

Аннотация

Макаров, В.Л. Моделирование социально-экономических процессов с использованием суперкомпьютерных технологий [Текст] : монография / В.Л. Макаров, А.Р. Бахтизин, Е.Д. Сушко. – Вологда : ИСЭРТ РАН, 2016. – 183 с.

В последние годы в число приоритетных направлений в информационных технологиях уверенно вошли отдельные отрасли гуманитарных наук, в частности, создание систем краткосрочного и долгосрочного предсказательного моделирования социальных явлений и событий с использованием суперкомпьютерных технологий.

По оценкам аналитической компании IDC (International Data Corporation), общий объем данных к 2020 г. достигнет 35 000 экзабайт, что означает 29-кратный рост за 10 лет (1 200 экзабайт в 2010 г.). Как полагают аналитики, наибольший вклад в этот огромный прирост внесут следующие сферы: интернет (блоги, социальные сети и т.д.); финансы (биржевые индексы, аналитические материалы и т.д.); здравоохранение (данные о пациентах, лекарствах, способах лечения и т.д.); астрономия (детализированные изображения галактик); биоинформатика (данные о 3,3 млрд. оснований нуклеиновых кислот в геноме человека, протеиновые последовательности и их анализ); библиотеки (текстовые данные, фото, карты и т.д.). Кроме того, следует учесть быстрое развитие электронного правительства, где создается единая база данных обо всех гражданах страны с подробной информацией о работе, о собственности, о доходах, различных платах и тому подобное.

Необходимость обработки такого объема данных обуславливает появление аналитических систем нового поколения, включающих усовершенствованные методы вычислений, распознавания образов, организации хранилищ, сбора статистики с целью извлечения смысла из данных и получения информационного контекста. Ставка делается в том числе и на агент-ориентированные модели (АОМ), относящиеся к классу моделей, основанных на индивидуальном поведении агентов и создаваемых для компьютерных симуляций. Основная идея, реализуемая в АОМ, заключается в построении вычислительного инструмента, представляющего собой совокупность агентов с определенным набором свойств и позволяющего проводить симуляции реальных явлений.

С помощью АОМ можно смоделировать систему, максимально приближенную к реальности. Появление АОМ следует трактовать как результат эволюции методологии моделирования: переход от мономоделей (одна модель – один алгоритм) к мультимоделям (одна модель – множество независимых алгоритмов). При этом агент в АОМ является автономной сущностью, имеющей, как правило, графическое представление, с определенной целью функционирования и возможностью обучения в процессе существования до уровня, определяемого разработчиками соответствующей модели. В качестве примеров агентов могут рассматриваться: 1) люди (или иные живые организмы), роботы, автомобили и другие подвижные объекты; 2) недвижимые объекты, а

также 3) совокупности однотипных объектов. В целом агентами в АОМ могут быть любые наблюдаемые в реальной жизни объекты, однако основной задачей учета этих объектов в рамках модели является их корректная спецификация. Отметим, что общая особенность и одновременно главное отличие всех АОМ от моделей других классов – наличие в них большого числа взаимодействующих друг с другом агентов.

В силу фактически экспоненциального роста данных обозначается дальнейший тренд развития АОМ – построение АОМ с использованием суперкомпьютерных технологий (в том числе на базе геоинформационных систем). Это направление активно развивается, а на мировых конгрессах, посвященных АОМ, оно уже давно обсуждается не только на специализированных сессиях, но и в рамках пленарных выступлений.

Актуальность использования суперкомпьютерных технологий для разработки АОМ обусловлена тем, что оперативная память обычного персонального компьютера не способна вместить то количество объектов программной среды, которое соответствует, например, населению земного шара или даже отдельных густонаселенных стран. Запуск оригинальной модели в специализированных средах для разработки АОМ с количеством агентов, превышающим несколько миллионов, уже приводит к исчерпанию оперативной памяти персонального компьютера.

Подобным же образом дела обстоят и с производительностью. Для пересчета состояния масштабной системы с нетривиальной логикой поведения и взаимодействия агентов требуются значительные вычислительные ресурсы, сопоставимые с потребностями вычислительных методов математической физики с аналогичным количеством расчетных ячеек. Но, в отличие от последних, поведение агентов включает элементы случайности, поэтому требуется провести целую серию расчетов и найти вероятностное распределение ключевых характеристик итогового состояния моделируемой среды.

Названные факторы обуславливают необходимость масштабных экспериментов с использованием суперкомпьютерных версий моделей, в которых популяция агентов распределяется по множеству узлов суперкомпьютера и расчеты выполняются параллельно. При этом возникает задача адаптации разрабатываемых в традиционных программных средах моделей для суперкомпьютеров.

Как и при создании суперкомпьютерных программ для решения многих физических задач, потенциал для распараллеливания многоагентных систем кроется в использовании локальности межагентного взаимодействия. В модели, как и в реальной жизни, большая часть взаимодействий происходит между субъектами, находящимися неподалеку. Это позволяет использовать подход распараллеливания «по пространству», то есть разместить популяцию агентов на узлах суперкомпьютера наиболее равномерно и с учетом близости их географического положения. Таким образом, разбиение территории, на которой проживают агенты, на так называемые кварталы, обеспечивает базовую возможность для распараллеливания задачи. Это наиболее часто встречающийся

на практике подход для случаев, когда взаимодействие элементов моделируемой системы, будь то агенты в многоагентных системах или отдельные расчетные ячейки с усредненными параметрами моделируемой физической среды, удовлетворяет принципу пространственной локализации: связи и обмен данными имеют место преимущественно для элементов с близкими координатами и обрабатываются практически мгновенно в пределах каждого вычислительного узла. При этом пространство надо понимать в широком смысле. Расстояние между агентами может измеряться не только расположением на карте, но и связями на графе или параметрами принадлежности к той или иной группе.

Монография организована следующим образом. В первой главе дан обзор и анализ зарубежных разработок в области построения АОМ с использованием суперкомпьютерных технологий, преимущественно тех, которые направлены на моделирование социальных систем и процессов, связанных с их функционированием. Вторая глава посвящена описанию разработанного к настоящему моменту специализированного программного обеспечения для технической реализации агентных моделей на суперкомпьютерах. В третьей главе рассмотрены наиболее интересные международные проекты, в рамках которых разрабатываются высокопроизводительные агентные модели, а в четвертой – вопросы использования графических процессоров для запуска АОМ.

И, наконец, в пятой главе дано описание построенных в ЦЭМИ РАН двух мультиагентных демографических моделей, различающихся уровнем детализации при имитации репродуктивного поведения людей. Далее, с использованием разработанной авторами технологии поддержки АОМ для суперкомпьютеров – STARS (Supercomputer Technology for Agent-oriented Simulation), проанализированы этапы и методы эффективного отображения счетного ядра мультиагентной системы на архитектуру современного суперкомпьютера.