

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.42:519.673

Мелещня
Дмитрий Владимирович

Алгоритмы и программные средства распараллеливания вычислений при моделировании процессов принятия решений в финансовых системах

АВТОРЕФЕРАТ

на соискание академической степени
магистра технических наук

по специальности 1-40 80 05 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Научный руководитель
Бранцевич П.Ю.
к.т.н., доцент

Минск 2015

КРАТКОЕ ВВЕДЕНИЕ

Вычислительная техника играет важнейшую роль во многих сферах деятельности человека будь то развлечения, медицина или промышленность. Функционирование современных финансовых структур, таких как банки, биржи, рынки ценных бумаг также невозможно представить без компьютерного обеспечения, которое служит для обработки, хранения информации, прогнозирования, оценки и принятия научно обоснованных решений.

Важнейшей нишей в финансовой сфере, которая стала гораздо эффективнее с развитием вычислительной техники – это предоставление финансовой информации для трейдеров. Именно информация зачастую становится ключом для совершения успешных сделок, определения тенденций изменения цены и покупки перспективного актива.

Однако анализа информации недостаточно для принятия решения о покупке того или иного актива либо ценной бумаги. Важным фактором в этом процессе является оценка рисков, связанных с приобретением или продажей этих активов. Для этого существует большое множество алгоритмов моделирования, которые зачастую являются достаточно ресурсоемкими в плане вычислительных мощностей. Поэтому перед трейдерами стоит сложная вычислительная задача, которая может быть решена с использованием параллельных вычислительных систем

OpenCL – фреймворк для написания компьютерных программ, связанных с параллельными вычислениями на различных графических (англ. GPU) и центральных процессорах (англ. CPU), а также FPGA. То есть при использовании данной платформы, алгоритмы финансового моделирования могут быть запущены на любых устройствах, реализующих стандарт.

Таким образом если финансовые модели будут адаптированы для параллельного выполнения, они могут быть запущены используя фреймворк OpenCL на различных вычислительных устройствах, позволяющих достичь высокой степени параллелизма. Такой механизм может позволить получить высокую эффективность вычислений, что в свою очередь может сделать возможным расчеты и моделирования в реальном времени.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью работы является адаптация алгоритма моделирования цены европейского опциона для запуска его на гетерогенной вычислительной системе используя программно-аппаратную платформу OpenCL. В составе аппаратной части могут присутствовать различные типы процессоров. Также необходимо доказать целесообразность использования гетерогенной системы в финансовом моделировании.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Определить возможности применения гетерогенных систем при решении задач, требующих больших вычислительных задач.
2. Выбрать алгоритм расчета необходимых величин, наиболее хорошо поддающийся распараллеливанию.
3. Разработать способы проверки корректности работы алгоритма.
4. Реализовать алгоритм расчета для центрального процессора.
5. Реализовать алгоритм моделирования с использованием OpenCL, а также оптимизировать алгоритм с использованием возможностей, предоставляемых фреймворком и платформой.
6. Сравнить результаты всех реализаций.
7. Разработать вариации алгоритма, позволяющие достичь максимальной эффективности от использования гетерогенной системы.

Объектом исследования алгоритмы моделирования, используемые в финансовых системах при принятии решений.

Предметом исследования является математическое и программное обеспечение компьютерных систем, алгоритмы и методы, предназначенные для моделирования основных финансовых показателей, используемых на рынках ценных бумаг.

Основной гипотезой, положенной в основу диссертационной работы, является возможность использования в финансовых расчетах гетерогенной вычислительной системы, в состав которой могут входить различные типы процессоров, в частности ЭВМ базирующаяся на центральном и графическом процессорах. Поскольку многие алгоритмы финансового моделирования достаточно ресурсоемкие и в то же время хорошо распараллеливаемы, то использование в таких расчетах параллельных вычислительных систем может позволить достичь высокой скорости расчета. Использование программно-аппаратной платформы OpenCL может позволить охватить достаточно широкий спектр оборудования, которое может быть использовано для

вычислений, а возможности оптимизации и различные архитектурные решения позволят получить еще большую производительность.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническим заданием и планом работ кафедры «Программное обеспечение информационных технологий» по теме «Разработать модели, методы, алгоритмы для оценки параметров, повышения надежности и качества функционирования аппаратно-программных средств систем и сетей сложной конфигурации и внедрить в современные обучающие комплексы» (ГБ № 11-2004, № ГР 20111065, научный руководитель НИР – В. В. Бахтизин).

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя П. Ю. Бранцевича, заключается в формулировке целей и задач исследования.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались и обсуждались на международной научной конференции «Информационные технологии и системы 2015» (Минск, Беларусь, 2015); на 3-й международной научно-практической конференции «Веб-программирование и интернет-технологии WebConf2015» (Минск, Беларусь, 2015); на 39-й международной научно-практической конференции «Научная дискуссия: Инновации в современном мире» (Москва, Россия, 2015).

Опубликованность результатов диссертации

По теме диссертации опубликовано 3 печатных работы в сборниках трудов и материалов международных конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников и списка публикаций автора. В первой главе представлен анализ предметной области, выявлены

основные существующие архитектурные, программные и аппаратные решения, позволяющие выполнять вычисления, требовательные к объему вычислительных ресурсов. Вторая глава посвящена обзору алгоритма Монте-Карло, используемому в финансовом моделировании. Также предложена формула Блэка-Шоулза, позволяющая аналитически проверить результаты моделирования. В третьей главе реализован алгоритм моделирования цены европейского опциона для центрального процессора. Далее полученный алгоритм адаптирован для запуска на параллельных вычислительных системах. Представлены результаты экспериментальных исследований моделирования с различными исходными параметрами.

Общий объем работы составляет 70 страниц, из которых основного текста – 58 страниц, 19 рисунков на 8 страницах, 1 таблица, список использованных источников из 46 наименований на 3 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** выполнен обзор эволюции параллельных вычислительных систем, а также рассмотрена возможность запуска сложных вычислительных запусков на гетерогенных устройствах. При этом упор сделан на чаще всего встречающуюся связку центральный процессор и один или несколько графических процессоров. Поэтому детально рассмотрена архитектура графического процессора, а также программных и аппаратных платформ, позволяющих выполнять вычисления общего назначения на GPU.

OpenCL – программный фреймворк для написания программ для гетерогенных систем, поддерживаемый некоммерческим консорциумом Khronos Group. Программы, разрабатываемые с использованием OpenCL, могут исполняться на ряде различных устройств различных производителей. Он поддерживает широкий диапазон уровней параллелизма и эффективно ложится на гомогенные и гетерогенные, с одним или несколькими вычислительными узлами системы. OpenCL предлагает и язык для написания кода на стороне устройства, и для слоя управления хостом.

Фреймворк OpenCL позволяет решать сложнейшие вычислительные задачи, требующие больших вычислительных ресурсов и находит применение в различных областях науки и техники, таких как, например, медицина, обработка видео и аудио, физическое и финансовое моделирование и др.

Одним из самых популярных алгоритмов в сфере финансового моделирования является метод Монте-Карло, суть которого заключается в генерировании большого количества реализаций некоторой случайной величины с заданным отклонением, а затем нахождение усредненного значения. Важной особенностью алгоритма является то, что выполнение вычислений, связанных с одним элементом данных никак не зависят от других элементов. Эти особенности указывают на то, что большая эффективность может быть достигнута, если операции над всеми элементами будут выполняться параллельно. И чем выше будет степень параллелизма, тем более эффективная реализация алгоритма будет получена. Для решения данной задачи прекрасно подходит фреймворк OpenCL.

Поскольку фреймворк OpenCL позволяет запускать вычисления на любых устройствах, реализующих стандарт, мы получим реализацию алгоритма Монте-Карло, которая будет являться переносимой и в то же время достаточно эффективной в случае запуска на устройстве, архитектура которого позволяет достичь высокой степени многопоточности.

Вторая глава посвящена обзору метода моделирования цены европейского опциона. Также разработан аналитический метод проверки корректности результатов, полученных в результате моделирования.

Опцион – это контракт, по которому одна сторона (покупатель) получает право совершить покупку или продажу некоторого актива по оговоренной цене в некоторый момент в будущем или на протяжении установленного периода, а вторая сторона (продавец) несет обязательство продать или купить актив согласно условиям договора. Существуют европейские и американские опционы. Европейский опцион исполняется только в дату его исполнения, оговоренную в контракте. В отличие от него, американский опцион может быть исполнен в любую дату в течение всего срока его действия

Механизм европейского опциона на покупку следующий. Покупатель рассчитывает, что в будущем цена актива ценой N вырастет. Поэтому он покупает опцион на покупку с ценой исполнения $N+M$, заплатив при этом премию P . Тогда, если в будущем цена актива будет меньше чем $N+M$, инвестору не следует исполнять опцион и сумма его потерь будет равной объему начальных инвестиций, т.е. P . Если же цена актива будет выше чем $N+P$, то инвестор купит акции по цене $N+M$ и немедленно продаст их на рынке. Таким образом, чем больше величина $M-P$, тем большую прибыль получит владелец опциона. Принцип действия опциона на продажу противоположен и держатель такого типа инструмента надеется, что цена базового актива будет

падать. Так, пусть инвестор покупает опцион продавца с ценой исполнения $N-M$, где N – цена актива в настоящий момент. Как и в случае с опционом покупателя, за право исполнить опцион в будущем выплачивается премия в размере P . Таким образом, поскольку опцион европейский, он будет исполнен в момент истечения срока его действия только в случае если его цена меньше $N-M$. Т.е. если на дату исполнения цена актива ниже чем $N-P$, покупатель опциона получит прибыль. Если же цена будет выше этой отметки, опцион исполняться не будет и его владелец потеряет средства в размере, равном премии опциона.

Марковский процесс – такой стохастический процесс, в котором каждое следующее значение переменной, зависит только от текущего. Принято считать, что цена актива представляет собой ни что иное, как Марковский процесс. Это согласуется с тем, что в какой-то момент времени она включает в себя все предшествующие значения из-за законов функционирования рынка. Учитывая эти особенности и сделав некоторые допущения, получим следующие аналитические выражения для расчета цены европейского опциона:

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2), \quad (1)$$

$$p = Ke^{-rT} N(-d_2) - S_0 N(-d_1). \quad (2)$$

В этой формуле:

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad (3)$$

$$d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}. \quad (4)$$

$N(x)$ – это интегральная функция нормально распределения с математическим ожиданием 0 и стандартным отклонением 1; c и p – это цены опционов на покупку и продажу активов; S_0 – первоначальная цена акции; K – цена исполнения; r – непрерывно начисляемая безрисковая процентная ставка; σ – волатильность цены акции; T – время, оставшееся до исполнения опциона.

В основе метода Монте-Карло лежат два важнейших закона теории вероятности: усиленный закон больших чисел и центральная предельная теорема. Суть метода заключается в генерировании большого количества

различных значений цены акции. При этом цена акции может зависеть от различных параметров.

Для определения цены данного дериватива методом Монте-Карло необходимо выполнить следующие шаги:

1. Генерируется случайная цена акции;
2. Вычисляется цена производного инструмента;
3. Предыдущие два шага выполняются многократно, генерируя таким образом большое количество выплат по деривативу;
4. Вычисляем среднее значение размера выплат по инструменту;
5. К полученной величине применяется дисконтирование.

Для вычисления цена актива в момент времени используется следующее выражение:

$$S(T) = S(0) \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) T + \sigma \varepsilon \sqrt{T} \right]. \quad (5)$$

Как видно из уравнения, каждая реализация случайной цены акции не зависит от других поэтому все они могут быть сгенерированы параллельно, а затем агрегировано результирующее значение. Таким образом получим алгоритм, время выполнения которого, в случае одновременного расчета всех путей симуляции, будет равно времени расчета одного пути плюс время агрегации результатов. Для проверки значения, полученного в результате моделирования, может быть использована аналитическая формула Блэка-Шоулза. Причем точность результатов симуляции должна расти с увеличением количества путей симуляции.

В третьей главе разработан алгоритм расчета цены опциона на центральном процессоре. Также предложен вариант данного алгоритма, адаптированный для запуска с использованием программно-аппаратной платформы OpenCL. Также проведены экспериментальные исследования направленные на определение эффективности использования гетерогенных систем в моделировании цены. Предложены способы оптимизации кода для устройств, а также архитектурные решения, позволяющие максимально извлечь преимущество от использования параллельных вычислительных систем.

Анализ метода Монте-Карло показывает, что все пути симуляции цены опциона никак не зависят друг от друга. Это значит, что все они могут быть

посчитаны параллельно. И чем больше количество путей, рассчитываемых одновременно, тем выше эффективность расчета. В идеальном случае время расчета будет равно времени расчета одного пути. Однако очевидно, что с увеличением количества путей, будет расти и время вычислений. Оценка сложности алгоритма в о-нотации дает нам сложность $O(N)$, т.е. время расчета растет пропорционально количеству полученных случайным образом значений цены. Рисунок 1 доказывает данный факт, поскольку из него также видно, что время расчета растет линейно с ростом количества путей.

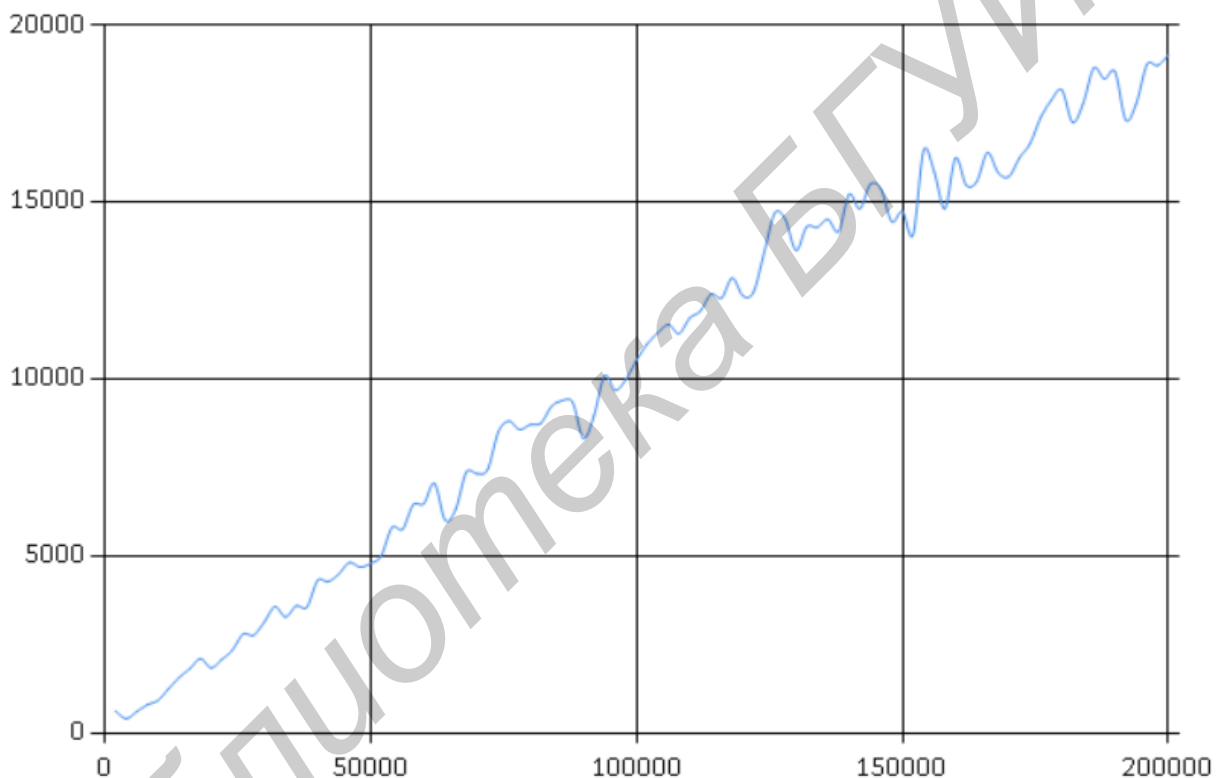


Рисунок 1 – Рост времени расчета с ростом количества путей.

Графический процессор не имеет такого сложного АЛУ как центральный. Однако даже имеющегося набора операция будет достаточно для расчета цены опциона. А поскольку количество ядер GPU значительно выше, чем у центрального процессора, то и степень параллелизма выше, а, следовательно, и теоретическая эффективность запуска расчета на GPU больше.

Фактически задача запуска алгоритма Монте-Карло на графическом процессоре состоит в том, чтобы загрузить исторические данные в память хоста, вычислить параметры симуляции, затем скопировать их в память

устройства. После того как параметры моделирования в памяти устройства можно запускать расчет, который описан функцией ядра. В результате работы программы на устройстве должны будут быть сгенерированы случайные значения цены опциона. Затем эти данные копируются в память хоста, где вычисляется среднее и после этого дисконтированное значение. Сначала складываются все элементы попарно и делятся на два, затем то же самое проделывается с результатом предыдущего шага и так пока не получится среднее по всем элементам.

Важным элементом алгоритма является генерация случайной величины, имеющей нормальное распределение. Однако при реализации данного алгоритма на графическом процессоре возникает проблема в виде отсутствия встроенных функций генераций случайного числа. Поэтому для этих целей случайная величина должна быть предварительно сгенерирована на стороне хоста и затем скопирована в память устройства. Поскольку генерация случайного числа с нормальным распределением основывается на нормально распределенной величине, часть расчета может быть возложена на устройство.

Для моделирования стандартных нормально распределённых случайных величин используется преобразование Бокса-Мюлера. Поскольку программа выполняется на процессоре с SIMD архитектурой, используется вариант преобразования не имеющий ветвлений:

$$z_0 = \cos(2\pi\varphi)\sqrt{-2 \ln r}, \quad (6)$$

$$z_1 = \sin(2\pi\varphi)\sqrt{-2 \ln r}. \quad (7)$$

Результаты экспериментальных исследований времени вычисления цены опциона используя алгоритм Монте-Карло на GPU и CPU представлены на рисунке 2.

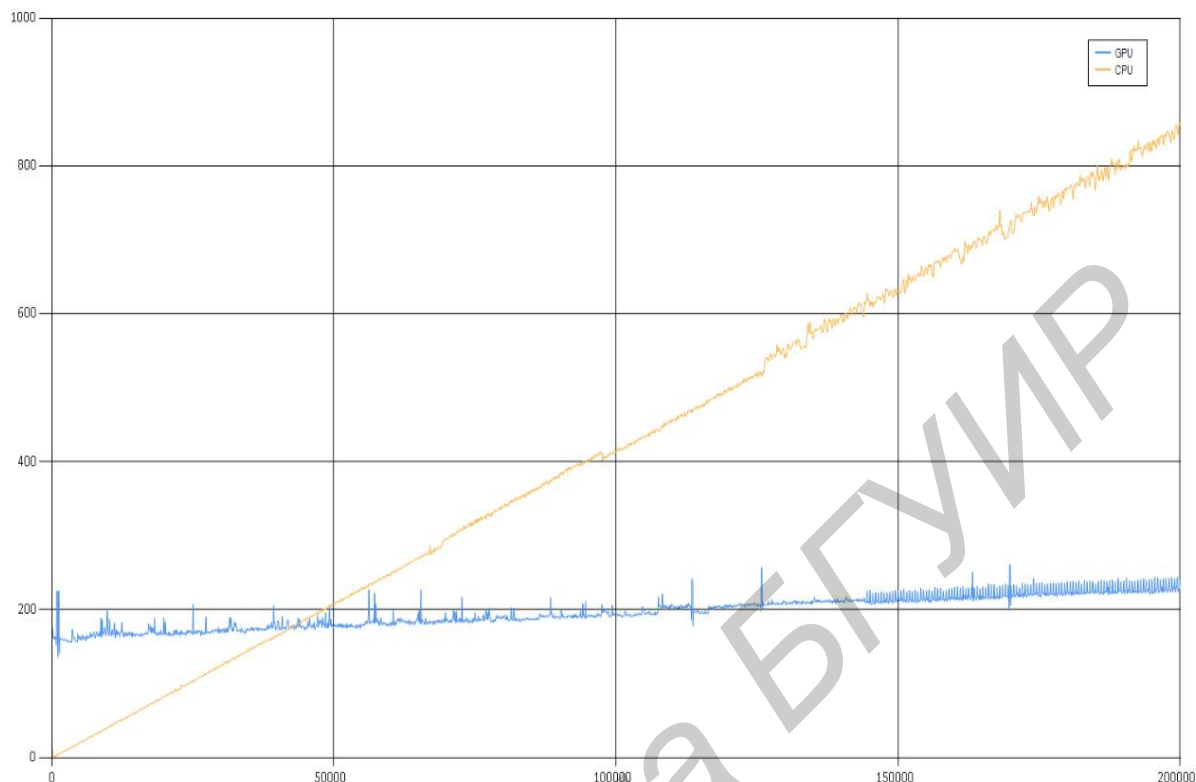


Рисунок 2 – Сравнение времени вычисления цены опциона используя алгоритм Монте-Карло на GPU и CPU

Как видно из графиков, на небольшом количестве путей реализация алгоритма для центрального процессора оказывается во много раз быстрее реализации на графическом адаптере. Однако с увеличением числа разница во времени уменьшается и уже на 50000-и тысячах путей вычисление на OpenCL оказываются эффективнее, а на двухстах тысячах путей время, занимаемое центральным процессором, становится больше практически в четыре раза.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Рассмотрены основные направления развития параллельных вычислительных систем, а также возможность применения гетерогенных

вычислительных систем в ресурсоемких вычислениях, таких как моделирования в финансовых системах.

2. Была разработана реализация алгоритма Монте-Карло, адаптированная для выполнения на гетерогенной вычислительной системе с использованием программно-аппаратной платформы OpenCL.

3. Предложен и реализован метод проверки корректности работы алгоритма моделирования. Для проверки правильности и точности расчета использовалась аналитическое значение, полученное с помощью формулы Блэка-Шоулза. Анализ результатов расчета с различным количеством путей симуляции доказывает теоретическое утверждение о том, что сходимость результатов, полученных путем моделирования цены, к результатам, полученным аналитические растёт с увеличением числа симуляций.

4. Было доказано теоретическое предположение о эффективности распараллеливания алгоритма Монте-Карло с применением гетерогенных вычислительных систем, в частности систем, использующих графический процессор. При небольшом числе путей симуляции центральный процессор вычисляет стоимость опциона быстрее чем графический. Однако с ростом количества шагов моделирования, время расчета на CPU приближается к времени расчета с использованием видеокарты. И при достаточно большом числе симуляций использование графического процессора становится оправданным, поскольку позволяют получить в несколько раз более эффективный расчет.

Сопоставив анализ эффективности алгоритмов и сходимости результатов в зависимости от количества реализаций случайной цены актива, можно сделать вывод о том, что при необходимости получить более точный результат использование графического процессора является оправданным, поскольку позволяет вычислить стоимость опциона со значительно меньшими временными затратами. Эксперименты на конкретном оборудовании показали, что использование GPU является целесообразным при числе путей моделирования превышающим пятьдесят тысяч.

5. Были предложены методы оптимизации расчетов, когда высокая точность вычислений является необязательной. В таком случае, фреймворк предоставляет оптимизированные математические функции, позволяющие еще больше сократить время вычислений, за счет использования реализованных производителями оборудования встроенных функций. Таких как например функции вычисления синуса, косинуса, а также некоторых

математических выражений. Конечно их использование влияет на точность расчетов, однако это может быть допустимо ввиду достаточно мягких требований к точности.

Рекомендации по практическому использованию результатов

1. Полученные результаты формируют теоретическую и практическую базу для разработки ПО компьютерных систем для решения задач моделирования процессов принятия решений в финансовых системах.

2. Разработанные алгоритмы могут быть применены в системах моделирования рисков в трейдинговых системах, для получения систем, способных моделировать необходимые величины в реальном времени.

3. Результаты работы могут использоваться в дальнейших исследованиях, связанных с ускорением финансового моделирования. В частности, перспективным направлением в решении сложных вычислительных задач с использованием платформы OpenCL, является запуск вычислений на FPGA. Это позволяет компилировать код, написанный на OpenCL в VHDL или Verilog описание архитектуры, которое затем используется для прошивки микросхемы. Это даёт возможность программисту использовать FPGA как ускоритель высокопроизводительных вычислений без знания HDL-языков. При этом OpenCL код будет скомпилирован в описание такой архитектуры, которая позволит максимально быстро и с наименьшими энергозатратами решить поставленную задачу. Это достигается благодаря тому, что архитектура, полученная таким образом, может содержать простейшие АЛУ, реализующие только необходимые функции. Однако количество ядер с такими АЛУ может быть достаточно большим и как следствие может быть достигнута достаточно высокая степень параллелизма.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Мелещя, Д.В. Расчет стоимости европейского опциона с использованием графического процессора / Д.В. Мелещя // Научная дискуссия: сборник статей по материалам 39-й научно-практической конференции. – Москва, 2015. – №7(38). – с. 43-46.

2. Мелещня, Д.В. Применение графического процессора для оценки производственных финансовых инструментов методом Монте-Карло / Д.В. Мелещня // WebConf2015: материалы 3-й международной научно-практической конференции. – Минск, 2015. – с. 98-99.

3. Мелещня, Д.В. Расчет стоимости европейского опциона с использованием программно-аппаратной платформы CUDA / Д.В. Мелещня // Информационные технологии и системы: материалы международной научной конференции. – Минск, 2015. – с. 264-266.

Библиотека БГУИР