

## ПОДХОД К СОЗДАНИЮ ИНФОРМАЦИОННОЙ СТРУКТУРЫ ПАРАЛЛЕЛЬНОГО СТАТИСТИЧЕСКОГО ЭКСПЕРИМЕНТА

**И.В. Тихонов**

*Кафедра «Автоматизированные системы», ГОУ ВПО «Иркутский  
государственный технический университет»; tihily@mail.ru*

*Представлена членом редколлегии профессором В.И. Коноваловым*

**Ключевые слова и фразы:** параллельные вычисления; статистические испытания.

**Аннотация:** Рассмотрена возможность и принципы организации серии статистических испытаний, исполняемых параллельно на множестве вычислительных узлов с единой коммуникационной средой. Проанализирована последовательность этапов метода статистических испытаний на предмет их возможной параллелизации без существенного ограничения общности.

---

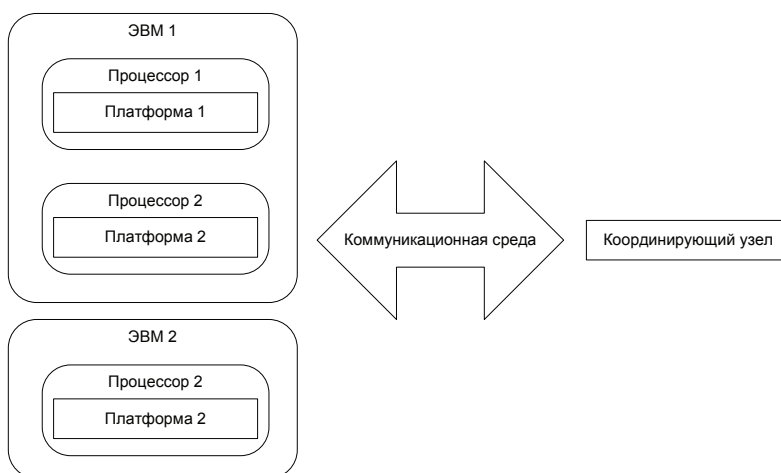
Проведены исследования с целью сокращения времени, затрачиваемого для получения результата в тех задачах имитационного моделирования, которые можно рассматривать с позиций статистического эксперимента ввиду случайного характера моделируемых процессов.

*Предпосылки к параллелизации статистического эксперимента.* Для проведения статистического эксперимента, ввиду его существенной вычислительной сложности, необходимы ЭВМ, которые на сегодняшний день зачастую являются параллельными [1]. В основе же статистического эксперимента лежит многократное повторение его независимых итераций, что дает возможность одновременного их исполнения с последующим сведением для получения оценки искомых выходных величин с требуемой точностью. С учетом этого является возможным параллельное проведение статистического эксперимента в несколько потоков (в том числе на сети ЭВМ), с целью либо сокращения времени расчета эксперимента, либо увеличения точности получаемых результатов [2].

*Компоненты окружения для параллельного проведения статистического эксперимента.* Логично предположить структуру программного средства по параллельному проведению статистических экспериментов на базе данного тезиса, состоящую:

- из платформы для прогонов имитационных моделей, выполняемых на независимых элементах некой вычислительной структуры;
- самой имитационной модели, согласно неким единообразным правилам;
- координирующего узла, управляющего процессом проведения статистического эксперимента и получающего результаты;
- коммуникационной среды для организации обмена сообщениями между платформой и координирующим узлом.

Общую схему взаимодействия вычислительных ресурсов, платформы и координирующего узла можно представить схемой, показанной на рис. 1.



**Рис. 1. Схема взаимодействия вычислительных ресурсов, платформы и координирующего узла**

*Этапы метода статистических испытаний.* Определим возможный набор правил разработки имитационных моделей, пригодных для параллельного исполнения на современных вычислительных средствах. Для этого рассмотрим последовательность этапов итерации статистического эксперимента согласно методу статистических испытаний:

1) получение входных воздействий, хотя бы одно из которых является случайным и подчиняется некоторому закону распределения вероятностей;

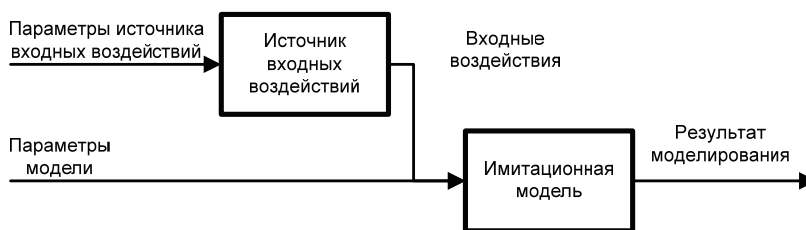
2) прогон имитационной модели с полученными на первом этапе входными воздействиями. Так как целью создания программного средства проведения статистических экспериментов является повышение эффективности данного процесса, целесообразно использовать универсальные языки программирования для описания поведения моделей [4];

3) обработка полученных случайных величин, где в качестве оценок результатов многократного моделирования используются их различные вероятностные характеристики. Это позволяет формировать оценки с учетом экономии памяти ЭВМ [5]. Получение данной оценки  $r$  на  $n$ -й итерации из исходных данных может быть описано следующим образом:  $r_n = v_1 \cdot v_2 \cdot \dots \cdot v_n$ , где  $v_i$  – выходная величина  $i$ -й итерации эксперимента; « $\cdot$ » – операция вычисления оценки;

4) принятие решения о продолжении или завершении эксперимента, или задача об определении размера выборки не может быть решена однозначно и однозначно, так как опирается на множество различных факторов. Рассматривая общеизвестные подходы [4], можно принять общее правило, согласно которому решение о завершении эксперимента принимается каждый раз после получения очередной оценки из этапа 3 согласно неким разработанным исследователем правилам. Это позволяет использовать как методы с априорным знанием требуемого количества итераций, так и методы автоостанова или даже комбинации различных методов.

Рассмотрим информационные связи между данными четырьмя этапами. При этом входные и выходные значения могут быть представлены как отдельными числовыми значениями, так и наборами этих значений без ограничения общности.

*Имитационная модель и источники входных воздействий.* Входные воздействия, как было сказано ранее, могут являться либо детерминированными, либо случайными, с собственным набором параметров закона распределения вероятностей (рис. 2Рис. 2).



**Рис. 2. Взаимодействие источника входных воздействий и имитационной модели**

Структура источника входных воздействий и имитационной модели (рис. 2) совпадают вплоть до используемых типов данных (что было оговорено выше). Поэтому возможно модифицировать понятие имитационной модели, рассматривая ее как ациклическую иерархическую структуру вычислительных элементов с множеством входов и единственным выходом, объединенных связями. Такая унификация позволяет повторно использовать реализацию общих для различных имитационных моделей алгоритмов, выделенных в отдельные компоненты, а также обеспечивает возможность более гибкого взаимодействия компонентов моделей и, наконец, упрощает построение моделируемых систем.

*Обработка результатов моделирования.* Результат моделирования, полученный на выходном компоненте имитационной модели, может быть обработан одним или несколькими независимыми алгоритмами для получения интересующих исследователя статистических оценок. Возможность независимого использования таких алгоритмов в совокупности с унифицированным характером выходных величин предоставляет широкие возможности по повторному использованию и комбинированию реализаций этих алгоритмов при проведении различных экспериментов.

*Принятие решения о продолжении или завершении эксперимента.* Итогом работы алгоритма принятия решения о продолжении или завершении эксперимента на базе полученных результатов моделирования, очевидно, может являться бинарное значение. В этом случае появляется возможность их комбинирования в бинарные выражения с использованием таких операций, как конъюнкция и дизъюнкция. Необходимость в этом может возникнуть в случае наличия нескольких различных условий для останова эксперимента, причем для исследователя существенен не столько факт останова, сколько условие, его вызвавшее.

*Коммуникация при проведении статистического эксперимента в распределенной среде. Ограничения.* Поскольку прогоны имитационной модели проводятся независимо и параллельно, они должны выполняться на платформе для проведения статистических экспериментов. Принятие решения о завершении эксперимента может быть выполнено только на координирующем узле, так как именно этот компонент обладает полной статистической информацией по результатам всех прогонов на каждой вычислительной платформе. Обработка же полученных выходных величин при проведении параллельного статистического эксперимента должна быть сопряжена с процессом передачи результатов моделирования на координирующий узел. Очевидно, что:

- 1) алгоритм статистической обработки полученных результатов моделирования расположен и исполняется на координирующем узле;
- 2) значения выходных величин, получаемые после прогона имитационной модели, с каждой вычислительной платформы передаются на координирующий узел при помощи коммуникационной среды, и последовательно обрабатываются в нем алгоритмом для получения текущей оценки.

При подобной реализации невозможно гарантировать порядок получения результатов параллельно выполняющихся итераций статистического эксперимента координирующим узлом, поэтому операция  $\cdot$  должна обладать свойством коммутативности, так как последовательность отдельных итераций статистического эксперимента не должна играть никакой роли в нахождении результата (такому условию удовлетворяют, например, методы отыскания оценок среднего значения, дисперсии, построения гистограммы распределения)

$$r_n = v_1 \cdot v_2 \cdot \dots \cdot v_{i-1} \cdot v_i \cdot \dots \cdot v_n = v_1 \cdot v_2 \cdot \dots \cdot v_i \cdot v_{i-1} \cdot \dots \cdot v_n \cdot$$

Однако этот способ коммуникации имеет следующие недостатки:

- все расчеты по статистической обработке полученных величин ложатся на координирующий узел, который может в такой ситуации стать сдерживающим фактором высокой производительности решения;

- данный способ не учитывает потенциально низкую производительность коммуникационной среды, что также может негативно отражаться на производительности.

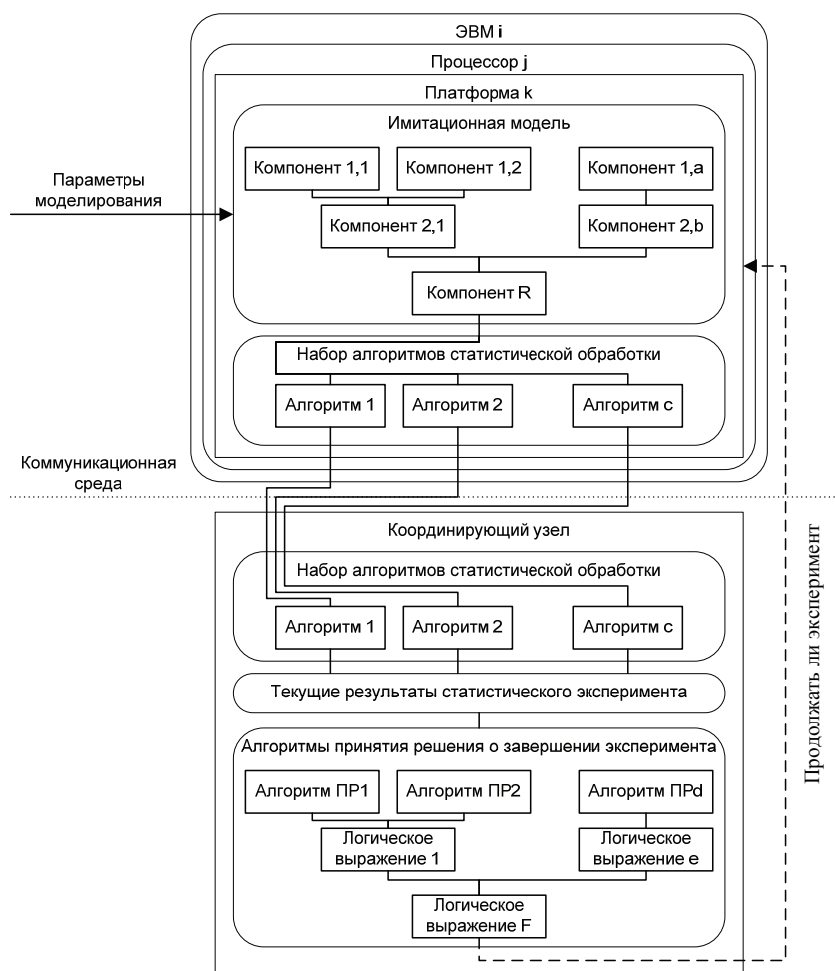
Возможным подходом к устранению этих недостатков может являться промежуточная статистическая обработка результатов итераций эксперимента на каждой из вычислительных платформ. Реализация данного промежуточного слоя с минимальными затратами возможна в случае, если операция « $\cdot$ » вычисления статистической оценки  $r$  является ассоциативной (например, нахождение оценок среднего значения, дисперсии, построение гистограммы распределения), тогда она может быть использована и для вычисления промежуточных оценок согласно следующему выражению:

$$r = v_{1,1} \cdot v_{1,2} \cdot \dots \cdot v_{1,i_1} \cdot v_{2,1} \cdot v_{2,2} \cdot \dots \cdot v_{2,i_2} \cdot \dots \cdot v_{m,im} = \\ = (v_{1,1} \cdot v_{1,2} \cdot \dots \cdot v_{1,i_1}) \cdot (v_{2,1} \cdot v_{2,2} \cdot \dots \cdot v_{2,i_2}) \cdot \dots \cdot v_{m,im} = r_1 \cdot r_2 \cdot \dots \cdot r_m,$$

где  $m$  – количество вычислительных платформ;  $r_p = v_{p,1} \cdot v_{p,2} \cdot \dots \cdot v_{p,i_p}$  – промежуточная статистическая оценка, полученная на  $p$ -й вычислительной платформе;  $i_p$  – количество итераций эксперимента, выполненных на  $p$ -й вычислительной платформе, позволяет балансировать нагрузку между коммуникационной средой и координирующим узлом.

*Структура параллельного статистического эксперимента.* Информационная структура параллельного статистического эксперимента с учетом вышесказанного представлена на рис. 3.

Программная реализация имитационной модели, представленной совокупностью взаимосвязанных компонентов, при очередном прогоне статистического эксперимента на платформе  $k$  генерирует выходное значение, которое попадает на вход требуемых алгоритмов статистической обработки. По прошествии  $i_k$  итераций полученные промежуточные статистики при помощи коммуникационной среды отправляются на координирующий узел. Там на базе результатов, полученных к текущему моменту со всех платформ, вычисляются итоговые статистики, на основе которых совокупностью определенных алгоритмов принимается решение о завершении и продолжении эксперимента с отправкой соответствующей команды на каждую из платформ. На базе предложенной структуры разработано программное средство, показывающее эффективность примененных подходов.



**Рис. 3. Информационная структура параллельного статистического эксперимента**

*Заключение.* Проанализирована последовательность этапов метода статистических испытаний на предмет их возможной параллелизации без существенного ограничения общности. Полученным новым результатом является информационная структура, описывающая такие ограничения, а также декомпозицию аспектов статистического эксперимента, при которых становится возможным его параллельное исполнение.

#### *Список литературы*

1. Bönisch, T. Parallel Hardware Architectures / T. Bönisch // V рос.-герм. шк. по параллельным вычислениям на высокопроизвод. вычисл. системах. Т. 1. – Новосибирск, 2008. – С. 7–20.
2. Тихонов, И.В. Вычислительный кластер статистических экспериментов / И.В. Тихонов // Тр. XIII Байкал. Всерос. конф. «Информационные и математические технологии в науке и управлении». – Иркутск : Изд-во Ин-та систем энергетики им. Л.А. Мелентьева. Сиб. отд-ние РАН, 2008. – Т. 1. – С. 177–183.
3. Эндрюс, Г.Р. Основы многопоточного, параллельного и распределенного программирования / Г.Р. Эндрюс. – М. : Вильямс, 2003. – 512 с.

4. Шеннон, Р. Имитационное моделирование систем – искусство и наука / Р. Шеннон. – М. : Мир, 1978. – 418 с.

5. Бусленко, Н.П. Моделирование сложных систем / Н.П. Бусленко. – М. : Наука, 1968. – 356 с.

---

## **Approach to the Creation of Information Structure of Parallel Statistic Experiment**

**I.V. Tikhonov**

*Department “Automated Systems”, Irkutsk State  
Technical University; tihlv@mail.ru*

**Key words and phrases:** parallel computing; statistical tests.

**Abstract:** The article considers the possibility and the principles of the statistical tests set organization, performing in parallel on a set of the computer cross points with the single communication media. The sequence of stages of the statistical tests method for the purpose of their possible parallelization without the significant restriction of the generality is analyzed.

---

## **Herangehen zur Schaffung der Informationsstruktur des parallelen statistischen Experimentes**

**Zusammenfassung:** Es sind die Möglichkeit und die Prinzipien der Organisierung der Reihe der statistischen Versuche, die auf der Menge der Berechnungsknoten mit dem einzigen Kommunikationsmedium parallel durchführen werden, betrachtet. Es ist die Reihenfolge der Etappen der Methode der statistischen Untersuchungen für ihre mögliche Parallelization ohne wesentlichen Beschränkung der Gemeinsamkeit analysiert.

---

## **Approche envers la création de la structure informatique de l'expérience statistique parallèle**

**Résumé:** Sont examinés la possibilité et les principes de l'organisation d'une série des essais statistiques effectués parallèlement sur une multitude des noeuds informatiques avec un milieu unique de communication. Est analysée la séquence des étapes de la méthode des essais statistiques au sujet de leur parallélisation possible sans une limitation substantielle de la communauté.

---

**Автор:** *Тихонов Илья Владимирович* – аспирант кафедры «Автоматизированные системы», ГОУ ВПО «Иркутский государственный технический университет».

**Рецензент:** *Петров Александр Васильевич* – доктор технических наук, профессор кафедры «Автоматизированные системы», ГОУ ВПО «Иркутский государственный технический университет».

---