

Раздел I. Принципы построения и архитектура суперкомпьютеров

УДК 004.382.2

DOI 10.18522/2311-3103-2016-12-616

А.К. Мельников

РАЗРАБОТКА ТРЕБОВАНИЙ К СРЕДСТВАМ АВТОМАТИЗАЦИИ ПРОЦЕДУРЫ АНАЛИЗА СООБЩЕНИЙ

Рассматривается разработанная автором процедура анализа сообщений, позволяющая выделить сообщения, содержащие смысловую нагрузку. Указанная процедура построена на основе обобщенного статистического метода анализа (ОСМА), позволяющего с помощью применения статистических критериев согласия с равновероятным распределением отсеивать из потока сообщений сообщения, имеющие распределение частот встречаемости входящих в него символов близкое к равновероятному распределению. Под эффективностью применения ОСМА понимается минимальность количества сообщений с равновероятным распределением символов, принятых за сообщения с не равновероятным распределением символов. Основой эффективного применения ОСМА и процедуры анализа сообщений в целом является использование в применяемом статистическом критерии согласия распределения вероятности значений статистики, соответствующего параметрам анализируемого сообщения. Для сохранения наибольшей эффективности процедуры анализа сообщений в ОСМА используются различные виды распределений вероятности значений статистики: точные, Δ -точные и предельные, позволяющие обеспечить соответствие параметров распределения вероятности значений статистики параметрам анализируемого сообщения, приведены методы расчета соответствующих видов распределений вероятности значений статистик, исследованы вычислительная сложность этих методов и значения границ параметров, в которых они применимы. На основе исследования алгоритмических свойств ОСМА сформулированы требования к языку программирования для получения программной реализации процедуры анализа сообщений и требования к вычислительным средствам для проведения практического анализа сообщений с помощью полученной программной реализации. Рассмотрение разработанных требований показало возможность использования в качестве языка программирования для получения программной реализации процедуры анализа сообщений языка описания множеств SET@L. Проведено сравнение характеристик языка программирования SET@L с характеристиками языков программирования Cryptol и COLAMO на предмет соответствия требованиям, предъявляемым алгоритмическими особенностями ОСМА. Проведен анализ понятий и программных конструкций языка SET@L, позволяющих обеспечить выполнение сформулированных требований. На основе исследования алгоритмических особенностей и вычислительной сложности процедуры анализа сообщений, сделан вывод о необходимости использования для проведения практического анализа сообщений гибридных высокопроизводительных многопроцессорных вычислительных систем на основе ПЛИС.

Статистический критерий; распределение вероятности; процедура анализа; программная реализация; язык программирования; вычислительные системы; программируемые логические интегральные схемы; производительность; вычислительная сложность.

A.K. Melnikov

DEVELOPMENT OF REQUIREMENTS TO AUTOMATIZATION TOOLS OF MESSAGE ANALYSIS PROCEDURE

In the paper we consider the procedure, which we developed for message analysis and which provides detection of messages with some meaning. The procedure is based on a generalized statistical method of analysis (GSMA), which with the help of statistical fitting criteria with equiprobable distribution provides selection of messages from the messages flow that have distribution of frequencies of their symbols occurrence close to equiprobable distribution. In our opinion, effectiveness of the GSMA application is the minimum number of messages with equiprobable distribution of symbols, which are taken for messages with non-equiprobable distribution of symbols. In total, the base of effective application of the GSMA and the procedure of message analysis is the use of statistics probability distribution corresponding to the parameters of the analyzed message in the used statistical fitting criterion. To keep the highest effectiveness of the message analysis procedure, various kinds of statistics probability distribution, such as exact, Δ -exact and limit distributions which provide matching between the parameters of statistics probability distribution and the parameters of the analyzed message, are used in the GSMA. Besides, we present methods for calculation of the corresponding kinds of statistics probability distribution, and analyze computational complexity of these methods and boundary parameter values these methods can be used for. On the base of analysis of algorithmic properties of a generalized statistical method of message analysis we have developed requirements to a programming language for software implementation of the message analysis procedure, and requirements to computer facilities for message practical analysis using the developed software implementation of the procedure. Analysis of the developed requirements proved that it is possible to use the set description language SET@L as a programming language for software implementation of the message analysis procedure. Also, we have compared the characteristics of the programming language SET@L with the characteristics of the programming languages Cryptol and COLAMO to check their correspondence to the requirements of the algorithmic features of the GSMA. We have analyzed definitions and program structures of the language SET@L, owing to which the declared requirements can be fulfilled. According to the study of algorithmic features and computational complexity of the message analysis procedure we can conclude that it is necessary to use high-performance FPGA-based hybrid computer systems for practical message analysis.

Statistical criterion; probability distribution; procedure of analysis; software implementation; programming language; computer systems; FPGA; performance; computational complexity.

Введение. При постоянно растущем объеме сообщений в глобальном информационном пространстве актуальной является задача выделения сообщений содержащих смысловую нагрузку. Процедура анализа сообщений, позволяющая выделить сообщения содержащие смысловую нагрузку, основана на разработанном в [1] обобщенном статистическом методе анализа (ОСМА) сообщений, позволяющем с помощью применения статистических критериев согласия (СКС) [2] с равновероятным распределением отсеивать из потока сообщений сообщения, имеющие близкое к равновероятному распределению входящих в него символов, т.е. не содержащие смысловой нагрузки.

Целью данной работы является разработка требований к средствам автоматизации процедуры анализа сообщений, т.е. требования к языку программирования, с помощью которого будет создаваться программная реализация процедуры, и требования к вычислительным средствам, на которых с помощью полученной программной реализации процедуры может выполняться анализ сообщений.

Постановка задачи. Перечень требований предъявляемых к средствам автоматизации процедуры анализа сообщений можно разделить на три группы:

- 1) требования к вычислительным средствам, в частности к их производительности и архитектуре, на которых с помощью программной реализации процедуры анализа сообщений может производиться обработка реального потока сообщений;

- 2) требования к языку программирования, позволяющему наиболее удобным образом реализовать процедуру анализа сообщений в соответствии с её алгоритмическими особенностями;
- 3) требования к языку программирования, определяемые удобством использования программной реализации процедуры анализа сообщений.

Для разработки требований, определяемых алгоритмическими особенностями процедуры анализа сообщений, рассмотрим математическое описание ОСМА, являющегося основой рассматриваемой процедуры.

Математическое описание ОСМА. Рассмотрим сообщение длины n знаков, которые принимают значения из множества (алфавита) $A_N = \{a_1, \dots, a_N\}$ мощности N , $|A|=N$. Необходимо проверить простую гипотезу, состоящую в том, все знаки сообщения принимают значения из алфавита A_N с равными вероятностями – $1/N$. Для этого в ОСМА применяется статистический критерий согласия с равновероятным распределением, а в качестве статистики критерия $S_{n,N}$ используется введенная Карлом Пирсоном [3, 4], статистика хи-квадрат – χ_n

$$\chi_n = \sum_{i=1}^N \frac{(h_i - np_i)^2}{np_i},$$

где h_i – частота встречаемости знака (исхода) a_i в первых n испытаниях, n – длина текста (объем выборки), N – число исходов полиномиальной схемы (мощность алфавита A_N) и p_i – вероятность a_i -го исхода.

Для применения статистики $S_{n,N}$ в статистическом критерии необходимо знать распределение вероятности значений данной статистики

$$P\{S_{n,N} \geq c\},$$

для равновероятной полиномиальной схемы $\{p_i = 1/N \mid i = 1, \dots, N\}$.

Под эффективностью применения критерия понимается минимальность ошибки второго рода – β , т.е. минимальное количество сообщений с равновероятным распределением знаков, принятых за сообщения с не равновероятным распределением знаков.

Основой эффективного применения ОСМА и процедуры анализа сообщений в целом является использование в применяемом критерии согласия распределения вероятности, соответствующего параметрам оцениваемого сообщения.

Вычислительная сложность расчета точного распределения статистики $S_{n,N} = \chi_n$ для n и N , оценивается в N^n операций. В [1] рассчитаны максимальные значения n и N , для которых на современном этапе (момент времени t), имея в своем распоряжении в течении месяца (30 дней) вычислительный ресурс $CP(t)$ производительностью 10^{15} операций в секунду, могут быть рассчитаны точные распределения статистики $P_T\{S_{n,N} \geq c\}$, их значения приведены в табл. 1 и на рис. 1.

Таблица 1

Максимальные значения параметров, для которых на современном этапе могут быть рассчитаны точные распределений статистик

Параметры	Значения								
Объем выборки (длина текста) n	58	29	19	17	12	11	9	8	7
Число исходов полиномиальной схемы (мощность алфавита) N	2	4	8	10	26	36	64	128	256

На рис. 1 показаны область точных распределений O_1 и её граница G_1 .

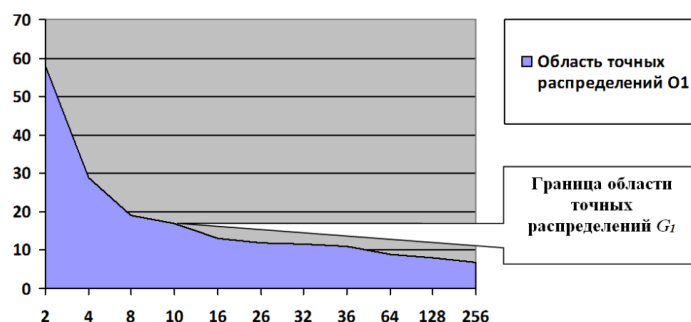


Рис. 1 Диаграмма параметров, для которых на современном этапе могут быть рассчитаны точные распределения статистик

Для расширения границ области расчета точных распределений и сохранения эффективности применения критерия в смысле минимальности ошибки второго рода – β автором в [1] на основе теоретической оценки распределения вероятности значений статистики максимальной частоты $M_n = \max_{i=1}^N h_i$ был предложен метод

расчета Δ -точных распределений $P_{\Delta}\{S_{n,N} \geq c\}$ – распределений, отличающихся от точных распределений не более чем на заранее заданную величину Δ

$$\left| P_{\Delta}\{S_{n,N} \geq c\} - P_T\{S_{n,N} \geq c\} \right| \leq \Delta.$$

Используя метод расчета Δ -точных распределений, для $\Delta = 10^{-5}$ был проведен расчет $P_{\Delta}\{S_{n,N} \geq c\}$ для параметров n и N , значения которых приведены в табл. 2 и на рис. 2.

Таблица 2

Значения параметров, для которых проведены расчеты Δ -точных распределений статистик $S_{n,N}$ для $\Delta = 10^{-5}$

Параметры	Значения							
Объем выборки (длина текста) n	150	80	70	60	50	50	45	40
Число исходов полиномиальной схемы (мощность алфавита) N	2	4	8	10	26	32	128	256

На рис. 2 показаны область точных распределений O_1 и её граница G_1 и область Δ -точных распределений O_2 и её граница G_2 .

Если параметры сообщения n и N находятся вне областей возможного расчета точных распределений O_1 и Δ -точных распределений O_2 , тогда в ОСМА [1] предлагается в качестве распределения статистики использовать предельное распределение – $P\{\chi^2_{(N-1)} \geq c\}$. Данное предположение основано на одном из свойств статистики χ_n^2 , таком, что при $m = \min_{i=1}^N np_i \rightarrow \infty$ ее предельное распределение не зависит от p_1, \dots, p_N и совпадает с χ^2 -распределением с $(N-1)$ степенью свободы [5]. В нашем случае для равновероятной полиномиальной схемы $\{p_i = 1/N \mid i = 1, \dots, N\}$

данное условие принимает вид: $m = n/N \rightarrow \infty$. Считаем, что условие применения предельного распределения выполнено [6, 7] при нахождении параметров n и N вне областей O_1 и O_2 , т.е. выше границы G_2 . Область применения предельных распределений статистик показана на рис. 3.

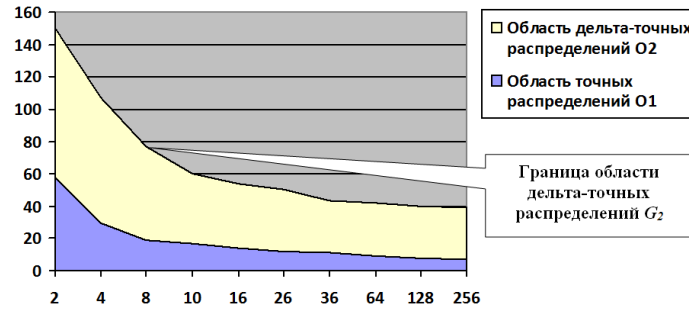


Рис. 2 Значения параметров, для которых рассчитаны Δ -точные распределения статистики $S_{n,N}$

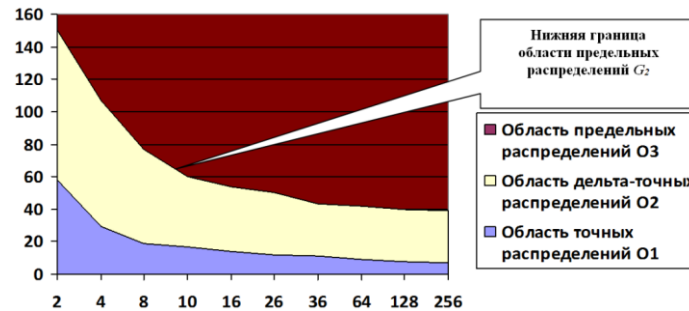


Рис. 3. Область предельных распределений, ограниченная снизу границей G_2

Из приведенных рассуждений следует, что в ОСМА присутствуют три вида распределения вероятностей значений статистик $P\{S_{n,N} \geq c\}$: точные распределения – $P_T\{S_{n,N} \geq c\}$, Δ -точные распределения – $P_\Delta\{S_{n,N} \geq c\}$ и предельные распределения – $P\{\chi^2_{(N-1)} \geq c\}$. По построению ОСМА, областью применения разного вида распределений вероятности значений статистики $P\{S_{n,N} \geq c\}$ является множество пар натуральных чисел (n, N) , для которых выполняются условия эффективности:

- а) Вид распределения определен и может быть рассчитан в момент время t с использованием вычислительного ресурса $CP(t)$;
- б) Применение данного вида распределения в статистическом критерии согласия минимизирует ошибку второго рода β .

Из условий эффективности следует:

1. Точное распределение $P_T\{S_{n,N} \geq c\}$ существует в области O_1 и не существует в областях O_2 и O_3 ;
2. Δ -точное распределение $P_\Delta\{S_{n,N} \geq c\}$ существуют в области O_2 и не существуют в областях O_1 и O_3 ;
3. Предельное распределение $P\{\chi^2_{(N-1)} \geq c\}$ существует в области O_3 и не существует в областях O_1 и O_2

Границы $G_1(t)$ и $G_2(t)$, разделяющие области (множества), зависят от имеющегося во время t вычислительного ресурса $CP(t)$ и с его ростом увеличиваются, сдвигаются вверх.

Таким образом, в ОСМА в зависимости от того, какой области $\{O_i | i=1, 2, 3\}$ принадлежат параметры анализируемого сообщения (n, N) , для сохранения наибольшей эффективности, выбирается и рассчитывается соответствующий вид распределения.

Из анализа математического описания ОСМА можно сделать вывод о следующих его алгоритмических особенностях определяющих требования как к производительности и архитектуре вычислительных средств, так и требования к языку программирования.

Требования к вычислительным средствам. Алгоритмические особенности ОСМА, определяющие требования к производительности и архитектуре вычислительных средств – 1-я группа требований:

1.1. большая вычислительная сложность алгоритма расчета распределений порядка $\sim N^n$;

1.2. структура алгоритма ОСМА содержит логические переходы.

Необходимо отметить, что поскольку процедура анализа сообщений предназначена для обработки сообщений из глобального информационного пространства, число сообщений и скорость передачи которых постоянно растут, вычислительные средства для её реализации должны быть самыми производительными. На современном этапе развития вычислительной техники самыми производительными являются многопроцессорные вычислительные системы.

Учитывая, что при обработке одного сообщения необходимо произвести процедуру расчета распределения вероятностей значений статистики сложностью $\sim N^n$ операций, требования к производительности вычислительных средств повышаются многократно. Поэтому для обеспечения требований ОСМА по производительности необходимо использовать многопроцессорные вычислительные системы на основе программируемых логических интегральных схем (ПЛИС) [7].

Анализ математического описания ОСМА показал наличие в структуре его алгоритма логических переходов, что требует применения для его выполнения универсальных процессоров, выполняющих функцию управления вычислительным процессом.

Таким образом, подытоживая все требования, предъявляемые ОСМА к вычислительным средствам, можно сделать вывод, что для эффективного применения программной реализации процедуры анализа сообщений необходимо применение гибридных высокопроизводительных многопроцессорных вычислительных систем на основе ПЛИС.

Требования к языку программирования. Алгоритмические особенности ОСМА, определяющие требования к языку программирования для создания программной реализации процедуры анализа сообщений, следующие – 2-я группа требований:

2.1. каждая область применения различного вида распределений $\{O_i | i=1, 2, 3\}$ представляет из себя множество множеств;

2.2. большая мощность областей применения различного вида распределений $\{O_i | i=1, 2, 3\}$ не позволяет разместить их целиком на запоминающих устройствах используемых и перспективных вычислительных средств;

2.3. для параметров сообщения $(n, N) \in O_{\nu}$, принадлежащих различным O_{ν} , соответствующие распределения вероятности значений статистики $P\{S_{n,N} \geq c\}$ имеет различный вид: $P_T\{S_{n,N} \geq c\}$, $P_{\Delta}\{S_{n,N} \geq c\}$ или $P\{\chi_{(N-1)} \geq c\}$, – оставаясь по своей сути одними и теми же информационно неразличимыми объектами – распределением вероятности значений статистики.

Ориентируясь на парадигму ресурснезависимого программирования, проработанную в [8, 9] для программирования прикладных программ на гибридных высокопроизводительных вычислительных системах различных архитектур и конфигураций, определим следующие требования к языку программирования – 3-я группа требований:

3.1. Программирование на гибридных вычислительных комплексах различных архитектур и конфигураций;

3.2. Возможность не изменять управляющие программы каждый раз под новую конфигурацию вычислительного комплекса;

3.3. Возможность не переделывать программы под новую архитектуру вычислительного комплекса.

Сравнение языков программирования. Сравним языки программирования, используемые для программирования аналогичных ОСМА методов, на предмет выполнения требований предъявляемых алгоритмическими особенностями ОСМА. В таблице 3 представлены характеристики сравниваемых языков программирования Cryptol [10], COLAMO [11], SET@L [12].

Таблица 3

Сравнение характеристик языков программирования

Характеристики языков программирования	Сравниваемые языки программирования		
	Cryptol	COLAMO	SET@L
Тип: язык высокого уровня	+	+	+
Вид представления данных:			
– массив	+	+	+
– множество / подмножество	-	-	+
– множество множеств	-	-	+
Понятия:			
– горизонт	-	-	+
– информационная неразличимость объектов	-	-	+
– вычислительный виртуальный ресурса	-	-	+
Генерация кода алгоритма для универсальных процессоров	+	+	+
Генерация кода алгоритма для вычислительных полей на ПЛИС	-	+	+
Итоговый рейтинг языка:	3	4	9

Наибольший рейтинг по итогам сравнения соответствия требований ОСМА к возможностям и характеристикам сравниваемых языков программирования набрал язык SET@L.

Возможность выполнения требований к средствам автоматизации процедуры анализа сообщений. По итогам сравнения характеристик языков программирования требования к языку программирования для создания программной реализации процедуры анализа сообщений наиболее полно реализованы в создан-

ном языке SET@L [13], позволяющем реализовать парадигму ресурсонезависимого программирования на гибридных высокопроизводительных вычислительных системах различных архитектур и конфигураций использующих ПЛИС.

В языке SET@L присутствуют программные конструкции, такие как множество и подмножество, реализована возможность построения множества множеств. Данные программные конструкции позволяют реализовать требование к виду представлений области применения распределений (требование 2.1).

В языке SET@L введено понятие горизонта, позволяющее хранить не всю последовательность элементов множества, а только её ограниченную часть и направление движения к горизонту. Указанное понятие позволяет организовать работу с данными большого объема (требование 2.2).

Понятие информационной неразличимости, введенное в языке SET@L, позволяет единообразно описать разные виды распределений, использующихся в ОСМА (требование 2.3).

В языке SET@L реализована парадигма, совокупность подходов, обеспечивающая ресурсонезависимое программирование за счет представления вычислительного ресурса вычислительного комплекса на ПЛИС как виртуального ресурса по аналогии с тем, как в вычислительных системах на универсальных процессорах рассматривается оперативная память (3-я группа требований).

Заключение и выводы.

1. Разрабатываемый для управления гибридными вычислительными комплексами язык программирования SET@L соответствует требованиям, предъявляемые процедурой анализа сообщений к её программной реализации.

2. Процедура анализа сообщений может быть компактно описана на языке SET@L и без семантического разрыва отображена на вычислительные компоненты гибридных вычислительных комплексов – ПЛИС и на их управляющие компоненты – универсальные процессоры.

3. Процедура анализа сообщений может эффективно применяться для анализа потока сообщений при использовании её программной реализации на гибридных высокопроизводительных вычислительных комплексах на основе ПЛИС.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Мельников А.К., Ронжин А.Ф. Обобщенный статистический метод анализа текстов, основанный на расчете распределений вероятности значений статистик // Информатика и её применения. – 2016. – Т. 10. – Вып. 4 (в печати). ISSN 1992-2264.
2. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с.
3. Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables in such that it can be reasonably supposed to have arisen from random sampling // Philosophical Magazine. Series 5. – 1900. – Vol. 50, No. 302. – P. 157-175.
4. Neyman F., Pearson E.S. On the use and interpretation of certain test criteria for purposes of statistical inference // Biometrika. – 1928. – Vol. 20-A. – P. 175-240, 264-299.
5. Smith P.F., Rae D.S., Manderscheid R.W., Silbergeld S. Exact and approximate distributions of the chi-squared statistic for equiprobability // Commun. Statist. – 1979. – Vol. B 8 (2), No. 1. – P. 131-149.
6. Фишер Р.А. Статистические методы для исследователей. – М.: Госстатиздат, 1958. – 73 р.
7. Кендалл М.Г., Стьюарт А. Теория распределений. – М.: Наука, 1966. – 302 с.
8. Каляев И.А., Левин И.И., Семерников Е.А., Шмойлов В.И. Реконфигурируемые мультиконвейерные вычислительные структуры. – 2-е изд., перераб., доп. / под общ. ред. И.А. Каляева. – Ростов-на-Дону: Изд-во ЮНЦ РАН 2009. – 344 с. ISBN 978-5-902982-61-6.
9. Гузик В.Ф., Каляев И.А., Левин И.И. Реконфигурируемые вычислительные системы: учеб. пособие / под общ. ред. И.А. Каляева. – Таганрог: Изд-во ЮФУ, 2016. – 472 с. ISBN 978-5-9275-1918-7.
10. http://www.galois.com/technology/communications_security/cryptol.

11. Левин И.И., Дордопуло А.И., Гудков В.А. Программирование реконфигурируемых вычислительных узлов на языке COLAMO: учеб. пособие. – 2-е изд. доп. – Таганрог: Изд-во ЮФУ, 2016. – 114 с.
12. Левин И.И., Мельников А.К. Методы управления гибридными высокопроизводительными вычислительными комплексами // Суперкомпьютерные технологии (СКТ-2014): Материалы 3-й Всероссийской научно-технической конференции: в 2 т. Т. 1. – Ростов-на-Дону. Изд-во ЮФУ, 2014. – С. 55-60. ISBN 978-9275-1283-6.
13. 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. – С. 23-28. ISSN 978-966-8736-05-6.
14. Левин И.И., Дордопуло А.И., Гудков В.А., Гуленок А.А., Бовкун А.В. Программирование вычислительных систем гибридного типа на языке высокого уровня COLAMO: учеб. пособие для вузов. – Ростов-на-Дону: Изд-во ЮФУ, 2016. – 92 с.
15. Dordopulo Alexey Igorevich, Kovalenko Vasily Borisovich, Gudkov Viacheslav Aikhandrovich, Slasten Liubov Mikhailovna. Porting of parallel applications to reconfigurable computer systems with various architectures and configurations // 2016 Sth International Conference on Informatics Electronics and Vision (ICIEV) Year: 2016. – P. 1122-1127. DOI: 10.1109/ICIEV.2016.7760174.
16. Дордопуло А.И., Левин И.И., Каляев И.А., Гудков В.А., Гуленок А.А. Ресурснезависимое программирование вычислительных систем гибридного типа на языке программирования COLAMO // Суперкомпьютерные дни в России. 26–27 сентября 2016 г., Москва, Россия. – С. 1087-1098.
17. Левин И.И., Дордопуло А.И., Каляев И.А., Доронченко Ю.И., Раскладкин М.К. Современные и перспективные высокопроизводительные вычислительные системы с реконфигурируемой архитектурой // Вестник Южно-Уральского государственного университета. Серия Вычислительная математика и информатика. – 2015. – Т. 4. № 3. – С. 24-37. ISSN: 2410-7034.
18. Левин И.И., Дордопуло А.И. Программирование перспективных вычислительных систем с реконфигурируемой архитектурой на языке COLAMO // Материалы XI Международной научной и практической конференции «Современная европейская наука (Modern european science -2015)», 30 июня – 7 июля 2015, Шеффилд, Англия. – С. 3-9. ISBN 978-966-8736-05-6.
19. Левин И.И., Доронченко Ю.И., Мельников А.К. Эффективная реализация алгоритмов с глубокими циклами на реконфигурируемых вычислительных системах // Материалы XI Международной научной и практической конференции «Современная европейская наука (Modern european science - 2015)», 30 июня – 7 июля 2015, Шеффилд, Англия. – С. 10-14. ISBN 978-966-8736-05-6.
20. Левин И.И., Дордопуло А.И., Каляев И.А., Гудков В.А., Гуленок А.А. Программирование вычислительных систем гибридного типа на основе метода редукции производительности // Труды Международной научной конференции «Параллельные вычислительные технологии (ПаВТ'2016)», г. Архангельск, 28 марта – 1 апреля 2016 г. – Челябинск: Издательский центр ЮУрГУ, 2016. – С. 131-140. ISBN 978-5-696-04663-1.

REFERENCES

1. Mel'nikov A.K., Ronzhin A.F. Obobshchennyu statisticheskiy metod analiza tekstov, osnovannyy na raschete raspredeleniy veroyatnosti znacheniy statistik [A generalized statistical method of analyzing texts based on the calculation of probability distributions of values of statistics], *Informatika i ee primeneniya* [Informatics and its applications], 2016, Vol. 10, Issue 4 (in press). ISSN 1992-2264.
2. Kramer G. Matematicheskie metody statistiki [Mathematical methods of statistics]. Moscow: Mir, 1975, 648 p.
3. Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables in such that it can be reasonably supposed to have arisen from random sampling, *Philosophical Magazine. Series 5*, 1900, Vol. 50, No. 302, pp. 157-175.
4. Neyman F., Pearson E.S. On the use and interpretation of certain test criteria for purposes of statistical inference, *Biometrika*, 1928, Vol. 20-A, pp. 175-240, 264-299.

5. *Smith P.F., Rae D.S., Manderscheid R.W., Silbergeld S.* Exact and approximate distributions of the chi-squared statistic for equiprobability, *Commun. Statist.*, 1979, Vol. B 8 (2), No. 1, pp. 131-149.
6. *Fisher R.A.* Statisticheskie metody dlya issledovateley [Statistical methods for researchers]. Moscow: Gosstatizdat, 1958, 73 p.
7. *Kendall M.G., St'yuart A.* Teoriya raspredeleniy [The theory of distributions]. Moscow: Nauka, 1966, 302 p.
8. Rekonfiguriruemye mul'tikonveyernye vychislitel'nye struktury [Multiconference reconfigurable computing structure]. 2nd ed., pererab., dop., ed. by I.A. Kalyaeva. Rostov-on-Don: Izd-vo YuNTs RAN 2009, 344 p. ISBN 978-5-902982-61-6.
9. *Guzik V.F., Kalyaev I.A., Levin I.I.* Rekonfiguriruemye vychislitel'nye sistemy: ucheb. posobie [Reconfigurable computing systems: the textbook], ed. by I.A. Kalyaeva. Taganrog: Izd-vo YuFU, 2016, 472 p. ISBN 978-5-9275-1918-7.
10. Available at: http://www.galois.com/technology/communications_security/cryptol.
11. *Levin I.I., Dordopulo A.I., Gudkov V.A.* Programmirovaniye rekonfiguriruemykh vychislitel'nykh uzlov na yazyke COLAMO: ucheb. posobie [Programming reconfigurable computational nodes in the language COLAMO: a training manual]. 2nd ed. Taganrog: Izd-vo YuFU, 2016, 114 p.
12. *Levin I.I., Mel'nikov A.K.* Metody upravleniya gibridnymi vysokoproizvoditel'nymi vychislitel'nymi kompleksami [Methods of management of hybrid high performance computing systems], Superkomp'yuternye tekhnologii (SKT-2014): Materialy 3-y Vserossiyskoy nauchno-tekhnicheskoy konferentsii [Supercomputer technologies (SCT-2014): Materials of the 3rd all-Russian scientific-technical conference]: in 2 vol. Vol. 1. Rostov-on-Donu. Izd-vo YuFU, 2014, pp. 55-60. ISBN 978-9275-1283-6.
13. *Levin I.I., Mel'nikov A.K.* Upravlenie gibridnymi vychislitel'nymi sistemami na yazyke SET@L [Control hybrid computer systems in the language of SET@L,] *Materiály XI mezinárodní vědecko-praktická konference «Aktuální vymoženosti vědy – 2015» Díl 7. Moderní informační technologie* [Materials XI international scientific-practical conference "Actual conveniences of science – 2015" Episode 7. The modern information technology]. Praha: Publishing House «Education and Science», 2015, 96 p., pp. 23-28. ISSN 978-966-8736-05-6.
14. *Levin I.I., Dordopulo A.I., Gudkov V.A., Gulenok A.A., Bovkun A.V.* Programmirovaniye vychislitel'nykh sistem gibridnogo tipa na yazyke vysokogo urovnya COLAMO: ucheb. posobie dlya vuzov [Programming the computing system of the hybrid type for high level language COLAMO: proc. the manual for high schools.]. Rostov-on-Don: Izd-vo YuFU, 2016, 92 p.
15. *Dordopulo Alexey Igorevich, Kovalenko Vasilii Borisovich, Gudkov Viacheslav Aixandrovich, Slasten Liubov Mikhailovna.* Porting of parallel applications to reconfigurable computer systems with various architectures and configurations, *2016 Sth International Conference on Informatics Electronics and Vision (ICIEV) Year: 2016*, pp. 1122-1127. DOI: 10.1109/ICIEV.2016.7760174.
16. *Dordopulo A.I., Levin I.I., Kalyaev I.A., Gudkov V.A., Gulenok A.A.* Resursonezavisimoe programmirovaniye vychislitel'nykh sistem gibridnogo tipa na yazyke programmirovaniya COLAMO [Resourcesavesize programming computing systems hybrid programming language COLAMO], *Superkomp'yuternye dni v Rossii. 26-27 sentyabrya 2016 g., Moskva, Rossiya* [Supercomputing in Russia. September 26-27, 2016, Moscow, Russia], pp. 1087-1098.
17. *Levin I.I., Dordopulo A.I., Kalyaev I.A., Doronchenko Yu.I., Raskladkin M.K.* Sovremennye i perspektivnye vysokoproizvoditel'nye vychislitel'nye sistemy s rekonfiguriruemoy arkhitekturoy [Modern and perspective high performance computer systems with reconfigurable architecture], *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya Vychislitel'naya matematika i informatika* [Vestnik of the South Ural state University. A series of Computational mathematics and Informatics], 2015, Vol. 4. No. 3, pp. 24-37. ISSN: 2410-7034.
18. *Levin I.I., Dordopulo A.I.* Programmirovaniye perspektivnykh vychislitel'nykh sistem s rekonfiguriruemoy arkhitekturoy na yazyke COLAMO [Programming perspective computer systems with reconfigurable architecture on the COLAMO language], *Materialy XI Mezhdunarodnoy nauchnoy i prakticheskoy konferentsii «Sovremennaya evropeyskaya nauka (Modern european science - 2015)», 30 iyunya – 7 iyulya 2015, Sheffield, Angliya* [Materials of XI International scientific and practical conference "Modern European science (Modern european science - 2015)", 30 June – 7 July 2015, Sheffield, England], pp. 3-9. ISBN 978-966-8736-05-6.

19. Levin I.I., Doronchenko Yu.I., Mel'nikov A.K. Effektivnaya realizatsiya algoritmov s glubokimi tsiklami na rekonfiguriruemyykh vychislitel'nykh sistemakh [Efficient implementation of algorithms with a deep cycle for reconfigurable computing systems], *Materialy XI Mezhdunarodnoy nauchnoy i prakticheskoy konferentsii «Sovremennaya evropeyskaya nauka (Modern european science - 2015)»*, 30 iyunya – 7 iyulya 2015, Sheffild, Angliya [proceedings of the XI International scientific and practical conference "Modern European science (Modern european science - 2015)", 30 June – 7 July 2015, Sheffield, England], pp. 10-14. ISBN 978-966-8736-05-6.
20. Levin I.I., Dordopulo A.I., Kalyaev I.A., Gudkov V.A., Gulenok A.A. Programmirovaniye vychislitel'nykh sistem gibridnogo tipa na osnove metoda reduktsii proizvoditel'nosti [Programming the computing system of the hybrid type based on the method of reduction performance], *Trudy Mezhdunarodnoy nauchnoy konferentsii «Parallelnyye vychislitel'nye tekhnologii (PaVT'2016)»*, g. Arkhangel'sk, 28 marta – 1 aprelya 2016 g. [Proceedings of International scientific conference "Parallel computational technologies (PCT' 2016)", Arkhangel'sk, March 28 – April 1, 2016]. Chelyabinsk: Izdatel'skiy tsentr YuUrGU, 2016, pp. 131-140. ISBN 978-5-696-04663-1.

Статью рекомендовал к опубликованию д.т.н., профессор И.И. Левин.

Мельников Андрей Кимович – ИТЦ ЗАО «ИнформИнвестГрупп»; e-mail: ak@iigroup.ru; 117587, Москва, Варшавское шоссе, 125, стр. 17; тел.: 84952870035; к.т.н.; доцент ВАК; г.н.с.

Melnikov Andrey Kimovitch – STC CLSC «InformInvestGroup»; e-mail: ak@iigroup.ru; 125, Varshavskoye road, build. 17, Moscow, 117587, Russia; phone: +74952870035; cand. of eng. sc.; associate professor of SAC; chief research officer.

УДК 004.382.2

DOI 10.18522/2311-3103-2016-12-1628

Д.А. Сорокин, А.Ю. Матросов, Е.Е. Семерникова, К.Н. Алексеев
СТРУКТУРНО-ПРОЦЕДУРНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА
ПРОГНОЗИРОВАНИЯ КРАТНЫХ ВОЛН НА ПЛИС*

Рассматриваются особенности решения на высокопроизводительных вычислительных системах задачи прогнозирования кратных волн-помех с помощью алгоритма SRMP. Данный алгоритм относится к классу вычислительно трудоемких сильносвязанных задач, в которых число информационных межпроцессорных обменов и обменов между процессорами и элементами памяти сравнимо или превышает число выполняемых операций. Для эффективной реализации данной задачи требуется обеспечение многоканальности в сочетании с нелинейным доступом к данным, что практически неосуществимо на вычислительных системах традиционной архитектуры. В связи с этим представлен альтернативный подход к решению задачи SRMP, основанный на идее построения единого вычислительного контура на реконфигурируемых вычислительных системах (PBC). Особенность PBC заключается в том, что данные вычислительные системы используют программируемые логические интегральные схемы (ПЛИС) в качестве основного вычислительного компонента и способны адаптироваться под вычислительную структуру решаемой задачи. Структурно-процедурная организацией вычислений, одна из самых эффективных для PBC, предполагает непосредственное отображение базового информационного графа задачи на вычислительное поле системы. Для такой реализации задачи SRMP требуется наличие вычислительного ресурса, которым современные PBC на данный момент не обладают. Принимая во внимание особенности алгоритма, был предложен подход, позволивший преобразовать базовый информационный граф задачи и разработать структуру вычислительного конвейера, так что в каждой вычислительной ПЛИС

* Работа выполнена при финансовой поддержке из бюджета Союзного государства в рамках реализации государственного контракта Министерства образования и науки Российской Федерации от 17 июня 2015г. № 14.964.11.0001.