Zagumennov Denis Vladimirovich – Southern Federal University; e-mail: zagumen.denis@gmail.com; Rostov-on-Don, Russia; phone: +79185820982; the department of algebra and discrete mathematics; junior researcher; graduate student.

Mkrtichyan Vyacheslav Vital'evich – e-mail: mkrtichan@list.ru; phone: +79034310555; the department of algebra and discrete mathematics; cand. of eng. sc.; associate professor.

УДК 519.224.22 DOI 10.18522/2311-3103-2022-4-50-62

А.К. Мельников, И.И. Левин, А.И. Дордопуло, Л.М. Сластен

ОЦЕНКА ВОЗМОЖНОСТЕЙ ПЕРСПЕКТИВНЫХ ВЫЧИСЛИТЕЛЬНЫХ ТЕХНОЛОГИЙ ДЛЯ РАСЧЕТА ТОЧНЫХ ПРИБЛИЖЕНИЙ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ ЗНАЧЕНИЙ СТАТИСТИК

Статья посвящена оценке аппаратного ресурса вычислительных систем для решения вычислительно-трудоемкой задачи – расчета распределений вероятностей значений статистик методом второй кратности на основе Δ-точных приближений для выборок объемом от 320 до 1280 знаков при мощности алфавита от 128 до 256 символов с точностью $\Delta=10^{-5}$*. Общее время решения не должно превышать 30 дней или* $2{,}592\cdot10^6$ *секунд при круглосуточном режиме вычислений. Использование свойств метода второй кратности позволяет привести вычислительную сложность расчета к диапазону* $9{,}68\cdot10^{22}$*–*$1{,}60\cdot10^{52}$ *операций с числом проверяемых векторов – от* $6{,}50\cdot10^{23}$ *до* $1{,}39\cdot10^{50}$*. Решение этой задачи для указанных параметров выборок в заданное время с помощью современных вычислительных средств (процессоров, графических ускорителей, программируемых логических интегральных схем) требует недостижимого на практике аппаратного ресурса. Поэтому в статье анализируются возможности перспективных квантовых и фотонных технологий для решения задачи с заданными параметрами. Основным преимуществом квантовых вычислительных систем является высокая скорость вычислений для всех возможных значений параметров. Однако, для расчета распределений вероятностей значений статистик квантовое ускорение не будет достигнуто из-за необходимости проверки всех полученных решений, число которых соответствует размерности задачи. Кроме того, текущий уровень развития элементной базы не позволяет создавать и использовать квантовые вычислители с разрядностью 120 кубитов, необходимой для решения рассматриваемой задачи. Фотонные вычислители могут обеспечить высокую скорость вычислений при низком энергопотреблении и для решения рассматриваемой задачи требуют наименьшее число узлов. Однако, нерешенные проблемы с физической реализацией элементов оперативного хранения данных и отсутствием доступной элементной базы не позволяют в обозримой перспективе (5–7 лет) использовать фотонные вычислительные технологии для расчета распределений вероятностей значений статистик, поэтому наиболее целесообразно применение гибридных вычислительных систем, содержащих узлы различных архитектур. Для реализации задачи на различных аппаратных платформах (универсальные процессоры, графические ускорители, программируемые логические интегральные схемы) и конфигурациях гибридных вычислительных систем предложено использование архитектурно-независимого языка программирования высокого уровня SET@L, объединяющего представление вычислений в виде множеств и совокупностей с помощью альтернативной теории множеств П. Вопенка с абсолютным параллелизмом информационного графа и парадигмами аспектно-ориентированного программирования.*

Вероятность; статистика; точное распределение; точное приближение; алгоритмическая сложность; квантовые вычисления; фотонные технологии; архитектурно-независимое программирование; язык Set@l.

A.K. Melnikov, I.I. Levin, A.I. Dordopulo, L.M. Slasten

ANALYSIS OF ADVANCED COMPUTER TECHNOLOGIES FOR CALCULATION OF EXACT APPROXIMATIONS OF STATISTICS PROBABILITY DISTRIBUTIONS

The paper is devoted to the evaluation of the hardware resource of computer systems for solving a computational-expensive problem such as calculation of the probability distributions of statistics by the second multiplicity method based on Δ-exact approximations for samples with a size of 320-1280 characters and an alphabet power of 128-256 characters, and with an accuracy of $\Delta=10^{-5}$. The total solution time should not exceed 30 days or $2.592\cdot10^6$ seconds for 24/7 computing. Owing to the use of the properties of the second multiplicity method, the computational complexity of the calculations can be brought to the range of $9.68\cdot10^{22}$-$1.60\cdot10^{52}$ operations with the number of tested vectors of $6.50\cdot10^{23}$-$1.39\cdot10^{50}$. The solution of this problem for the specified parameters of samples during the given time requires the hardware resource which cannot be provided by modern computer means such as processors, graphics accelerators, programmable logic integrated circuits. Therefore, in the paper we analyze the possibilities of promising quantum and photon technologies for solving the problem with the given parameters. The main advantage of quantum computer systems is the high speed of calculations for all possible parameter values. However, quantum acceleration will not be achieved to calculate the probability distributions of statistics due to the need to check all the obtained solutions. Here, the number of obtained solutions corresponds to the dimension of the problem. In addition, due to the current development level of the quantum hardware components, it is impossible to create and use the 120-qubit quantum computers for the solution of the considered problem. Photon computers can provide high computation speed at low power consumption and require the smallest number of nodes to solve the considered problem. However, unsolved problems with the physical implementation of efficient memory elements and the lack of available hardware components make the use of photon computer technologies impossible for calculation of the probability distributions of statistics in the near future (5-7 years). Therefore, it is most reasonable to use hybrid computer systems containing nodes of different architectures. To solve the problem on various hardware platforms (general-purpose processors, GPUs, FPGAs) and configurations of hybrid computer systems, we suggest to use an architecture independent high-level programming language SET@L. The language combines the representation of calculations as sets and collections (based on the alternative set theory of P. Vopenka), the absolutely parallel form of the problem represented as an information graph, and the paradigm of aspect-oriented programming.

Probability; statistic; exact distribution; exact approximation; algorithmic complexity; quantum calculations; photonic technologies; architecture independent programming; Set@l language.

Введение. Для решения прикладных задач, требующих статистической обработки текстов как осмысленных символьных последовательностей, используются критерии, минимизирующие количество ложных решений о справедливости проверяемых гипотез. Наибольшей относительной эффективностью обладают критерии на основе точных распределений эталонных статистик [1], но расчет точных распределений – это вычислительно трудоемкая задача [1, 2], зависящая от мощности алфавита N и объема выборки n (длины текстовой последовательности).

Вычислительную сложность задачи можно сократить, если вместо точных распределений использовать их точные *приближения* (предельные распределения), которые минимально снижают эффективность используемых критериев [1, 2]. В качестве точных приближений распределений эталонных статистик используются Δ-точные распределения [2], отличающиеся от исходных на заранее задаваемую сколь угодно малую величину Δ. Для расчета точных распределений существует ряд методов, таких как метод вычисления точного распределения статистик типа Колмогорова-Смирнова [3–4], известный классический метод Монте-Карло [5] и другие. Наиболее предпочтительным является метод второй кратности [2], позволяющий при одинаковом ресурсе рассчитывать точные приближения распределений для максимальных значений параметров выборок. Метод второй кратности, основанный на решении систем линейных уравнений, имеет полиномиальную сложность, но его вычислительная трудоемкость для реальных прикладных задач все равно достаточно велика [2], поэтому расчет точных приближений в оперативно-приемлемое время с помощью современных вычислительных средств затруднен. Оценки необходимого вычислительного ресурса и времени решения задачи с помощью современных процессоров, графических ускорителей и ПЛИС для требуемых на практике параметров распределений приведены в [2]. В данной статье рассматриваются оценки необходимого ресурса и возможного времени решения задачи расчета точных приближений распределений вероятностей значений статистик с помощью перспективных вычислительных технологий – квантовых и фотонных вычислительных систем.

Постановка задачи. Распределения вероятностей значений статистик рассматриваются для алфавита $A_N = \{a_1, \dots, a_N\}$ мощностью $|A_N| = N$ и его выборки объёмом *n*. Для расчета точных приближений распределений вероятностей значений статистики $P_\Delta\{S_{N,n} \ge c\}|$, отличающихся от точных распределений $P_T\{S_{N,n} \ge c\}$ на задаваемую заранее сколь угодно малую величину Δ

$$| P_T\{S_{N,n} \ge c\} - P_\Delta\{S_{N,n} \ge c\}| \le \Delta \tag{1}$$

применяется метод второй кратности (МВК), основанный на решении системы линейных уравнений

$$\begin{cases} \mu_0^{(\upsilon)} + \mu_1^{(\upsilon)} + \dots + \mu_n^{(\upsilon)} = N \quad , \\ 1 \cdot \mu_1^{(\upsilon)} + 2 \cdot \mu_2^{(\upsilon)} + \dots + n \cdot \mu_n^{(\upsilon)} = n. \end{cases} \tag{2}$$

где $\mu_j^{(\upsilon)}$ – это количества знаков алфавита A_N, встретившихся в υ-й выборке объёма *n* ровно *j* раз.

Метод МВК [2] основан на последовательном переборе векторов $\mu^{(\upsilon)} = \{\mu_0^{(\upsilon)}, \mu_1^{(\upsilon)}, \dots, \mu_n^{(\upsilon)}\}$ и проверке каждого $\mu^{(\upsilon)}$ на принадлежность к решениям системы линейных уравнений (2). Подробный теоретический обзор применимости и особенностей реализации МВК рассмотрен в [2]. Алгоритмическая сложность метода МВК определяется выражением

$$\begin{aligned} C_{MBK}(P_\Delta\{S_{N,n} \ge c\}) = L_{\mu(N,n,r)} \times 5 \cdot r + \\ + K_\mu(N,n,r) \cdot (\, 5 \cdot (r+1) + 2(N+r) + 3\,) + \\ + 2 \cdot K_\mu(N,n,r) \cdot \log_2 K_\mu(N,n,r) + 2 \cdot K_\mu(N,n,r), \end{aligned} \tag{3}$$

где $L_{\mu(N,n,r)}$ – это сокращенное количество проверяемых векторов возможных решений с.л.у. (2) с ограничениями $\{r \le n \mid \forall i = \overline{r+1, n},\ \mu_i^{(\upsilon)} = 0\}$, а $K_\mu(N,n,r)$ – число неотрицательных целочисленных решений (2) с ограничениями *r*. Основную трудоёмкость выражения (3) определяет первый член многочлена [2]

$$L_{\mu(N,n,r)} = (N+1)^{min([n/N],r)+1} \cdot \frac{(min([n/N],r))!}{(\, n+min([n/N],r)\,)!} \cdot \frac{(n+r)!}{r!}. \tag{4}$$

Из формул (3) и (4) следует, что алгоритмическая сложность расчета точных приближений $C_{MBK}(P_\Delta\{S_{N,n} \ge c\})$ – это многочлен, зависящий не только от параметров выборки *N*, *n*, но и от параметра ограничения *r*, который, в свою очередь, функционально зависит от параметров выборки и точности вычислений Δ, то есть *r*=*m*(*N*,*n*,Δ).

Наиболее трудоемкая часть метода МВК – это процедура получения векторов потенциальных решений и проверки их принадлежности множеству решений. Для практических задач мощность алфавита N находится в диапазоне от 128 до 256, а объёмы выборки – от 320 до 1280 при требуемом общем времени вычислений не более 30 дней. Поэтому для оценки алгоритмической сложности в [2] были определены заданные в формате *(мощность алфавита, объём выборки)* следующие выборки: (256, 1280), (128, 640), (128, 320) и (192, 320) с точностью вычислений $\Delta=10^{-5}$. Алгоритмическая сложность и требуемая производительность задачи для этих параметров выборок приведены в табл. 1.

Таблица 1

Характеристики метода расчета точных приближений для различных параметров выборок

№ п/п	Параметры выборки N/n	Параметр ограничения объема выборки r	Число проверок сокращенное $L_{\mu(N,n,r)}$	Число решений $K_\mu(N,n,r)$	Общая сложность $C_{MBK}(P_\Delta\{S_{N,n}\geq c\}$ (операций)	Необходимая производит. для расчета за 30 дней (оп/сек)
1	256 / 1280	23	$1{,}39\times10^{50}$	$5{,}60\times10^{25}$	$1{,}60\times10^{52}$	$6{,}15\times10^{45}$
2	128 / 640	22	$2{,}67\times10^{27}$	$1{,}76\times10^{20}$	$2{,}94\times10^{29}$	$1{,}13\times10^{23}$
3	192 / 320	14	$6{,}50\times10^{23}$	$4{,}67\times10^{12}$	$4{,}55\times10^{25}$	$1{,}75\times10^{19}$
4	128 / 320	16	$1{,}21\times10^{21}$	$2{,}23\times10^{13}$	$9{,}68\times10^{22}$	$3{,}72\times10^{16}$

Из табл. 1 следует, что вычислительная сложность расчета варьируется от $9{,}68\times10^{22}$ до $1{,}60\times10^{52}$ операций, средняя – порядка $4{,}55\times10^{25}$ операций, при этом нужно проверить от $6{,}50\times10^{23}$ до $1{,}39\times10^{50}$ векторов и получить от $4{,}67\times10^{12}$ до $5{,}60\times10^{25}$ решений.

Количество проверяемых переменных в (2) не превышает (r+1), т.е. для приведенных в табл. 1 параметров не превышает 24: $\mu^{(\upsilon)}=\{\mu_0^{(\upsilon)},\mu_1^{(\upsilon)},\dots,\mu_{23}^{(\upsilon)}\}$, а остальные переменные равны нулю: $\{\mu_j^{(\upsilon)}=0\mid j=24,,\dots,n\}$. Значения $\mu_i^{(\upsilon)}$ целые и лежат в пределах $\{0\leq\mu_i^{(\upsilon)}\leq 23\mid i=\overline{0,23}\}$. Для размещения одного $\mu_i^{(\upsilon)}$ требуется не менее 5 бит ($2^4<23<2^5$), а для размещения всего вектора $\mu^{(\upsilon)}$, состоящего из 24-х компонентов $\mu_i^{(\upsilon)}$, требуется не менее $5\times24=120$ бит.

Важным свойством метода является возможность его распараллеливания по данным, так как проверка любых $\mu^{(i)}$ и $\mu^{(j)}$ на принадлежность к решениям системы при $i\neq j$ может проводиться независимо или параллельно.

Рассмотрим возможность применения перспективных технологий - квантовых и фотонных вычислительных систем – для решения задачи расчета точных приближений распределений.

Применение квантовых вычислительных технологий для расчета точных приближений распределений. Квантовые вычислительные технологии были предложены в 80-х годах 20 века целым рядом известных ученых, таких как Ричард Ф. Фейнман [6], Пол Бениофф [7], К.А. Валиев и А.А. Кокин [8], Ю.И. Манин [9]. Основная идея квантовых вычислений состоит в том, что квантовая система из q кубитов, работающих когерентно, имеет 2^q линейно независимых состояний. Согласно принципу квантовой суперпозиции, пространство состояний такого квантового регистра является 2^q-мерным гильбертовым пространством [10]. Одна операция над группой из q кубитов вычисляется сразу над всеми возможными её значениями, в отличие от группы классических битов, когда существует только одно текущее значение. Так обеспечивается максимально возможное распараллеливание вычислений по данным и рост производительности до 2^q, которое называют «квантовым ускорением». Физическими системами, реализующими кубиты, могут быть любые объекты, имеющие два квантовых состояния: поляризационные состояния фотонов, электронные состояния изолированных атомов или ионов, спиновые состояния ядер атомов, и т.д.

Структурная схема квантового компьютера, предложенная русскими учеными К.А. Валиевым и А.А. Кокиным [8], представлена на рис. 1.

Р. Фейнман сформулировал принцип [6], что для любого алгоритма можно получить реализацию на квантовой системе, которая будет не хуже его реализации на классической фон-неймановской системе. Однако одновременно показано [9], что «квантовое ускорение» возможно не для всякого алгоритма и для произвольного алгоритма возможность получения квантового ускорения не гарантируется. Особенностью квантовых вычислений является вероятностный характер результата вычислений: получаемый результат верен только с некоторой вероятностью.

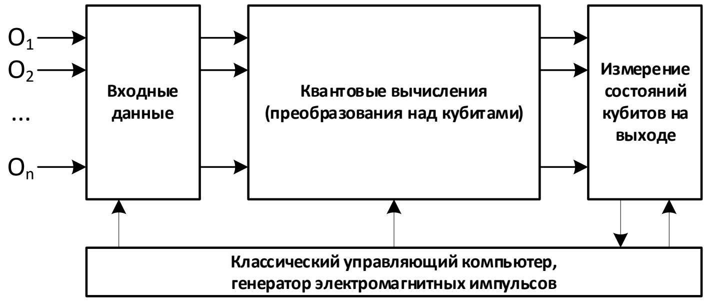

Рис. 1. Структурная схема квантового компьютера

Квантовые вычислители созданы в Гарвардском университете (система из 51 кубита) [11] и в Политехнической школе Университета Париж-Сакле (система из 70 кубит) [12]. Американская компания IonQ [13] анонсировала первый коммерческий квантовый компьютер на ионных ловушках, который содержит 160 кубит, но для квантовых операций используются 79 кубитов, а произвольные квантовые алгоритмы реализуются всего на 11 кубитах. В 2020 году фирма IBM [14] представила самый мощный квантовый компьютер на 64 кубитах. В России создание универсального квантового вычислителя из 100 кубитов ожидается к 2024 году [12], а квантового вычислителя, способного оперировать с данными в 120 бит, в настоящее время в мире не существует.

Оценим возможность решения задачи расчета точных приближений распределений на квантовом компьютере. Структурная схема процедуры проверки векторов в задаче расчета точных приближений распределений [2] соответствует структурной схеме квантового компьютера (рис. 1), что позволяет ожидать значительного эффекта при решении задачи на квантовых компьютерах. По минимальной оценке, для размещения вектора проверки решений $\mu^{(\upsilon)}$ требуется 120 бит, поэтому в квантовом компьютере нужно иметь 120 кубит, работающих совместно (когерентно). Предположим, что квантовый вычислитель, оперирующий с 120 кубитами (назовем его ГКВ-120), существует, и для него решена проблема достоверности результатов. При расчете точных приближений на ГКВ-120, для вектора проверок $\mu^{(\upsilon)} = \{\mu_0^{(\upsilon)}, \mu_1^{(\upsilon)}, \ldots, \mu_{23}^{(\upsilon)}\}$ длиной 120 бит, получим все $\{\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(2^{120})}\}$, т.е. $(2^{120} \cong 1{,}3292 \times 10^{36})$ возможных решений, из которых нужно выделить $K_\mu(N, n, r)$ истинных решений. Для N=256 и n=1280 из таблицы 1 $K_\mu(256,1280,23)=5{,}60075\cdot10^{25}$. Таким образом, при измерении значения i-го возможного решения получаем состояние системы $\{\underbrace{\mu^{(i)}, \mu^{(i)}, \dots, \mu^{(i)}}_{2^{120}}\}$. Для получения ($i$+1)-го возможного решения получаем состояние системы $\{\mu^{(1)}, \mu^{(2)}, \dots, \mu^{(2^{120})}\}$ и считываем (измеряем) (i+1)-е возможное решение. Для проверки всех возможных решений и получения $K_\mu(N, n, r)$ истинных решений необходимо выполнить 2^{120} обращений к ГКВ-120, что кардинально сокращает эффект от одновременного расчета 2^{120} возможных решений.

Необходимость проверки всех полученных решений, число которых соответствует размерности задачи, не позволяет достичь квантового ускорения и является существенным и принципиальным ограничением применения перспективных квантовых вычислительных технологий для расчета точных приближений распределений. Отсутствие технической и технологической базы не только в Российской Федерации, но и в мире, для создания квантовой вычислительной системы, оперирующей 120 кубитами, необходимыми для решения задачи, является дополнительным, технологическим ограничением. Поэтому перспективы применения квантовых вычислительных систем на существующем уровне развития науки и техники для решения задачи расчета точных приближений распределений вероятностей значений статистик достаточно скромны, несмотря на потенциально высокую производительность. Возможно, с развитием работ в области квантовых вычислений, приведенные оценки для рассматриваемой задачи могут быть пересмотрены.

Применение фотонных компьютеров для расчета точных приближений распределений. Еще одним актуальным направлением развития перспективных вычислительных средств, основанных на новых физических принципах, является создание вычислительных систем, использующих эффекты взаимодействия когерентных световых волн, порождаемых лазерным излучением и его носителями - фотонами [15, 16].

Структурная схема фотонного компьютера, разработанная во Всероссийском научно-исследовательском институте экспериментальной физики (г. Саров) [15], представлена на рис. 2. Фотонный компьютер (рис. 2) состоит из четырех крупных блоков: устройства преобразования задачи в программу для фотонного вычислителя (УПЗ), источника лазерного излучения (ИЛИ), устройства ввода-вывода (УВВ), а также фотонного процессора (ФП). В свою очередь, фотонный процессор содержит четыре процессорных элемента, которые содержат арифметико-логические устройства (АЛУ), устройства управления (УУ) и коммутаторы (К). Блоки фотонного компьютера соединены между собой электронными и оптическими каналами, а компоненты процессорного элемента – только оптическими каналами.

Блок УПЗ передает фотонную программу в блок ИЛИ. Лазерное излучение генерируется блоком ИЛИ и поступает в УВВ, где делится на световые лучи по количеству разрядов, одновременно подаваемым в ПЭ. Модуль УВВ формирует фотонную программу и передает ее в фотонный процессор, в котором вычисления выполняются процессорными элементами. Световые лучи взаимодействуют в пределах фотонного процессора, а синхронизацию выполняют оптические линии задержки [16]. В пределах фотонного процессора ПЭ можно соединить оптическими каналами в многопроцессорную среду любой топологии [16]. Одним из немаловажных достоинств фотонных вычислительных систем является низкое энергопотребление при высокой вычислительной производительности.

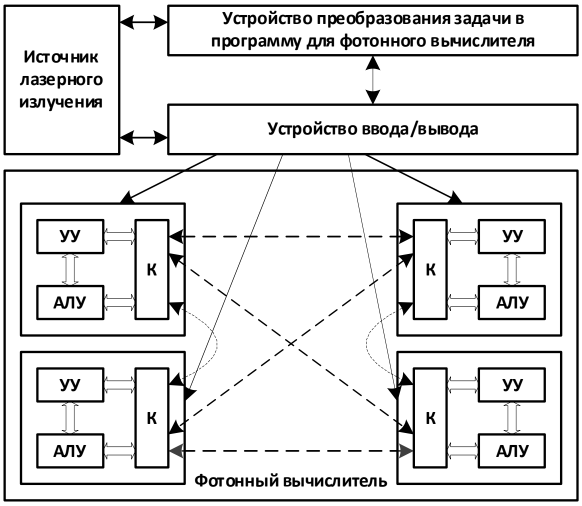

Рис. 2. Структурная схема фотонного компьютера

Показано, что производительность фотонного вычислительного узла может достигать до 10^{18} оп./сек. на 100 Вт потребляемой мощности [17].

Для достижения требуемой производительности фотонного компьютера P_{PHOTON}, нужно использовать $N_{PH\_УЗ}$ узлов производительностью $P_{PH\_УЗ}$=10^{18} оп/сек каждый:

$$P_{PHOTON} = N_{PH\_УЗ} \times P_{PH\_УЗ}. \qquad (5)$$

Уровень производительности $P_{PHOTON} \geq 1{,}75 \times 10^{19}$ для расчетов точных приближений (параметры выборки № 3 в табл. 1) обеспечат

$$N_{PH\_УЗ} \geq 1{,}75 \times 10^{19} / P_{PH\_УЗ} = 1{,}75 \times 10^{19} / 1{,}0 \times 10^{18} \cong 17{,}5$$

узлов, то есть не менее 18 фотонных вычислительных узлов. Для достижения производительности, позволяющей вычислять точные приближения для всего спектра параметров рассматриваемых выборок $\{N = \overline{2,\, 256},\, n = \overline{1,\, 5N}\}$, необходимо иметь не менее

$$N_{PH\_УЗ} \geq 6{,}15 \times 10^{45} / P_{PH\_УЗ} = 6{,}15 \times 10^{45} / 1{,}0 \times 10^{18} = 6{,}15 \times 10^{27}$$

вычислительных узлов, что существенно ниже, чем во всех рассмотренных ранее [2] технологиях.

В отличие от квантовых вычислительных систем, данных о действующих образцах цифровых фотонных вычислительных систем на настоящий момент не приводится. В большинстве работ [15–17] моделируется функционирование отдельных узлов и оцениваются возможные параметры вычислителя. Известны работы академика В.А. Сойфера [18–20] и ряда других исследователей в области аналоговой фотоники, но каждый вычислительный узел создается для задачи с определенными параметрами, а технология разработки и создания аналогового фотонного вычислительного узла для произвольной задачи требует целого цикла научно-исследовательских и опытно-конструкторских работ. Поэтому действующих образцов фотонных вычислительных систем, технологической базы и/или коммерчески доступных компонентов на сегодняшний день не существует. Текущий уровень развития науки и техники не позволяет в ближайшие 5–7 лет рассматривать применение фотонных вычислительных систем для решения задачи расчета точных приближений распределений вероятностей значений статистик, хотя потенциально такие вычислительные системы обладают очень высокой производительностью и требуют наименьшее число вычислительных узлов с низким энергопотреблением. Возможно, развитие технологической элементной базы для фотонных вычислительных технологий позволит пересмотреть и улучшить представленные оценки для рассматриваемой задачи в ближайшие несколько лет.

Архитектурно-независимое представление задачи расчета точных приближений для гибридных вычислительных систем. Одним из перспективных направлений расчета точных приближений является создание специализированной гибридной вычислительной системы (ГВС) с конвергентной архитектурой, которая может содержать вычислительные узлы современных архитектур: универсальные процессоры, графические ускорители, программируемые логические интегральные схемы. Поскольку вычисление распределений производится методом, единым для всех параметров выборок, то для такой системы особенно важна возможность программирования различных вычислительных архитектур в едином контуре, которая обеспечивается в архитектурно-независимой парадигме программирования [21]. С учетом возможного использования перспективных архитектур квантовых вычислителей и/или фотонных вычислительных узлов, архитектурно-независимое представление задачи расчета точных приближений для гибридных вычислительных систем становится особо значимым.

Такие возможности обеспечивает архитектурно-независимый аспектно-ориентированный язык SET@L, позволяющий сосредоточить внимание разработчика на методах параллелизма при решении задачи, а не на их технической реализации в выбранной вычислительной архитектуре. Язык Set@l позволяет свести проблему переноса программного обеспечения между разными по конфигурации и архитектуре ГВС к формированию аспектов (описаний) технических средств, в которых описываются ключевые моменты разбиения метода по параллелизму на данных технических средствах. При этом исходная программа, реализующая метод обработки данных, остается неизменной [22].

Язык Set@l основан на парадигме аспектно-ориентированного программирования (АОП) [23], согласно которой алгоритм решения прикладной задачи и особенности его реализации описываются в виде отдельных модулей программы. Программа на языке Set@l представляет информационный граф решаемой вычислительной задачи в виде множеств, разбиение и типизация которых задают разные варианты распараллеливания и другие аспекты реализации алгоритма. В отличие от других языков программирования на основе теории множеств, например, языков SETL, SETL2 и SETLX, Set@l использует типизацию множеств по различным критериям и позволяет оперировать нечеткими совокупностями в соответствии с альтернативной теорией множеств [24].

Разделение описаний алгоритма и особенностей его реализации на ВС в языке Set@l обеспечивается использованием парадигмы АОП, согласно которой сквозная функциональность программы, обуславливающая негативные эффекты спутывания и разбрасывания кода, представляется в виде отдельных программных модулей – аспектов. Исходный код, реализующий основную функциональность программы, содержит создаваемую пользователем разметку, которая определяет его взаимодействие с аспектами при трансляции или исполнении. Анализируя разметку и исходный код, транслятор-препроцессор формирует новую исполняемую программу с вплетенной сквозной функциональностью. Применение технологии АОП, как правило, упрощает разработку и дальнейшую поддержку ПО и повышает адаптируемость программ к различным изменениям.

Для неявного описания различных способов распараллеливания алгоритмов в языке архитектурно-независимого программирования ВС Set@l введена классификация совокупностей по типу параллелизма их элементов при обработке. В тех случаях, когда характер параллелизма элементов совокупности четко определен, используются типы «множество» (par – параллельно-независимая обработка), «кортеж» (seq – последовательная обработка), «кортеж-конвейер» (pipe – конвейерная обработка) и «множество обработки по итерациям» (conc – параллельно-зависимая обработка). Однако в некоторых аспектах программы тип параллелизма ряда совокупностей нельзя определить однозначно, так как неизвестна архитектура ВС, на которой будет реализован алгоритм. В данном случае используется особый тип imp (неопределенный), а типизация совокупностей уточняется в других аспектах программы с помощью специальных синтаксических конструкций. Если аспекты программы на языке Set@l не изменяют алгоритм в процессе его адаптации к архитектуре ВС, то решение задачи может быть представлено в рамках классической теории множеств Кантора–Больцано. Однако в ряде случаев целесообразно модернизировать алгоритм в соответствии с особенностями его реализации на ВС с определенной архитектурой. В таких случаях некоторые совокупности выделены нечетко и не являются множествами, поэтому их невозможно задать как объекты классической теории множеств. Set@l позволяет описать разные способы реализации одного и того же алгоритма в единой аспектно-ориентированной программе. Для этого вводится классификация совокупностей по четкости выделения их элементов, предложенная в рамках альтернативной теории множеств П. Вопенка. Тип «множество» (set) соответствует классической четко выделенной совокупности элементов, тип «полумножество» (sm) – нечеткой совокупности, а тип «класс» (cls) – совокупности объектов, тип и разбиение которой нельзя однозначно определить на данном уровне абстракции. Примером вычислительной задачи, для описания которой на языке программирования Set@l необходимы нечеткие совокупности, является алгоритм решения СЛАУ методом Якоби [25].

В отличие от других подходов к параллельному программированию высокопроизводительных ВС, язык программирования Set@l задает не только граничные случаи реализации алгоритма, но и семейство промежуточных вариантов, которые неразличимы с точки зрения процедурного программирования, но обеспечивают непрерывность описания модели вычислений. Использование языка Set@l позволяет синтезировать множество вариантов решения задачи и осуществлять переход между его элементами в зависимости от архитектуры и конфигурации ВС. Указанная возможность проиллюстрирована на примере программы быстрого преобразования Фурье: в зависимости от доступного объема памяти ВС комплексные поворотные коэффициенты вычисляются заранее и загружаются из памяти или рассчитываются через базовые и вспомогательные составляющие в процессе вычислений [26].

Таким образом, алгоритм решения задачи в языке Set@l представляется в виде архитектурно-независимого исходного кода с помощью аспектно-ориентированного программирования, теоретико-множественного представления кода и реляционного исчисления. Особенности реализации алгоритма выделяются в отдельные аспекты, задающие разбиение на подмножества и типизацию основных совокупностей программы. Программа может оперативно портироваться между вычислительными системами с разными архитектурами и адаптируется к любым изменениям аппаратных ресурсов с помощью аспектов метода обработки, архитектуры и конфигурации. Благодаря предлагаемому подходу к параллельному программированию высокопроизводительных вычислительных систем на языке Set@l, открываются новые перспективы для разработки архитектурно- и ресурсонезависимого программного обеспечения.

Язык Set@l успешно применялся для решения алгоритмически сложных и ресурсоемких задач, таких как решение систем линейных алгебраических уравнений методом Гаусса [22] и Якоби [25], реализации алгоритма быстрого преобразования Фурье [26] и других, которые структурно не отличаются от решаемой задачи. Поэтому возможность использования языка Set@l при расчете точных приближений распределений на ГВС позволит существенно облегчить сложность программирования вычислительных узлов различных архитектур: универсальных процессоров, графических ускорителей, программируемых логических интегральных схем, квантовых вычислителей и/или фотонных вычислительных узлов.

Заключение и выводы. В статье проанализирована возможность использования перспективных вычислительных технологий для решения вычислительно трудоемкой задачи расчета Δ-точных приближений распределений вероятностей значений статистик методом второй кратности, основанным на решении системы линейных уравнений второй кратности. Метод имеет полиномиальную сложность и допускает распараллеливание на фрагменты по данным.

Оценены возможности перспективных вычислительных технологий на основе квантовых и фотонных компьютеров для расчета распределений методом второй кратности. Анализ возможностей квантовых и фотонных вычислительных технологий показал, что они обладают большим потенциалом для решения задачи, но в настоящее время не позволяют решать задачу вычисления точных приближений распределений выборок с мощностью алфавита не более 256, объемом до 1280 знаков и ограничениями точности вычислений $\Delta=10^{-5}$. Проведен теоретический анализ квантовых вычислительных систем для решения задачи, который показал, что квантовое ускорение невозможно. Так, текущий уровень развития квантовых компьютеров является недостаточным, а также не существует алгоритма и критерия выбора подходящего решения из огромного множества получаемых решений задачи, а уровень развития фотонных компьютеров не позволяет создать вычислитель с требуемым числом вычислительных узлов.

Поддержка. Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-07-00545.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Мельников А.К.* Сложность расчета точных распределений вероятности симметричных аддитивно разделяемых статистик и область применения предельных распределений // Доклады ТУСУР. – Томск, 2017. – Т. 20, № 4, – С. 126-130. – ISSN 1818-0442.
2. *Мельников А.К., Левин И.И., Дордопуло А.И., Писаренко И.В.* Анализ возможностей современных вычислительных технологий для расчета точных приближений распределений вероятностей значений статистик // Известия ЮФУ. Технические науки. – 2021. – № 7 (224). – С. 6-19. – ISSN 1999-9429. – DOI: 10.18522/2311-3103-2021-7-6-19.
3. *Тимонин В.И., Черномордик О.М.* Метод вычисления точного распределения статистик типа Колмогорова-Смирнова при альтернативах Лемана // Теория вероятностей и ее применение. – 1985. – Т. 30. – Вып. 3. – С. 572-573.
4. *Тянникова Н.Д, Тимонин В.И.* Метод вычисления точных распределений статистик типа Колмогорова-Смирнова в случае нарушения однородности и независимости анализируемых выборок // Наука и образование. – 2014. – № 11. – С. 227–237. – DOI: 10.7463/1114.0740251 ISSN 1994-0448.
5. Расчет распределений методом Монте-Карло.
6. *Feynman R.P.* Simulating physics with computers (англ.) // International Journal of Theoretical Physics. – 1982. – Vol. 21, Iss. 6. – P. 467-488. – DOI: 10.1007/BF02650179.
7. *Benioff P.* Quantum mechanical hamiltonian models of turing machines (англ.) // Journal of Statistical Physics (англ.) русск.: journal. – 1982. – Vol. 29, No. 3. – P. 515-546. – DOI: 10.1007/BF01342185. – Bibcode: 1982JSP....29..515B.
8. *Валиев К.А., Кокин А.А.* Квантовые компьютеры: надежды и реальность. – Ижевск: РХД, 2004. – 320 с.

9. *Манин Ю.И.* Вычислимое и невычислимое. – М.: Сов. радио, 1980. – 128 с. – (Кибернетика).
10. *Морен К.* Методы гильбертова пространства. – М.: Мир, 1965. – 570 с.
11. Физики из России и США создали первый 51-кубитный квантовый компьютер. – https://ria.ru/20170714/1498476410.html? (дата обращения: 2 мая 2022).
12. Как Россия потратит 15 млрд. на создание квантового компьютера. – https://gov.cnews.ru/articles/2020-01-9_kak_rossiya_potratit_15_mlrd_na_sozdanie? (дата обращения: 2 мая 2022).
13. IonQ заявила о создании самого мощного в мире квантового компьютера. – https://3dnews.ru/1060908/kompaniya-ion-q-soobshchila-o-sozdanii-samogo-moshchnogo-v-mire-kvantovogo-kompyutera? (дата обращения: 2 мая 2022).
14. IBM продемонстрировали квантовый компьютер с «квантовым объемом» 64. – https://habr.com/ru/company/raiffeisenbank/news/t/516062/? (дата обращения: 2 мая 2022).
15. *Степаненко С.А.* Фотонная вычислительная машина. Принципы реализации. Оценка параметров // Вычислительная техника. Доклады Академии наук. – М., 2017. – Т. 476, № 4. – С. 389-394. – DOI: 10.1134/S1064562417050234.
16. *Степаненко С.А.* Фотонный компьютер: структура и алгоритмы, оценка параметров // Фотоника. – 2017. – № 7 / 67. – С. 72-83. – DOI: 10.22184/1993-7296.2017.67.7.72.83.
17. *Степаненко С.А*. Интерференционные логические элементы / представлено академиком РАН Ю.А. Трутневым. Доклады Академии наук. Математика, информатика, процессы управления. – М., 2020. – Т. 493. – С. 64-69. – DOI: 10.31857/S2686954320040189.
18. *Soifer V.A., Kharitonov S.I., Khonina S.N. etc*. Spiral caustics of vortex beams // Photonics. – 2021. – Vol. 8, Issue 1. – P. 1-20. – DOI: 10.3390/photonics8010024.
19. *Soifer V.A., Doskolovich L.L., Bezus E.A. etc*. Spatial differentiation of optical beams using a resonant metal-dielectric-metal structure // Journal of Optics. – 2021. – Vol. 23, Issue 2. – DOI: 10.1088/2040-8986/abe63b.
20. *Soifer V.A., Doskolovich L.L., Bezus E.A. etc.* An Optical Differentiator Based on a Three-Layer Structure with a W-Shaped Refractive Index Profile // Journal of Experimental and Theoretical Physics. – 2018. – Vol. 127, Issue 2. – P. 202-209. – DOI: 10.1134/S1063776118080174.
21. *Левин И.И., Мельников А.К.* Управление гибридными вычислительными системами на языке SET@L // Materiály XI mezinárodní vědecko-praktická konference «Aktuální vymoženosti vědy – 2015» Díl 7. Moderní informační technologie. Praha. Publishing House «Education and Science», 2015. – 96 с. – P. 23-28. – ISSN 978-966-8736-05-6.
22. *Левин И.И., Дордопуло А.И., Писаренко И.В., Мельников А.К.* Подход к архитектурно-независимому программированию вычислительных систем на основе аспектно-ориентированного языка Set@l // Известия ЮФУ. Технические науки. – 2018. – № 3. – С. 46-58.
23. *Safonov V.O.* Using Aspect-Oriented Programming for Trustworthy Software Development. – New York: John Wiley & Sons, 2008. – 352 p.
24. *Вопенка П.* Альтернативная теория множеств: Новый взгляд на бесконечность: пер. со славац. – Новосибирск: Изд-во Института математики, 2004. – 612 с.
25. *Левин И.И., Дордопуло А.И., Писаренко И.В., Мельников А.К.* Описание алгоритма решения систем линейных алгебраических уравнений методом Якоби на языке архитектурно-независимого программирования Set@l // Известия ЮФУ. Технические науки. – 2018. – № 5. – С. 34-48.
26. *Левин И.И., Дордопуло А.И., Писаренко И.В., Мельников А.К.* Архитектурно-независимая программа быстрого преобразования Фурье на языке программирования SET@L // Вестник РГРТУ. – 2019. – № 68. – С. 28-36. – ISSN 1995-4565. – DOI: 10.21667/1995-4565-2019-68-2-28-36. – http://vestnik.rsreu.ru/ru/archive/2019/vypusk-68/805-1995-4565-2019-68-2-28-36.

REFERENCES

1. *Mel'nikov A.K.* Slozhnost' rascheta tochnykh raspredeleniy veroyatnosti simmetrichnykh additivno razdelyaemykh statistik i oblast' primeneniya predel'nykh raspredeleniy [The complexity of calculating the exact probability distributions of symmetric additively separable statistics and the scope of marginal distributions], *Doklady TUSUR* [TUSUR reports]. Tomsk, 2017, Vol. 20, No. 4, pp. 126-130. ISSN 1818-0442.

2. *Mel'nikov A.K., Levin I.I., Dordopulo A.I., Pisarenko I.V.* Analiz vozmozhnostey sovremennykh vychislitel'nykh tekhnologiy dlya rascheta tochnykh priblizheniy raspredeleniy veroyatnostey znacheniy statistik [Analysis of the possibilities of modern computing technologies for calculating accurate approximations of probability distributions of statistical values], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2021, No. 7 (224), pp. 6-19. ISSN 1999-9429. DOI: 10.18522/2311-3103-2021-7-6-19.
3. *Timonin V.I., Chernomordik O.M.* Metod vychisleniya tochnogo raspredeleniya statistik tipa Kolmogorova-Smirnova pri al'ternativakh Lemana [A method for calculating the exact distribution of Kolmogorov-Smirnov type statistics with Lehman alternatives], *Teoriya veroyatnostey i ee primenenie* [Probability theory and its application], 1985, Vol. 30, Issue 3, pp. 572-573.
4. *Tyannikova N.D, Timonin V.I.* Metod vychisleniya tochnykh raspredeleniy statistik tipa Kolmogorova-Smirnova v sluchae narusheniya odnorodnosti i nezavisimosti analiziruemykh vyborok [Method for calculating exact distributions of Kolmogorov-Smirnov type statistics in case of violation of uniformity and independence of the analyzed samples], *Nauka i obrazovanie* [Science and Education], 2014, No. 11, pp. 227–237. DOI: 10.7463/1114.0740251 ISSN 1994-0448.
5. Raschet raspredeleniy metodom Monte-Karlo [Calculation of distributions by the Monte Carlo method].
6. *Feynman R.P.* Simulating physics with computers, *International Journal of Theoretical Physics*, 1982, Vol. 21, Iss. 6, pp. 467-488. DOI: 10.1007/BF02650179.
7. *Benioff P.* Quantum mechanical hamiltonian models of turing machines, *Journal of Statistical Physics*, 1982, Vol. 29, No. 3, pp. 515-546. DOI: 10.1007/BF01342185. Bibcode: 1982JSP....29..515B.
8. *Valiev K.A., Kokin A.A.* Kvantovye komp'yutery: nadezhdy i real'nost' [Quantum computers: Hopes and Reality]. Izhevsk: RKHD, 2004, 320 p.
9. *Manin Yu.I.* Vychislimoe i nevychislimoe [Computable and non–computable]. Moscow: Sov. radio, 1980, 128 p. (Cybernetics).
10. *Moren K.* Metody gil'bertova prostranstva [Methods of Hilbert space]. Moscow: Mir, 1965, 570 p.
11. Fiziki iz Rossii i SSHA sozdali pervyy 51-kubitnyy kvantovyy komp'yuter [Physicists from Russia and the USA have created the first 51-qubit quantum computer]. Available at: https://ria.ru/20170714/1498476410.html? (accessed 2 May 2022).
12. Kak Rossiya potratit 15 mlrd. na sozdanie kvantovogo komp'yutera [How will Russia spend 15 billion? to create a quantum computer]. Available at: https://gov.cnews.ru/articles/2020-01-9_kak_rossiya_potratit_15_mlrd_na_sozdanie? (accessed 2 May 2022).
13. IonQ zayavila o sozdanii samogo moshchnogo v mire kvantovogo komp'yutera [IonQ has announced the creation of the world's most powerful quantum computer]. Available at: https://3dnews.ru/1060908/kompaniya-ion-q-soobshchila-o-sozdanii-samogo-moshchnogo-v-mire-kvantovogo-kompyutera? (accessed 2 May 2022).
14. IBM prodemonstrirovali kvantovyy komp'yuter s «kvantovym ob"emom» 64. Available at: https://habr.com/ru/company/raiffeisenbank/news/t/516062/? (accessed 2 May 2022).
15. *Stepanenko S.A.* Fotonnaya vychislitel'naya mashina. Printsipy realizatsii. Otsenka parametrov [Photonic computing machine. Principles of implementation. Estimation of parameters], *Vychislitel'naya tekhnika. Doklady Akademii nauk* [Computer engineering. Reports of the Academy of Sciences]. Moscow, 2017, Vol. 476, No. 4, pp. 389-394. DOI: 10.1134/S1064562417050234.
16. *Stepanenko S.A.* Fotonnyy komp'yuter: struktura i algoritmy, otsenka parametrov [Photonic computer: structure and algorithms, estimation of parameters], *Fotonika* [Photonics], 2017, No. 7 / 67, pp. 72-83. DOI: 10.22184/1993-7296.2017.67.7.72.83.
17. *Stepanenko S.A.* Interferentsionnye logicheskie elementy [Interference logic elements], presented by Academician of the Russian Academy of Sciences Yu.A. Trutnev. *Doklady Akademii nauk. Matematika, informatika, protsessy upravleniya* [Reports of the Academy of Sciences. Mathematics, Computer Science, Management processes]. Moscow, 2020, Vol. 493, pp. 64-69. DOI: 10.31857/S2686954320040189.
18. *Soifer V.A., Kharitonov S.I., Khonina S.N. etc*. Spiral caustics of vortex beams, *Photonics*, 2021, Vol. 8, Issue 1, pp. 1-20. DOI: 10.3390/photonics8010024.
19. *Soifer V.A., Doskolovich L.L., Bezus E.A. etc*. Spatial differentiation of optical beams using a resonant metal-dielectric-metal structure, *Journal of Optics*, 2021, Vol. 23, Issue 2. DOI: 10.1088/2040-8986/abe63b.

20. *Soifer V.A., Doskolovich L.L., Bezus E.A. etc.* An Optical Differentiator Based on a Three-Layer Structure with a W-Shaped Refractive Index Profile, *Journal of Experimental and Theoretical Physics,* 2018, Vol. 127, Issue 2, pp. 202-209. DOI: 10.1134/S1063776118080174.
21. *Levin I.I., Mel'nikov A.K.* Upravlenie gibridnymi vychislitel'nymi sistemami na yazyke SET@L [Management of hybrid computing systems in the SET@L language], *Materiály XI mezinárodní vědecko-praktická konference «Aktuální vymoženosti vědy – 2015» Díl 7. Moderní informační technologie* [Materials XI international scientific and practical conference "current achievements of science-2015" Part 7. Modern information technology"]. Praha. Publishing House «Education and Science», 2015. – 96 s. – P. 23-28. – ISSN 978-966-8736-05-6.
22. *Levin I.I., Dordopulo A.I., Pisarenko I.V., Mel'nikov A.K.* Podkhod k arkhitekturno-nezavisimomu programmirovaniyu vychislitel'nykh sistem na osnove aspektno-orientirovannogo yazyka Set@l [Approach to architecture-independent programming of computing systems based on aspect-oriented language Set@l], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2018, No. 3, pp. 46-58.
23. *Safonov V.O.* Using Aspect-Oriented Programming for Trustworthy Software Development. New York: John Wiley & Sons, 2008, 352 p.
24. *Vopenka P.* Al'ternativnaya teoriya mnozhestv: Novyy vzglyad na beskonechnost' [Alternative Set Theory: A New View of Infinityъ: Transl. from Slavack. Novosibirsk: Izd-vo Instituta matematiki, 2004, 612 p.
25. *Levin I.I., Dordopulo A.I., Pisarenko I.V., Mel'nikov A.K.* Opisanie algoritma resheniya sistem lineynykh algebraicheskikh uravneniy metodom Yakobi na yazyke arkhitekturno-nezavisimogo programmirovaniya Set@l [Description of the algorithm for solving systems of linear algebraic equations by the Jacobi method in the language of architecture-independent programming Set@l], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2018, No. 5, pp. 34-48.
26. *Levin I.I., Dordopulo A.I., Pisarenko I.V., Mel'nikov A.K.* Arkhitekturno-nezavisimaya programma bystrogo preobrazovaniya Fur'e na yazyke programmirovaniya SET@L [An architecture-independent fast Fourier transform program in the SET@L programming language], *Vestnik RGRTU* [Vestnik of RSREU], 2019, No. 68, pp. 28-36. ISSN 1995-4565. DOI: 10.21667/1995-4565-2019-68-2-28-36. Available at: http://vestnik.rsreu.ru/ru/archive/2019/vypusk-68/805-1995-4565-2019-68-2-28-36.

Статью рекомендовал к опубликованию д.т.н. Э.В. Мельник.

Мельников Андрей Кимович – АО «Вычислительные решения»; e-mail: anta-mak@umail.ru; ak@comp-sol.ru; г. Москва, Россия; г.н.с.; к.т.н.; доцент ВАК.

Левин Илья Израилевич – Южный федеральный университет; e-mail: iilevin@sfedu.ru; г. Таганрог, Россия; зав. кафедрой ИМС; д.т.н.; профессор.

Дордопуло Алексей Игоревич – НИЦ супер-ЭВМ и нейрокомпьютеров; e-mail: dordopulo@superevm.ru; г. Таганрог, Россия; начальник отдела; к.т.н.

Сластен Любовь Михайловна – e-mail: lmslasten@yandex.ru; научный сотрудник; к.т.н.

Melnikov Andrey Kimovich – JSC "Computing Solutions"; e-mail: anta-mak@umail.ru; ak@comp-sol.ru; Moscow, Russia; chief researcher; cand. of eng. sc.; associate professor of the Higher Attestation Commission.

Levin Il'ya Izrailevich – Southern Federal University; e-mail: iilevin@sfedu.ru; Taganrog, Russia; head of department of intelligent and multiprocessor systems; dr. of eng. sc.; professor.

Dordopulo Alexey Igorevich – Supercomputers and Neurocomputers Research Center; e-mail: dordopulo@superevm.ru; Taganrog, Russia; head of department; cand. of eng. sc.

Slasten Lyubov' Mikhaylovna – e-mail: lmslasten@yandex.ru; cand. of eng. sc.