

ИНСТРУМЕНТАЛЬНАЯ СРЕДА МОДЕЛИРОВАНИЯ МЕТОДОМ МОЛЕКУЛЯРНОЙ ДИНАМИКИ

Буза М.К., Зимянин Л.Ф.
Беларусь, г. Минск

Введение.

Компьютерное моделирование методом молекулярной динамики (МД - моделирование) является одним из мощных и универсальных методов исследования в современной физике и химии и относится к классу «больших» и в определенной мере «хорошо распараллеливаемых» задач [1].

В молекулярной динамике строится численная модель молекулярной системы (твердое, жидкое или газообразное физическое тело). Атомы заменяются «идеальными атомами», то есть материальными точками, которые характеризуются положением в пространстве и скоростью, и взаимодействуют друг с другом в соответствии с межатомным потенциалом. Потенциал позволяет рассчитать силы, действующие на каждый атом со стороны других атомов (как правило, не всех, а находящихся в пределах радиуса взаимодействия). Траектория движения каждого атома определяется законом Ньютона и может быть найдена как результат интегрирования дифференциального уравнения второго порядка.

Таким образом, задав начальные положения и скорости всех атомов, мы однозначно определяем динамику всей системы. Содержательные результаты для рассматриваемой задачи получаются при количестве атомов порядка $N=10^4-10^5$. В связи с этим для решения задачи требуются значительные вычислительные ресурсы.

Одним из подходов по сокращению времени МД-моделирования является распараллеливание алгоритмов. При этом, в качестве базы для проведения моделирования может быть использованы высокоскоростные корпоративные локальные сети, поскольку специфика данной задачи не требует интенсивного сетевого взаимодействия.

В связи с этим необходимо разработать архитектуру инструментальной среды для проведения MD-моделирования на кластерных системах.

1. Метод молекулярной динамики.

Метод молекулярной динамики основан на гипотезах, обсуждение которых приведено в [2]. Важным свойством MD-методов является их замкнутость. Действительно, варьируя начальными данными, числом частиц, формой пространства, в котором они находятся, а также меняя точность интегрирования системы уравнений движения при постоянном значении полной энергии системы,

можно только средствами численного эксперимента получить представление о качестве того или иного результата. В этом смысле, вопрос о правильности выводов, сделанных по результатам численного эксперимента, всегда сводится к вопросу о правдоподобности гипотез. Замкнутость методов и очевидная простота их оснований по сравнению со статистико-механическими теориями, опisyвающими такую же физическую систему, позволяют использовать эти методы для проверки последних наравне с обычным физическим экспериментом. Ниже рассмотрим математическую модель MD-метода.

Пусть в пространстве задана система из N частиц с координатами $x_i^{(\alpha)}$, где $x_i^{(\alpha)}$ есть α -компонента вектора x_i положения i -й частицы ($\alpha = 1, 2, 3$, $i = 1, \dots, N$). Уравнение движения системы взаимодействующих частиц может быть представлено как

$$\frac{d^2 x_i^{(\alpha)}}{dt^2} = - \frac{\partial}{\partial x_i^{(\alpha)}} U(x_i^{(\alpha)}, i = \overline{1, N}; \alpha = 1..3) \quad (1)$$

$U(x_i^{(\alpha)}, i = \overline{1, N}; \alpha = 1..3)$ – потенциальная энергия системы.

Часть потенциальной энергии системы, связанная с межмолекулярным взаимодействием, обычно предполагается зависящей от расстояния между парами взаимодействующих молекул, т.е.

$$U(x_i^{(\alpha)}, i = \overline{1, N}; \alpha = 1..3) = \frac{1}{2} \sum_{i,j} \sum_{\alpha=1}^3 U^{\alpha}(|x_i - x_j|) + \sum_i \Phi(x_i) \quad (2)$$

где $U(|x_i - x_j|)$ – потенциал взаимодействия молекул i и j , $\Phi(x_i)$ – потенциал внешнего поля, действующий на i -ю частицу.

Для Леннард-Джонсовского потенциала

$$U(|x_i - x_j|) = 4\varepsilon [(\sigma / r_{ij})^{12} - (\sigma / r_{ij})^6] \quad (3)$$

σ, ε – параметры, $r_{ij} = |x_i - x_j|$.

Решение задачи (1) для системы уравнений движения выполняется с начальными условиями:

$$x_i(t) |_{t=t_0} = x_i(t_0);$$

$$\frac{d}{dt} x_i(t) |_{t=t_0} = v_i(t_0), \quad (4)$$

где $v_i(t)$ – вектор скорости i -частицы ($i=1, \dots, N$).

В результате интегрирования системы $3N$ уравнений движения с граничными условиями получаем траектории движения частиц. Эти траектории частиц представляют значительный интерес. Следует отметить, что для получения траекторий частиц используются численные методы интегрирования, что приводит к необходимости постоянной оценки ошибок расчетов. При проведении

расчетов оказываются важными следующие переменные: шаг интегрирования, общее время интегрирования. Важно, однако, получать в результате численного эксперимента значения величин термодинамических потенциалов и таких параметров, как давление, температура и так далее или хотя бы их оценки. Учитывая это обстоятельство, а также то, что в реальном физическом эксперименте невозможно одновременно измерять координаты и импульсы, в методах молекулярной динамики необходимо получать величины, либо вовсе не зависящие от начальных условий, либо не зависящие от начальных условий, выбранных из некоторой области в фазовом пространстве. Поэтому решение задачи (1) при начальных условиях (4) – значительная серия расчётов при различных начальных условиях.

2. Схемы численных экспериментов.

1. Систему $3N$ дифференциальных уравнений движения второго порядка (1) можно записать в виде системы $6N$ дифференциальных уравнений первого порядка:

$$\begin{aligned} \frac{d}{dt} x_i^{(\alpha)} &= v_i^{(\alpha)}; \\ \frac{d}{dt} v_i^{(\alpha)} &= F_i^{(\alpha)}, F_i^{(\alpha)} = \sum_{j \neq i} \varphi_{ij}^{(\alpha)} + \Phi_i^{(\alpha)}. \end{aligned} \quad (5)$$

Запись (5) полностью эквивалентна (1), но более удобна, так как учитывает скорости v , в явном виде: $v_i^{(\alpha)}$ – α -компонента скорости i в текущий момент времени, $\alpha = 1+3; i=1,2,\dots,N$.

2. Выбор разностной схемы для численного интегрирования (5) является важным для проведения MD -моделирования. Поэтому при разработке инструментальной среды моделирования предполагается ее параметризация, т.е. пользователь имеет возможность выбора схемы из набора предлагаемых либо создает свою собственную. В работе использована простая разностная схема, которая базируется на исследованиях, приведенных в [2].

$$\begin{aligned} x_i^{(\alpha)}(t+\tau) - x_i^{(\alpha)}(t) &= \tau(v_i^{(\alpha)}(t+\tau) - v_i^{(\alpha)}(t))/2, \\ v_i^{(\alpha)}(t+\tau) - v_i^{(\alpha)}(t) &= \tau(F_i^{(\alpha)}(t+\tau) - F_i^{(\alpha)}(t))/2, \end{aligned} \quad (6)$$

где шаг τ – шаг по времени разностной схемы (6); $F_i^{(\alpha)}(t+\tau)$ и $F_i^{(\alpha)}(t)$ – функции от множеств $x_i^{(\alpha)}(t+\tau)$ и $x_i^{(\alpha)}(t)$ соответственно. Система уравнений (5) определяет координаты и скорости частиц в момент $(t+\tau)$, если координаты и скорости этих частиц в момент времени t известны. Она решается итерациями.

3. Наибольших затрат времени в процессе проведения метода требует вычисление межмолекулярных сил. Проведение этого вычисления предполагает два этапа: 1) поиск взаимодействующих частиц; 2) вычисление действующих сил.

Данная процедура параметризуется. В работе рассматриваются варианты поиска взаимодействующих между собой частиц в случае трёхмерной задачи со всеми периодическими граничными условиями, с учетом отражающих стенок и др. Кроме того, в [2] предложены ряд частных решений, которые позволяют повысить быстродействие этих процедур.

4. Проблема выбора начальных данных и шага τ по времени моделирования представляет ответственный и трудоемкий процесс. Перед началом каждого моделирования приходится решать самостоятельную задачу по генерации исходных значений координат и скоростей, удовлетворяющих определённым условиям (могут быть заданы значения полной энергии системы, импульса или момента импульса, плотность атомов и т.д.). Начальные данные могут быть случайными. Часто используют способ, который состоит в том, что $x_i^a(t_0)$, соответствует пространственной решётке, такой, что период этой решётки $l \geq 2l/6\sigma$.

Выбор величины шага τ определяется как физическими требованиями (например, величиной временного масштаба, характеризующего систему), так и требованиями точности вычислений.

3. Распараллеливание алгоритмов метода молекулярной динамики.

Как видно из выше приведенного, для проведения MD-моделирования требуются существенные вычислительные ресурсы. Поэтому одним из основных способов сокращения времени расчетов является организация вычислений на системах параллельного действия, в частности, на кластерных архитектурах. Для этого необходимо схемы численных экспериментов адаптировать под такие системы.

Существует достаточно большое число подходов по распараллеливанию алгоритмов молекулярной динамики [3-5].

В данной работе используется алгоритм пространственной декомпозиции (ПД) вычислительной нагрузки. Физическое пространство моделирования разделяется на трёхмерные ячейки-боксы. Каждый бокс ставится в соответствие одному процессору, который вычисляет силы и обновляет координаты и скорости для всех атомов, находящихся в его боксе на каждом шаге моделирования. Основная сложность данного подхода в том, что атомы могут мигрировать в соседние боксы, кроме того, при вычислении межмолекулярного взаимодействия необходимо учитывать влияние атомов соседних боксов, попадающих в зону их взаимодействия. В работе определение списка соседей реализовано по алгоритму Верлета [1]. Размер и форма каждого бокса зависит от количества частиц N , количества процессоров P и формы области моделирования, которая обычно имеет форму прямоугольного параллелепипеда. Согласно с данными требованиями число процессоров P выбирается так, чтобы форма бокса была ближе к кубической, что это позволит уменьшить затраты на коммуникацию за счёт меньшей площади поверхности боксов, а следовательно и меньшей вероятности перехода атома из одного бокса в другой. Реализация параллельной обработки выполнено с использованием коммуникационной библиотеки *MPI* [6].

4. Программная среда моделирования

Схема работы среды включает следующие этапы:

1. Ввод данных. Для ввода данных имеется пользовательский интерфейс (рис. 1), который позволяет задать начальную расстановку $^0x_i(t_0)$ N атомов и их скоростей $^0v_i(t_0)$; параметры атомов: массу m , радиус действия межатомных сил r , σ и ϵ ; параметры алгоритма (шаг τ , количество шагов Q , точность алгоритма, данные о вычислительных узлах кластера P_i , ($i=1, \dots, P$)). Также выбираются варианты результатов, которые хотелось бы получить.

2. Нормализация данных. На этом этапе выполняется инициализация *MPI*, для равномерной загрузки всех узлов кластера подсчитывается наилучшее число атомов для системы, вычисляется наиболее подходящий шаг и количество шагов моделирования для получения корректных результатов, вычисляется количество боксов B , на которое необходимо будет разбить область моделирования (по умолчанию – единичный куб с центром в точке $(0,0,0)$). Тут же вычисляются наборы из меток боксов $\{B_j\}$ ($j=1, \dots, B$), которые будут привязаны к узлам кластера P_i .

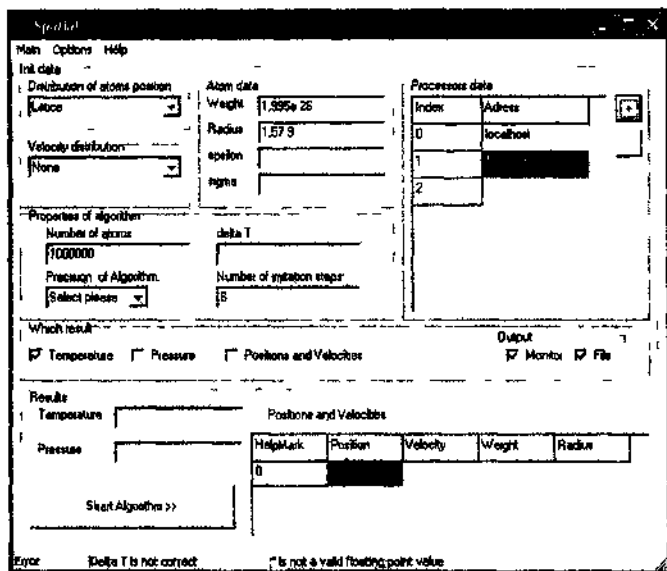


Рис. 1. Пользовательский интерфейс среды MD-моделирования

Меткой является уникальный идентификатор, который однозначно определяет бокс в системе. Заметим, что у атомов так же есть уникальные метки для идентификации внутри системы и внутри бокса, в котором она находится на данный момент. Метки атомов показывают принадлежность к определенному боксу и могут меняться в течение всего алгоритма, а метки идентифицирующие атом в системе постоянны. Кроме того, каждый бокс характеризуется кроме метки ещё углом бокса с минимальными координатами и длинами его рёбер (v

основном они равны, так как в качестве стандартного бокса берётся куб). Есть взаимно однозначное соответствие между данным углом и вспомогательной меткой. Идея меток представлена на рисунке 2.

3. Назначение боксов процессорам. Каждому процессору P_i посылаются входные данные необходимые для работы алгоритма, а так же набор индивидуальных меток для боксов $\{B_j\}^i$, к которым процессор P_i будет привязан в процессе моделирования.

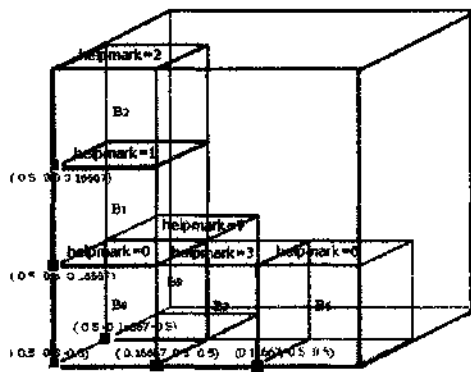


Рис. 2. Вспомогательные метки боксов для случая $P=27$, где символом ■ помечены идентифицирующие углы для боксов

4. Инициализация боксов. В соответствии с пришедшими данными происходит инициализация боксов с метками из множества $\{B_j\}^i$ на каждом процессоре P_i и подготовка к первому шагу.

5. Процесс моделирования. Это основной этап работы среды, который представляет последовательность итераций по пересчёту боксов на данном шаге имитации $n, n=1, \dots, Q$. Каждую итерацию можно разбить на несколько отдельных фаз:

а) актуализация окружения – в процессе этой фазы, каждый процессор P_i синхронизирует данные с другими процессорами $P_j, i \neq j, i, j \in \{1..P\}$, которые привязаны к боксам $\{B_k\}^j, k \in \{1..B\}$, являющиеся в свою очередь соседними для боксов $\{B_j\}^i$. Таким образом, реализуется идея связанных ячеек. В результате мы имеем актуальное состояние бокса $B_k \in \{B_j\}^i$, и набор соседних атомов из соседних боксов на каждом процессоре на момент времени $(n-1)\tau$;

б) пересчёт боксов – на этой фазе на основе актуальных данных полученных на фазе а), происходит пересчёт физических параметров для атомов соответствующих боксов, а точнее позиций атомов $^n x_i^a(n\tau)$ и их скоростей $^n v_i^a(n\tau)$. Заметим, что процессор P_i меняет параметры только “своих” атомов, которые принадлежат назначенным ему боксам $\{B_j\}^i$. Остальные атомы-отражения из “чужих” боксов участвуют только для моделирования общности системы. Причём те атомы, что вышли за границы бокса, формируются в список, который в дальнейшем будет использован на следующей итерации.

б) Вывод результатов. После этапа 5) мы имеем состояние множества блоков $\{B_j\} (Q\tau)$, на момент времени окончания моделирования $Q\tau$, на основании которого вычисляются требуемые характеристики моделируемой системы.

Заключение.

Предложенные в работе инструментальные средства позволяют существенно сократить время МД-моделирования. Теоретические расчеты [5] и проведенное тестирование показывают, что реализуемый параллельный алгоритм пространственной декомпозиции позволяет достичь ускорения вычислений порядка $\sim N/P$. В связи с этим, разработанные инструментальные средства могут найти применение в различных областях физики, химии, биологии, техники для исследования процессов в твердых телах, жидкостях и газах, для прогноза состояния технических систем при различных внешних воздействиях, в микроэлектронике при моделировании технологических процессов проектирования топологии печатных плат.

Литература

1. M.P. Allen. Introduction to Molecular Dynamics Simulation. Computational. Lecture Notes. John von Neumann Institute for Computing, Julich, NIC Series, Vol. 23, 2004 – 28с.
2. Метод молекулярной динамики в физической химии. – М.: Наука, 1996.
3. Daan Frenkel and Berend Smit. Understanding Molecular Simulation: From Algorithms to Application. – San Diego–London–Boston–New York–Sydney–Tokyo–Toronto, – AP, 1998.
4. W.Smith. Molecular dynamics on hypercube parallel computer. // Computer Physics Communications 62 (1991) 229-248.
5. Steve Plimpton. Fast Parallel Algorithms for Short-Range Molecular Dynamics. Journal of Computational Physics vol 117, 1995. – 40 с.
6. Шпаковский Г. И., Серикова Н.В. Программирование для многопроцессорных и систем в стандарте MPI. Мн: Белгосуниверситет, 2002. – 323 с.

СИСТЕМА ВЕКТОРИЗАЦИИ МОНОХРОМНЫХ ТЕХНИЧЕСКИХ ДОКУМЕНТОВ

Л. И. Минченко, М. В. Стержанов
Беларусь, г. Минск, БГУИР

Актуальна задача реализации наиболее эффективных методов векторизации растровых изображений в одной среде, которая является надстройкой над целевой САПР. Имеющиеся на рынке пакеты векторизации классифицируют изображения, каждый класс обрабатывается наиболее подходящим методом. Следует отметить, что трудоемкость используемых в большинстве программных продуктов алгоритмов является сдерживающим фактором к увеличению разрешающей способности сканирования[1].