей требования разработчиков по информационной чувствительности, а по сути — особенности проблемной области применения алгоритма. Нормированные значения трудоёмкости вычисляются по формуле

$$x_f = \frac{f - f_A^{\vee}(n)}{f_A^{\wedge}(n) - f_A^{\vee}(n)} = \frac{f - f_A^{\vee}(n)}{R(n)}. \qquad (5)$$

Содержательно такая количественная оценка с обозначением $\delta_{IQ}(\gamma)$ [8] есть доля теоретического сегмента варьирования трудоёмкости, в которой с заданной вероятностью γ будут наблюдаться значения трудоёмкости алгоритма на произвольных входах фиксированной длины. В общем случае задача определения сегмента, по которому интегрируется заданная вероятность при известной функции плотности, имеет бесконечное множество решений [9]. Традиционное решение этой задачи — определение длины сегмента, задающего вероятность γ, симметрично относительно медианы распределения. При известной функции распределения вероятностей эта задача решается на основе вычисления $1/2 \pm \gamma/2$-квантилей этого распределения. Такая оценка информационной чувствительности является характеристикой алгоритма для входов фиксированной длины. Очевидно, что для практического применения и теоретического исследования алгоритма необходимо рассматривать $\delta_{IQ}(\gamma)$ не только как функцию вероятности γ, но и как функцию длины входа n, т. е. $\delta_{IQ} = \delta_{IQ}(\gamma, n)$.

Пусть $\Phi = \{f(n, x)\}$, $x \in [0, 1]$ есть семейство непрерывных функций плотности распределения, аппроксимирующих нормированные значения трудоёмкости, параметризированное аргументом n — длиной входа алгоритма. Зафиксируем некоторое значение длины входа n, тогда для любой непрерывной, не обращающейся в нуль на интервалах, функции плотности распределения вероятностей $f(n, x)$ из семейства Φ соответствующая интегральная функция распределения вероятностей $F(n, x)$ является монотонно возрастающей, в силу чего имеет обратную функцию. Обозначим через $F^{-1}(n, x)$ функцию, обратную к $F(n, x)$, тогда согласно [9]

$$\delta_{IQ}(\gamma, n) = F^{-1}\left(n, \frac{1}{2} + \frac{\gamma}{2}\right) - F^{-1}\left(n, \frac{1}{2} - \frac{\gamma}{2}\right). \quad (6)$$

Интуитивно понятно, что с уменьшением разброса значений трудоёмкости относительно математического ожидания, приводящего к уменьшению дисперсии, информационная чувствительность алгоритма по трудоёмкости будет падать. Заметим, что использование медианы, а не математического ожидания в определении квантильной оценки информационной чувствительности связано с тем, что медиана и математическое ожидание не обязательно совпадают для несимметричных функций плотности распределения с ограниченной вариацией. Использование в этом случае значения математического ожидания как центра сегмента интегрирования функции плотности может привести, для значений γ, близких к единице, к выходу границ этого сегмента за границы определения функции плотности с ограниченной вариацией.

Использование квантильной оценки позволяет разработчикам, варьируя значение вероятности γ, гибко оценивать информационную чувствительность, учитывая требования технического задания на разработку программ. Заметим, что значение $\gamma = 1$ приводит к значению $\delta_{IQ}(1, n) = 1 \forall n$, т. е. информационной чувствительности по полному нормированному размаху варьирования. Отметим, что эта оценка не содержит информацию о положении γ-квантиля закона распределения на нормированном сегменте. Однако этот недостаток легко устраним при построении функционала операционной оценки.

Количественная оценка операционной чувствительности. Для более детального анализа алгоритма для разработчиков представляет интерес не только собственно длина сегмента, но и его положение, равно как и зависимость этих величин от длины входа. Таким образом модифицированная квантильная оценка информационной чувствительности запишется в виде:

$$\delta^*_{IQ}(\gamma, n) = \left(x_\gamma^{(1)}(n),\ x_\gamma^{(2)}(n),\ x_\gamma^{(2)}(n) - x_\gamma^{(1)}(n)\right), \quad (7)$$

т. е. количественная оценка $\delta^*_{IQ}(\gamma, n)$ задаётся тремя числами: значениями левой и правой границ нормированного сегмента, по которому интегрируется заданная вероятность γ, и длиной этого сегмента.

Однако для разработчиков более актуальной является информация, представленная не в относительных, а в абсолютных единицах — их интересует чувствительность алгоритма в элементарных операциях модели вычислений, на основе которой можно перейти к оценке разброса времени выполнения (оценке стабильности по времени программной реализации алгоритма). Такой переход от трудоёмкости к временной эффективности может быть осуществлён, например, по методике, предложенной в работе [3].

По замыслу авторов, операционная чувствительность должна показывать границы изменения трудоёмкости алгоритма на фиксированной длине входа при заданной исследователем доверительной вероятности. Для этого авторы предлагают использовать оценку $\delta^*_{IQ}(\gamma, n)$, прийдя в ней к абсолютным значениям. Обращение формулы (5) приводит к тому, что значению $x_\gamma^{(1)}(n)$ (рис. 1) соответствует $f_A^\vee(n) + x_\gamma^{(1)}(n)R(n)$, значению $x_\gamma^{(2)}(n)$ — $f_A^\vee(n) + x_\gamma^{(2)}(n)R(n)$, а длине нормированного сегмента $x_\gamma^{(2)}(n) - x_\gamma^{(1)}(n)$ — сегмент значений трудоёмкости, связанный с информационной

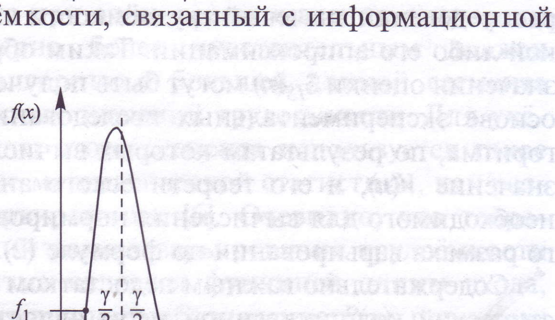

Рис. 1. Квантильная количественная мера информационной чувствительности: $\delta_{IQ}(\gamma, n) = x_2 - x_1$

чувствительностью $\delta_{IQ}(\gamma, n)$ соотношением $\delta_{IQ}(\gamma, n)R(n)$, где $R(n)$ определится по формуле (1).

Таким образом, вводим оценку операционной чувствительности с обозначением $\delta_{OP}(\gamma, n)$ на основе модифицированной квантильной оценки информационной чувствительности, размаха варьирования и обращения формулы (5):

$$\delta_{OP}(\gamma, n) = \\ = \begin{pmatrix} f_A^{\vee}(n) + x_{\gamma}^{(1)}(n)R(n), \\ f_A^{\vee}(n) + x_{\gamma}^{(2)}(n)R(n), \delta_{IQ}(\gamma, n)\ R(n) \end{pmatrix}. \quad (8)$$

Более корректно $\delta_{OP}(\gamma, n)$ есть функционал, отображающий алгоритм и множество входов фиксированной длины в три числовых значения, т.е. отображающий пару (A, D_n) в R^3, где A — алгоритм, D_n — множество входов фиксированной длины.

Методика экспериментального определения операционной чувствительности. Непосредственное применение формулы (8) приводит к необходимости построения аппроксимирующей функции плотности для наблюдаемых значений трудоёмкости. Если исследователь может быть удовлетворён менее точным определением границ сегмента информационной чувствительности, то в этом случае предлагается следующий, более простой метод определения значений $\delta_{OP}(\gamma, n)$:

1. Фиксация интересующей исследователя длины входа n и проведение экспериментального исследования программной реализации алгоритма для определения числа заданных элементарных операций. Рекомендуемое число экспериментов m не менее 10 000, при этом генерация входов может выполняться либо на основе псевдослучайного равномерного генератора, либо с учётом особенностей входов, определяемых областью применения.

2. Построение гистограммы относительных частот полученных значений трудоёмкости при разбиении полученного экспериментального размаха варьирования не более чем на $\sqrt{m}$ (правая граница рекомендуемого числа полусегментов [9]) равных полусегментов. При m = 10 000 получим 100 полусегментов гистограммы.

3. Задание доверительной вероятности γ для операционной чувствительности и определение значения $\alpha = (1 - \gamma)/2$.

4. Суммирование относительных частот в левых и правых полусегментах гистограммы вплоть до достижения полученного значения α. Тем самым отбрасываем левый и правый хвосты распределения (см. рис.1), имеющие вероятность α. Полученные таким образом границы и являются границами в $\delta_{OP}(\gamma, n)$.

Модельный пример. Проиллюстрируем предложенную оценку на основе данных экспериментального исследования [8] алгоритма Кнута—Морриса—Пратта для поиска подстроки в строке [10]. Этот алгоритм использует префиксную функцию для предобработки строки поиска. Теоретические функции трудоёмкости алгоритма Кнута—Морриса—Пратта в случае однократного вхождения подстроки, необходимые для нормирования значений трудоёмкости, получены в работе [11]. При этом трудоёмкость в лучшем случае имеет вид $f_A^{\vee}(n,m) = 14n + 17m - 7$, а функция трудоёмкости в худшем случае равна $f_A^{\wedge}(n,m) = 24n + 17m - 27$, где n — длина строки, m — длина подстроки. Заметим, что функция трудоёмкости для этого алгоритма имеет два аргумента. Для значений n = 10 000, m = 20 получаем теоретический минимум и максимум трудоёмкости для однократного вхождения:

$$f_A^{\vee}(10\,000, 20) = 140\,333,$$
$$f_A^{\wedge}(10\,000, 20) = 240\,313.$$

Экспериментальное исследование было проведено при n = 10 000 и m = 20 по методике, предложенной в работе [12]. Генерировались случайные входы, для которых случайно выбиралась подстрока с однократным вхождением. Проведено 20 000 экспериментов, в каждом из которых фиксировалось значение трудоёмкости — число элементарных операций (в модели вычислений [1]), заданных алгоритмом. На основе этих данных построена гистограмма относительных частот W значений трудоёмкости (рис. 2). Отметим, что поведение

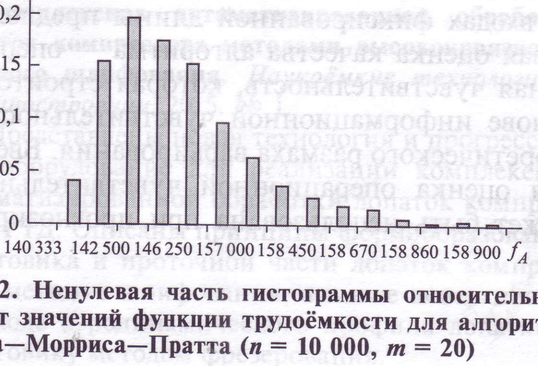

Рис. 2. Ненулевая часть гистограммы относительных частот значений функции трудоёмкости для алгоритма Кнута—Морриса—Пратта (n = 10 000, m = 20)

значений трудоёмкости как ограниченной случайной величины имеет для этого алгоритма и данных параметров входа ярко выраженную асимметрию, в данном случае — левую.

Нормирование значений трудоёмкости в сегмент [0, 1] проведено по теоретически полученным границам. Была выдвинута гипотеза H_0 об аппроксимации гистограммы нормированных относительных частот функцией плотности бета-распределения. Методом моментов [12] определены параметры аппроксимирующего бета-распределения: $\alpha = 1{,}42$; $\beta = 18{,}47$. После вычисления теоретических относительных частот для функции плотности бета-распределения с данными параметрами гипотеза H_0 подтверждена критерием согласия Пирсона. Для вероятности $\gamma = 0{,}95$ вычислено значение $\delta_{IQ}(\gamma) = 0{,}2104$, при этом границы сегмента, соответствующего $1/2 \pm \gamma/2$-квантилям бета-распределения (см. рис. 1), оказались равными: $x_1 = 0{,}0049$, $x_2 = 0{,}2153$. Операционная чувствительность, определённая по формуле (8), для $n = 10\,000$ при $m = 20$

$$\delta_{OP}(0{,}95\ 10\,000,\ 20) = \\ = (140\,823,\ 163\,349,\ 21\,526). \qquad (9)$$

Таким образом, для данного модельного примера при размахе варьирования в 99 980 элементарных операций операционная чувствительность при доверительной вероятности $\gamma = 0{,}95$ составляет 21 526 элементарных операций (менее 22 % размаха варьирования), а ожидаемое число элементарных операций при данной вероятности колеблется от 140 823 до 163 349, при этом положение сегмента операционной чувствительности близко к теоретическому лучшему случаю. Переходя в (9) от обобщённых операций к времени выполнения (для данного компьютера и данного языка программирования), можно получить значения временной чувствительности.

Заключение. На основе вероятностного подхода к описанию трудоёмкости алгоритма на входах фиксированной длины предложена новая оценка качества алгоритма — операционная чувствительность, которая строится на основе информационной чувствительности и теоретического размаха варьирования. Введённая оценка операционной чувствительности может быть использована при прогнозировании временной эффективности компьютерных алгоритмов, а также при решении задачи рационального выбора алгоритмов по критерию временной устойчивости (стабильности времени выполнения на разных входах фиксированной длины), равно как и при решении других задач прикладной теории алгоритмов.

Библиографические ссылки

1. **Ульянов М.В.** Ресурсно-эффективные компьютерные алгоритмы. Разработка и анализ. М.: ФИЗМАТЛИТ, 2008. 304 с.

2. **Ульянов М.В.** Метод прогнозирования временных оценок программных реализаций алгоритмов на основе функции трудоёмкости // Информационные технологии. 2004. № 5. С. 54—62.

3. **Ульянов М.В.** Классификация и методы сравнительного анализа вычислительных алгоритмов. М.: ФИЗМАТЛИТ, 2004. 212 с.

4. **Ульянов М.В., Головешкин В.А.** Информационная чувствительность функции трудоёмкости алгоритмов к входным данным // Новые информационные технологии: Сб. тр. VII Всероссийской науч.-техн. конф. М.: МГАПИ, 2004. С. 19—26.

5. **Ульянов М.В.** Система обозначений в анализе ресурсной эффективности вычислительных алгоритмов // Вестник МГАПИ. Серия: Естественные и инженерные науки. 2004 № 1 (1). С. 42—49.

6. **Гмурман В.Е.** Теория вероятностей и математическая статистика: учеб. пособие для вузов, 9-е изд., М.: Высш. шк., 2003. 479 с.

7. **Ульянов М.В., Петрушин В.Н., Кривенцов А.С.** Доверительная трудоёмкость — новая оценка качества алгоритмов // Информационные технологии и вычислительные системы. 2009. № 2. С. 23—37.

8. **Ульянов М.В., Алексеенко А.С.** Вероятностный подход к определению количественной меры информационной чувствительности компьютерных алгоритмов // Автоматизация и современные технологии. 2009. № 10. С.24—32.

9. **Лагутин М.Б.** Наглядная математическая статистика: учеб. пособие. М.: БИНОМ. Лаборатория знаний, 2007. 472 с.

10. **Гасфилд Д.** Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология. СПб.: Невский диалект, 2003. 654 с.

11. **Алексеенко А.С.** Информационная чувствительность алгоритма Кнута—Морриса—Пратта // Задачи системного анализа, управления и обработки информации: межвуз. сб. науч. тр. Вып. 3 / под общ. ред. Е.В. Никульчева. М.: МГУП, 2010. С. 7—10.

12. **Петрушин В.Н., Ульянов М.В.** Планирование экспериментального исследования трудоёмкости алгоритмов на основе бета-распределения // Информационные технологии и вычислительные системы. 2008. № 2. С. 81—91.