УДК 519.8; 368; 65.0+004.421.2

НОРКИН Б.В.

ОБ АКТУАРНЫХ ВЫЧИСЛЕНИЯХ С ИСПОЛЬЗОВАНИЕМ ГРАФИЧЕСКИХ ПРОЦЕССОРОВ

В статье рассматриваются проблемы выполнения актуарных вычислений на графических ускорителях. Оценка вероятности разорения и других показателей функционирования страховой компании может быть проведена методом статистического моделирования Монте- Карло. Во многих случаях это единственно применимый метод. В связи с тем, что вероятность разорения достаточно мала, для достижения приемлемой точности может потребоваться астрономическое число испытаний. Распараллеливаемость метода Монте Карло делает возможным проведение расчетов на графическом ускоритель. Это позволяет получить достаточно точный результат за разумное время. В работе представлены результаты численных экспериментов на разработанной системе актуарного моделирования RMS 0.1, использующей GPU с архитектурой NVIDIA CUDA 4.2.

The paper considers issues of performing actuarial calculations on graphical processing units (GPU). Ruin probability estimation for an insurance company can always be made by means of Monte-Carlo method. In some cases it is the only applicable method. Smallness of ruin probability may require astronomical number of Monte-Carlo simulations to achieve an acceptable accuracy. Effective parallelization of the Monte Carlo method and utilizing calculation power of modern GPU makes it possible to obtain good accuracy within reasonable real time limits. A software system RMS 0.1 utilizing NVIDIA CUDA 4.2 architecture for modeling reserve evolution and performance of an insurance company was developed. Results of numerical experiments by means of this system are presented.

Введение[1]

Суть страхового бизнеса состоит в получении максимума чистой прибыли при достаточных страховых резервах для покрытия страховых требований. При этом важно соблюсти тонкий баланс между прибыльностью и рисками. В качестве меры риска может выступать вероятность разорения или неплатежеспособности компании, а прибыльность можно оценивать по распределению собранных дивидендов и остатка капитала в конце планового периода времени.

Для формального описания деятельности страховой компании часто используется классический случайный процесс риска (модель Крамера-Лундберга), моделирующий стохастическую эволюцию капитала страховой компании [1]. В этой модели, с одной стороны капитал монотонно и линейно возрастает с течением времени за счет непрерывно поступающих премий, а с другой стороны, в случайные моменты времени (прихода страховых требований) он убывает на случайную величину (требования). Компания разоряется, если капитал становится меньше нуля.

Очевидно, что данный процесс не отражает многих аспектов деятельности страховой компании, например, перестрахование, инвестиции, выплату дивидендов и др. Поэтому для моделирования реальной страховой компании используют стохастические имитационные модели динамического финансового анализа [2], которые учитывают влияние многих управляющих и случайных факторов на динамику резервов компании. В качестве управляющих параметров могут выступать доли отчислений в страховой резерв, на инвестиционную деятельность, на дивиденды, на перестрахование, на текущие расходы и другие стратегические параметры управления. В качестве важных случайных факторов выступают уровень страховых выплат (по отношению к страховым премиям), доходность инвестиций, уровень инфляции, уровень выплат по договорам перестрахования и др. При этом результаты моделирования при заданных значениях управляющих параметров не являются однозначными, а есть случайные величины. Однократное моделирование траектории изменения показателей функционирования компании во времени не дает полного представления о качестве управления. Необходимо промоделировать большое число траекторий, чтобы получить представление о

[1] Данная статья является расширенной версий доклада [1] на международной конференции «Высокопродуктивные вычисления» HPC-UA 2012, 8-9 октября 2012 г., Киев, Украина.

вероятностном распределении результатов. Тогда можно вычислить любые детерминированные характеристики выбранной стратегии управления: вероятность разорения, ожидаемые дивиденды, ожидаемое значения остаточного резерва и др. Однако проблема состоит в том, что для оценки этих показателей, например, вероятности разорения, может потребоваться астрономическое число симуляций, которое невозможно выполнить за разумное время даже на современных компьютерах. Эта проблема осознана в страховой математике, поэтому, например, для классического процесса риска получены разнообразные формулы приближенной оценки вероятности разорения [1]. Но эти формулы не работают для более общих, а тем более реальных моделей страховой деятельности.

Универсальным, а зачастую, единственным методом моделирования сложных динамических стохастических систем является метод статистического имитационного моделирования Монте-Карло. Этот метод требует астрономического числа симуляций для достижения достаточной точности, поскольку вероятность разорения является малой величиной. Однако, отличная распараллеливаемость метода делает возможным решить эту проблему за счет наращивания вычислительной мощности. В работах автора [4 – 7] параллельная версия метода Монте-Карло наряду с параллельным методом последовательных приближений, реализованные на кластере из нескольких персональных компьютеров, применялись для нахождения вероятности разорения как функции от начального капитала и других параметров управления страховой компанией. В [8] обсуждаются близкие вопросы параллельного моделирования случайных блужданий на графических ускорителях с архитектурой NVIDIA CUDA [9]. В данной работе метод Монте-Карло для решения актуарных задач реализован на графическом ускорителе с архитектурой NVIDIA CUDA 4.2 [9], за счет чего время расчетов сократилось на один-два порядка по сравнению с расчетами на центральном процессоре. Более того, скорость вычислений делает возможным изучать зависимости вероятности разорения и других показателей не только от начального капитала, а от нескольких других параметров работы компании.

Обобщенный процесс риска как модель страховой компании

Страховая компания обязана поддерживать некоторый уровень страховых резервов для покрытия текущих случайных страховых требований. Математическая модель стохастической эволюции резервов x_t страховой компании имеет вид:

$$x_t = u + \int_0^t \left(c - d(x_s)\right) ds - S_t, \quad 0 \le t \le T, \qquad (1)$$

где t – временной параметр $\left(0 \le t \le T\right)$; $x_0 = u \ge 0$ – начальный капитал (страховой резерв); $S_t = \sum_{k=1}^{N_t} z_k$ – агрегированные случайные страховые требования; z_k – случайные требования в моменты t_k с функцией распределения $\overline{F}_{t_k}(\cdot)$; N_t – число поступивших к моменту t случайных требований; c – агрегированная страховая премия в единицу времени; $d(\cdot)$ – некоторая функция, выражающая интенсивность выплат дивидендов и других отчислений в зависимости от текущих резервов, $0 \le d(\cdot) \le c$. Здесь, функция $d(\cdot)$ имеет смысл позиционного управления дивидендами. Например,

$$d(x) = \begin{cases} a\left(x - b(x)\right), & x \ge b(x), \\ 0, & x < b(x), \end{cases}$$

где $b(\cdot)$ – некоторая монотонно возрастающая функция, называемая дивидендным барьером, a – доля капитала, отчисляемая на непроизводственные нужды. В классической модели Крамера-Лундберга [1] (с вычитанием постоянных дивидендов d, $0 \le d \le c$) уравнение эволюции капитала имеет вид:

$$x_t = u + (c - d)t - \sum_{k=1}^{N_t} z_k, \quad 0 \le t \le T, \qquad (2)$$

где $\{z_k, k = 1, 2, ...\}$ – независимые одинаково распределенные наблюдения случайной величины требований с общей функцией распределения $F(\cdot)$ и средним μ, N_t – целочисленная случайная величина с распределением Пуассона с параметром α (временная интенсивность прихода требований в экспоненциальном распределении).

На практике финансовое состояние компании регистрируется в дискретные моменты времени, например, поквартально. В этом случае математическая модель стохастической эволюции резервов x_t страховой компании имеет вид:

$$x_{t+1} = x_t + c_t - x_t - d_t = \\ = u + \sum_{i=1}^{t} c_i - \sum_{i=1}^{t} x_i - \sum_{i=1}^{t} d_i, \qquad (3)$$

где $t = 0, 1, \ldots, T - 1$ – дискретный временной параметр; $x_0 = u \geq 0$ – начальный капитал (резерв); x_t – текущий страховой резерв в момент t; c_t, d_t, x_t – соответственно суммарные квартальные премии, дивидендные выплаты и случайные страховые требования за период t. Распределение величин x_t находится из страховой статистики [9].

Рассмотрим вероятность разорения $\psi(\cdot) = \Pr\left\{\inf_{0\le t\le T} x_t < 0\right\}$ как функцию от параметров процесса. Эта вероятность может быть использована как мера риска при управлении страховой компанией. Например, вероятность разорения (на бесконечном интервале времени) в классической модели страховой компании (2) с экспоненциальным распределением требований имеет вид [2]:

$$\psi(u,c,d,\alpha,\mu) = \begin{cases} \dfrac{1}{1+\rho}\exp\left(-\dfrac{\rho u}{(1+\rho)\mu}\right), & \rho > 0, \\ 1, & \rho \le 0, \end{cases} \qquad (4)$$

где $\rho = (c-d)/(\alpha\mu) - 1$. В данном случае зависимость $\psi(u,c,d,\alpha,\mu)$ известна в явном виде. В более общей модели (1) такая зависимость не известна и может быть получена только методом Монте-Карло. Формула (4) в данной работе используется для тестирования и настройки метода Монте-Карло. Кроме вероятности разорения нас может интересовать распределение капитала в некоторый момент времени и собранные дивиденды, их средние значения и дисперсии в этот момент, а также зависимость этих величин от разнообразных параметров.

Параллельный метод статистического моделирования (метод Монте-Карло)

Метод состоит в независимом параллельном моделировании большого числа N траекторий стохастической эволюции резервов x_t страховой компании на заданном интервале времени $[0,T]$ для заданного набора параметров (u,c,d,α,μ) и вычислении доли $p_N(t)$ неразорившихся траекторий к моменту времени t, а также среднего чистого дохода (собранных дивидендов) (38)

$$D_N(T) = (1/N)\sum_{n=1}^{N}\int_0^{\tau_n} d(x_s^n)ds,$$

где $\left\{x_s^n, 0 \le s \le \tau_n\right\}$ – траектория процесса риска (1), (2) или (3) в n-ом испытании, τ_n – момент разорения или $\tau_n = T$, если разорения до момента T в n-ом испытании не произошло. В процессе параллельного моделирования вычислительные ядра не общаются, а по завершении моделирования информация о траекториях собирается на одном ядре и строятся функции $p_N(u,c,d,\alpha,\mu,T)$ и $D_N(u,c,d,\alpha,\mu,T)$ как функции того или иного параметра. Результаты моделирования отображаются в плоскости «изменяемый параметр – вероятность разорения» и в пространстве «риск – доход», т.е. в плоскости «вероятность разорения – собранные дивиденды». Точность метода Монте-Карло может быть оценена с помощью экспоненциального неравенства Хефдинга:

$$\Pr\left\{\left|p_N(u,c,d,\alpha,\mu,T) - \psi(u,c,d,\alpha,\mu,T)\right| \ge \delta\right\} \le 2e^{-N\delta^2/2},$$

откуда (10^{-k})–доверительная граница для $\left|p_N(u,t) - \varphi(u,t)\right|$ имеет вид

$$\delta_k(N) = \sqrt{2\left(k\ln 10 + \ln 2\right)}\Big/\sqrt{N}.$$

Программная реализация параллельного метода Монте-Карло на GPU

Для проведения расчетов была создана программная система страхового моделирования “RMS 0.2”. Графический интерфейс системы (Рис.1, 2) позволяет изучать зависимость функции полезности (к примеру, величины собранных дивидендов или остаточного резерва) от различных параметров работы компании. Можно также строить зависимость вероятности разорения (используемой в качестве меры риска) от вышеуказанных параметров. Для написания интерфейса был использован язык C#, расчётная часть написана на CUDA C.

В системе реализованы следующие функции:

выбор модели (классическая или модель с дискретным временем);

загрузка статистических данных и стандартный анализ данных;

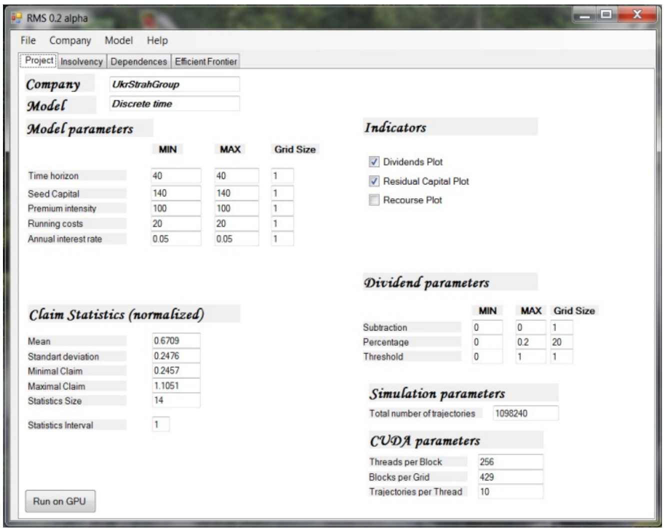

Рис.1. Окно задания параметров в RMS 0.1

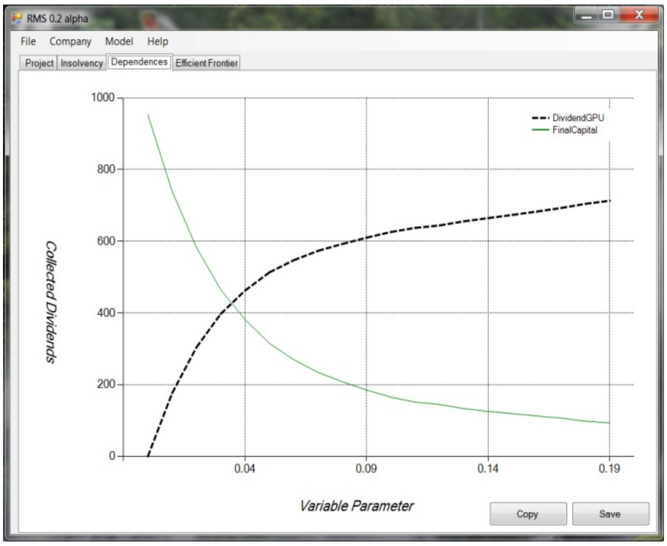

Рис.2. Отображение результатов моделирования

задание интервалов и сетки для изменяемых параметров модели;

задание параметра статистического моделирования (числа симуляций);

задание параметров для распараллеливания вычислений (распределение нагрузки между вычислительными ядрами);

сохранение и загрузку проекта (данных и параметров модели);

построение и графическое представление зависимостей вероятности разорения, суммарных дивидендов и остаточного резерва от любого исследуемого параметра;

отображение результатов моделирования (эффективной границы) в плоскости «вероятность разорения – суммарные дивиденды»;

вывод результатов в память и на печать.

Система работает на реальных статистических данных для конкретных компаний, полученных путем обработки информации из украинских журналов "Страхова справа" и "Insurance TOP". Пример таких данных приведен на рис. 3.

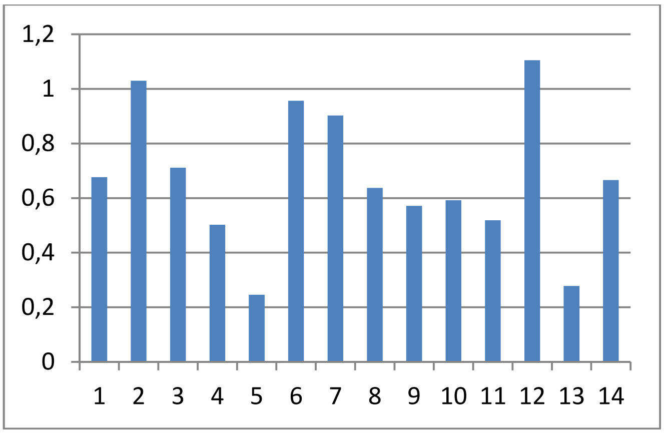

Рис. 3. Эмпирическое распределение нормированных квартальных требований

Для реализации параллельного метода Монте-Карло ключевое значение имеет проблема параллельного генерирования большого числа длинных независимых числовых последовательностей, состоящих из независимых случайных чисел. В CUDA эта проблема решается с помощью библиотеки генерирования псевдослучайных числе CUDA CURAND [8]. В данной системе каждая «нить» (thread) в цикле генерировала от 10 до 1000 траекторий, 256 нитей объединялись в блоки. Проводились эксперименты с одинарной и двойной точностью представления числовых данных.

Результаты численных экспериментов

Была рассмотрена задача нахождения вероятности разорения и чистого дохода страховой компании на конечном интервале времени $[0,T]$. Для проведения тестовых численных экспериментов использовалась классическая модель Крамера-Лундберга (2) с экспоненциально распределенными величинами требований и ее временная дискретизация (3), так как для вероятности разорения $\psi(u,c,d,\alpha,\mu)$ модель (2) допускает простое аналитическое решение (4). Численные эксперименты производились на конфигурации AMD Athlon 64 3200+ 1.5Gb Ram, графический ускоритель – Nvidia

GeForce GTX 560 2Gb с использованием технологии NVIDIA CUDA 4.2.

1. Ускорение вычислений на GPU. Было произведено сравнение времени оценки вероятности разорения для процесса (2) с параметрами $c=1$, $d=0$, $\alpha=\mu=1$, $u=1$, $T=100$ на CPU и GPU. В эксперименте наблюдалось ускорение вычисления вероятности разорения с помощью GPU на один-два порядка.

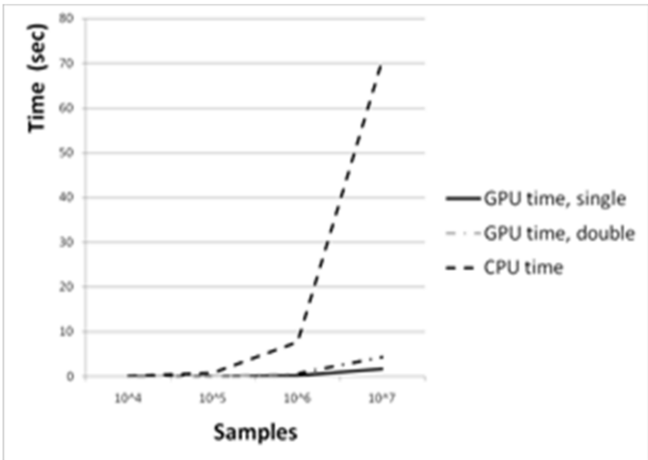

Рис. 4. Сравнение скорости вычислений

2. Точность и время оценки малых вероятностей разорения на GPU. Проводилось также сравнение оценки вероятности разорения процесса (2) для $T=100$ с ее точным значением при $T=\infty$, полученным по формуле (4), для параметров $c=2$, $\alpha=\mu=1$, $u=10$.

Табл. 1. Точность и время вычисления вероятности разорения

Число траекторий	10^6	10^7
Время (в миллисекундах) одинарная точность	199	1959
Время (в миллисекундах) двойная точность	581	5019
Относительная ошибка, одинарная точность %	3.35	0.654401
Относительная ошибка, двойная точность %	1.03	0.322772
Теоретическая вероятность разорения	0.003369	

Этот эксперимент показывает, что с помощью GPU можно с высокой точностью оценивать малые вероятности разорения в реальном времени.

3. Исследование зависимости вероятностей разорения и чистого дохода от параметров с помощью GPU. Разработання система страхового моделирования RMS 0.1 позволяет в реальном времени исследовать зависимость вероятности разорения и чистого дохода (суммарных дивидендов) страховой компании как функцию от любого параметра модели. Для этого достаточно в окне интерфейса системы задать минимальное и максимальное значение изменяемого парамтера, а также число промежуточных значений параметра. Сокращение времени вычислений достигается за счет массивного распараллеливания метода статистического моделирования с помощью GPU.

На Рис. 5 показана (пунктиром) зависимость вероятности разорения $\psi(u,c,d,\alpha,\mu,T)$ в модели (2) от интенсивности прихода требований $\mu\in[0.5,2.2]$. Сплошная линия отображает точное значение вероятности разорения при $T=\infty$, полученное по формуле (3). Остальные параметры принимали следующие значения $u=10$, $c=2$, $d=0$, $\alpha=1$, $T=100$. Несовпадение оценки вероятности разорения и ее точного значения объясняется конечностью временного интервала $[0,100]$ в методе статистического моделирования.

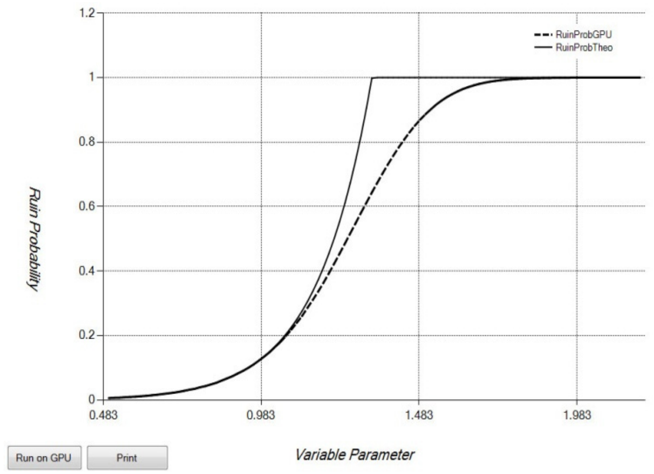

Рис. 5. Зависимость вероятности разорения от интенсивности прихода требований

На рис. 6 – 7 представлены результаты моделирования процесса (1) с дивидендной стратегией

$$d(x)=\begin{cases}0, & x<b,\\ d', & x\ge b,\end{cases}$$

где величина порога отбора дивидендов $b\in[5,50]$ выступает в качестве изменяемого параметра. Остальные параметры модели принимали следующие значения $c=2$, $\alpha=2$, $u=10$, $\mu=1.5$, $T=100$, $d'=1$.

Рис. 8 сопоставляет на одном графике вероятность разорения и собраные дивиденды при изменении параметра b.

Выводы

Использование графических ускорителей позволяет производить численные расчеты в сложных общих актуарных моделях методом Монте Карло, при этом приемлемая относительная точность порядка 1% на вероятностях разорения порядка 10^{-3}

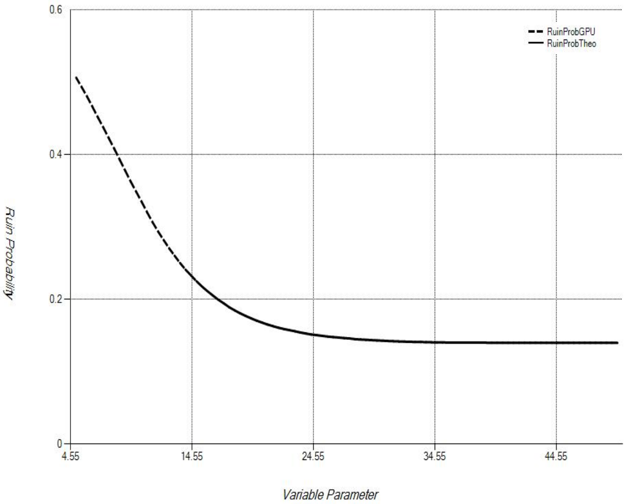

Рис.5. Зависимость вероятности разорения от порога b.

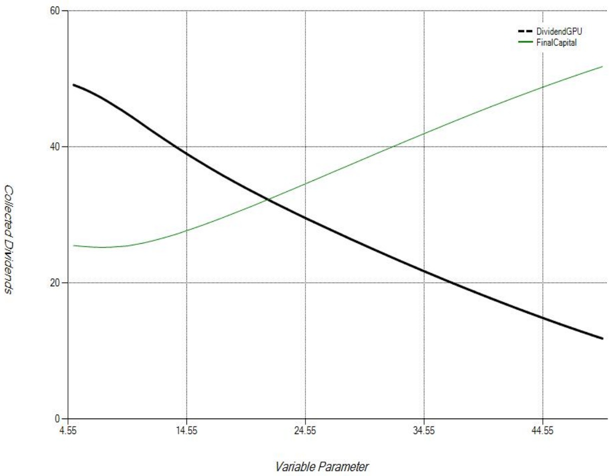

Рис.6. Зависимость величины собранных дивидендов и остаточного капитала от порога b.

достигается практически в режиме реального времени. Следует принимать во внимание, что скорость работы графических ускорителей Nvidia GeForce при решении актуарных задач

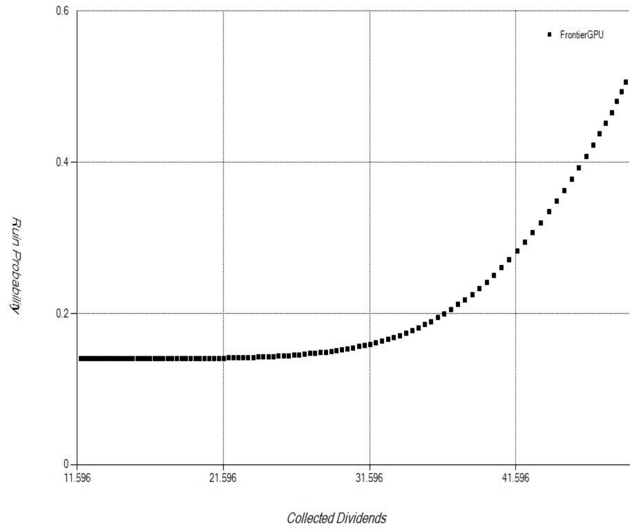

Рис.7. Отображение результатов моделирования в плоскости «доходность-риск»

отличается при использовании операций с числами одинарной и двойной точности (теоретически – в 8 раз), поэтому для каждого конкретного случая необходимо оценивать целесообразность применения расчетов с числами двойной точности.

Разработанная система актуарного моделирования RMS 0.1 позволяет в реальном времени за счет ускорения вычислений на GPU исследовать зависимость вероятности разорения и ожидаемого чистого дохода страховой компании как функцию любого параметра управления компанией. Тем самым она позволяет исследовать влияние факторов управления на функционирование страховой компании. Система позволяет также сопоставлять в реальном времени доходность и риск при выборе параметров управления компанией.

Список литературы

1. Норкин Б.В. Об актуарных вычислениях с использованием графических процессоров // Труды международной конференции «Высокопродуктивные вычисления» (HPC-UA 2012, Украина, Киев, 8-10 октября 2012 г.). – Киев: Национальная академия наук Украины, 2012. – С.268-274.
2. Леоненко М.М., Мішура Ю.С., Пархоменко Я.М., Ядренко М.Й. Теоретико-ймовірнісні та статистичні методи в економетриці та фінансовій математиці. – К.: Інформтехніка, 1995. – 380с.
3. Kaufmann R., Gadmer A., Klett R. Introduction to dynamic financial analysis // ASTIN Bulletin. – 2001. – Vol. 31. – No. 1. – P. 213-249.
4. Норкин Б.В. Математические модели оптимизации страхового дела // Кибернетика и системный анализ. – 2011. – № 1. – С. 128-145.

5. Norkin B. Parallel computations in insurance business optimization //Proceedings of the 1-st International Conference on High Performance Computing. October 12-14, 2011, Kyiv, Ukraine. – P. 33-39.
6. Норкин Б.В. Распараллеливание методов оценки риска банкротства страховой компании // Теорія оптимальних рішень. – Київ : Інститут Кібернетики, 2010. – Стор. 33-39.
7. Норкин Б.В. О вероятности разорения управляемого процесса авторегрессии // Комп’ютерна математика. Ін-т кібернетики ім. В.М. Глушкова. Київ, 2011. – С. 142-150.
8. Haizhen Wu. Parallel Computing Using GPUs. – March 1, 2011. http://ecs.victoria.ac.nz/twiki/pub/EResearch/EcsTeslaResource/Parallel.Computing.Using.Graphics.Cards.pdf
9. Боресков А. В., Харламов А. А., Марковский Н.Д. и др. Параллельные вычисления на GPU. Архитектура и программная модель CUDA. – М.: Издательство Московского университета, 2012. – 336 с.
10. Залетов А.И. Страхование в Украине. – Киев: Международная агенция «BeeZone», 2002. – 452 с.
11. NVIDIA, CUDA CURAND Library, 2010.