

СЕРВИС ОБРАБОТКИ ДАННЫХ ПО РАСТВОРИМОСТИ МНОГОКОМПОНЕНТНЫХ ВОДНО-СОЛЕВЫХ СИСТЕМ

М.С. Моисеенков

ВВЕДЕНИЕ

В статье рассматривается вопрос создания проблемно-ориентированного сервиса по обработке данных по растворимости многокомпонентных водно-солевых систем простого эвтонического типа без образования кристаллогидратов и ограниченных рядов твёрдых растворов. Создаваемое клиент-серверное приложение реализует многомерную модель изотермы-изобары растворимости водно-солевых систем рассматриваемого типа, инвариантную по отношению к количеству компонентов. Унифицированный способ хранения и представления экспериментальных данных по растворимости позволяет автоматизировать построение фазовой диаграммы и решать базовые прикладные задачи на её основе.

ВВЕДЕНИЕ В ПРОБЛЕМНУЮ ОБЛАСТЬ

Состояния многокомпонентной физико-химической системы описываются правилом фаз Гиббса, принципами соответствия и непрерывности, предложенными Н.С. Курнаковым. Прежде всего, фазовые диаграммы используются для их анализа с целью определения составов и соотношений равновесных фаз и решения различных прикладных задачи химической технологии. Построение математических моделей позволяет быстро устанавливать природу, состав и соотношение равновесных фаз.

Анализ таких систем осложняется тем, что размерность пространства фазовой диаграммы зависит от количества компонентов водно-солевой системы. Если диаграмма тройной системы, исследованной в изобарно-изотермических условиях, может быть изображена на плоскости, то диаграмма четверной системы является трёхмерным объектом (рис.1). Наглядное представление фазовых диаграмм большей размерности возможно только в виде оптимальных проекций. Помимо увеличения размерности фазового пространства добавление новых компонентов в систему ведёт к усложнению структуры самой диаграммы, увеличению числа фазовых областей, что создаёт дополнительные трудности при обработке и анализе экспериментальных данных по растворимости водно-солевой системы.

Построение изотермы-изобары растворимости водно-солевой систем простого эвтонического типа на плоскости в случае трёх компонентов (две соли S_1 , S_2 и вода H_2O) не вызывает затруднений. Построение фазовой диаграммы четырёхкомпонентной системы $S_1-S_2-S_3-H_2O$ в трёхмерном пространстве требует проведения дополнительных математических выкладок и использования программного обеспечения для работы с трёхмерной графикой. Поскольку в настоящее время не существует специализированного ПО, автоматизирующего построение фазовой диаграммы, то на плечи специалиста-химика ложится вся рутинная работа по обработке экспериментальных данных и их графическому представлению. Переход в четырёх- и более- мерное фазовое пространство значительно увеличивает объём трудозатрат и усложняет анализ системы, поскольку полная визуализация всех данных возможна только через построение оптимальных проекций.

СТРУКТУРА ИЗОТЕРМЫ-ИЗОБАРЫ РАСТВОРИМОСТИ МНОГОКОМПОНЕНТНОЙ ВОДНО-СОЛЕВОЙ СИСТЕМЫ ПРОСТОГО ЭВТОНИЧЕСКОГО ТИПА

Многомерное обобщение математической модели изотермы-изобары растворимости четырёхкомпонентной водно-солевой системы позволяет свести работу со сложно-структурированным многомерным

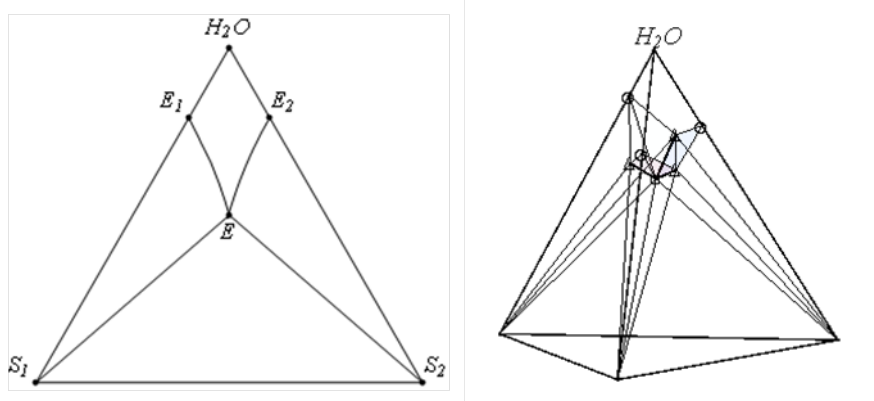


Рис.1. Схема фазовой диаграммы водно-солевой системы простого эвтонического типа для случаев трёх (слева) и четырёх (справа) компонентов

объектом к работе с массивом простых геометрических фигур (симплексов) — объём фазовой диаграммы делится на фазовые области, которые в свою очередь разбиваются на неперекающееся множество симплексов.

Примем следующие обозначения: k — число компонентов водно-солевой системы, d — размерность пространства фазовой диаграммы. Тогда в соответствии с методом Гиббса-Розебома фазовая диаграмма k -компонентной водно-солевой системы изображается в виде d -мерного координатного симплекса, причём справедливо равенство:

$$d = k - 1 \quad (1)$$

Фазовое пространство данной системы ограничено объёмом координатного симплекса и делится на 2^{k-1} фазовые области. Область ненасыщенных растворов, соответствующая растворам, в которых все соли находятся в жидкой фазе, отделена от других фазовых областей границей ликвидуса. Геометрическим образом границы ликвидуса на фазовой диаграмме четырёхкомпонентной системы является поверхность кристаллизации, которая представляется как «мозаика», составленная плоскостями, построенными по трём ближайшим друг к другу экспериментальным точкам, лежащим на этой поверхности. При переходе к системам с большим числом компонентов гиперповерхности моделируются как совокупности гиперплоскостей, построенных по четырём, пяти и т. д. точкам этой гиперповерхности.

Для приближения поверхностей кристаллизации в общем случае предлагается использовать многомерную триангуляцию Делоне, поскольку она максимизирует наименьший угол симплексов разбиения, что позволяет приблизить их к равносторонним и повысить качество интерполяции.

Обозначим за t количество солей, находящихся в твёрдой фазе в данной фазовой области кристаллизации. В силу того, что компоненты системы не образуют новых веществ, возможно выпадение солей в осадок в любом сочетании. Поэтому каждому значению $t = \overline{1; k-1}$ соответствует $C_{k-1}^t = \frac{(k-1)!}{t!(k-t-1)!}$ гиперповерхностей кристаллизации, размерность которых определяется с помощью правила фаз Гиббса с учётом постоянного давления и температуры:

$$v = k - t - 1 \quad (2)$$

Таким образом, для построения фазовой диаграммы k -компонентной водно-солевой системы потребуется построить $\sum_{t=1}^{k-1} C_{k-1}^t$ v -мерных триангуляций Делоне, размерность v которых определяется соотношением (2).

ОЦЕНКА ВЫЧИСЛИТЕЛЬНОЙ СЛОЖНОСТИ АЛГОРИТМОВ

Известно, что v -мерную триангуляцию Делоне множества n точек можно построить за линейное время по $(v+1)$ -мерной выпуклой оболочке (выпуклой оболочкой множества точек S , принадлежащих евклидову пространству E^d , называется граница наименьшей выпуклой области в E^d , которая охватывает S), которая может быть построена за время $O(n^{\lfloor (v+1)/2 \rfloor + 1})$. Тогда, с учётом (2) на построение v -мерной триангуляции Делоне поверхности кристаллизации t солей потребуется время:

$$T_t(k, n) \sim O(C_{k-1}^t n^{\lfloor (k-t)/2 \rfloor + 1}) \quad (3)$$

И тогда общее время, необходимое для построения фазовой диаграммы, равно сумме:

$$T(k, n) = \sum_{t=1}^{k-1} T_t(k, n) \quad (4)$$

Так, для формирования изотермы-изобары растворимости трёхкомпонентной водно-солевой системы простого эвтонического типа потребуется время $T(k=3, n) \sim O(kn^2)$. Увеличение числа компонентов системы повышает вычислительную сложность алгоритма и увеличивает минимально необходимый объём исходных данных — например, в случае девятикомпонентной системы потребуется время порядка $T(k=9, n) \sim O(k^3 n^4)$.

Построенная фазовая диаграмма позволяет решать широкий спектр прикладных задач, среди которых наиболее часто возникает задача определения фазовой области, в которой лежит фигуративная точка, отвечающая исходной смеси компонентов. Концентрации компонентов в смеси определяют координаты фигуративной точки на фазовой диаграмме, попадание которой в ту или иную фазовую область позволяет заключить о том, какие из компонентов системы выпадают в осадок.

В соответствии с предлагаемой многомерной моделью изотермы-изобары растворимости всё пространство фазовой диаграммы разбивается на множество неперекающихся симплексов, благодаря чему задача сводится к поиску того симплекса разбиения, в объёме которого находится исходная фигуративная точка. В процессе перебора симплексов разбиения проводится проверка положения фигуративной точки относительно каждого из них. Каждая проверка потребует время порядка $O(d^3)$. Число таких проверок зависит от количества симплексов разбиения, которое в случае d -мерного пространства оценивается величиной $O(n^{\lfloor d/2 \rfloor})$, где n — число точек, по которым строится d -мерная триангуляция Делоне. Поскольку в соответствии с формулой (2) наибольшую размерность будут иметь гиперповерхности кристаллизации одной

соли ($t = 1$) $v = k - 2$, количество которых составит величину $C_{k-1}^1 = k - 1$, то время, требуемое для решения этой задачи в худшем случае оценивается величиной:

$$T_{ph}(k, n) \sim O(kn^{\lceil (k-2)/2 \rceil} d^3) \sim O(k^4 n^{\lceil (k-2)/2 \rceil}) \quad (5)$$

Оценки (3), (5) показывают, что объём расчётов значительно увеличивается как при росте числа компонентов водно-солевой системы k , так и при увеличении объёма исходных данных n . Точность проводимых расчётов зависит от экспериментальной базы. В настоящее время накоплен достаточно большой объём экспериментальных данных, который тем не менее продолжает расти. Поэтому имеет место тенденция к росту величины n , что ведёт к повышению вычислительной сложности автоматической обработки данных по растворимости.

ИССЛЕДОВАНИЕ ВОПРОСА СОЗДАНИЯ ПРОБЛЕМНО-ОРИЕНТИРОВАННОГО СЕРВИСА

Построенная многомерная математическая модель изотермы-изобары растворимости простой эвтонической водно-солевой системы, унифицируя представление экспериментальных данных для любого числа компонентов, допускает программную реализацию на языке C++ в виде прикладной библиотеки классов, инкапсулирующих универсальные методы представления многомерных сложно-структурированных данных по растворимости и методы их обработки, инвариантные к размерности задачи.

В силу большого объёма экспериментальных данных по растворимости многокомпонентной водно-солевой системы время, необходимое для проведения расчётов, может превысить приемлемое значение. Поэтому возникает спрос на ускорение вычислений за счёт их распараллеливания. Поскольку на практике большинство химиков не владеют навыками параллельного программирования и не всегда есть в наличии параллельная вычислительная система, то актуальным является создание высокопроизводительного проблемно-ориентированного сервиса обработки данных по растворимости. Имеющееся на базе Научно-образовательного центра «Параллельные и распределённые вычисления» (НОЦ ПиРВ) ПГНИУ высокопроизводительное вычислительное оборудование позволяет создавать и поддерживать подобные сервисы.

При создании сервиса высокопроизводительной обработки данных учитывается два уровня параллелизма: параллелизм на уровне методов обработки экспериментальных данных и параллелизм на уровне одновременной работы нескольких пользователей.

Предлагаемый подход представления данных по растворимости, основанный на разбиении сложной структуры на совокупность простых объектов, позволяет достигнуть низкой информационной зависимости между однотипными операциями, выполняемыми над составными частями (симплексами) разбиения:

- При построении фазовой диаграммы каждая из $(2^{k-1} - 1)$ поверхностей насыщения t солей представляется в виде v -мерной триангуляции Делоне, построенной по множествам точек, лежащим на этой поверхности. Организация бесконфликтного доступа к памяти несколькими нитями или процессами позволит сократить время построения фазовой диаграммы в среднем в $(2^{k-1} - 1)$ раз за счёт распараллеливания получения триангуляций Делоне нескольких поверхностей;
- Задача определения фазы произвольной фигуративной точки является переборной задачей, поскольку для определения фазы этой точки необходимо проверить её положение относительно всех симплексов разбиения. При этом на практике часто возникает задача определения фазы сразу для целого набора фигуративных точек. Поскольку между фигуративными точками нет информационной зависимости и прогноз фаз для каждой из них обладает высоким потенциалом распараллеливания, то эта задача может быть заметно ускорена за счёт применения массивных параллельных вычислительных систем, в т.ч. на основе технологии NVIDIA CUDA.
- Расчёт составов равновесных твёрдой и жидкой фаз, на которые распадается состав исходной смеси компонентов], сводится к решению СЛАУ с квадратной матрицей размера $d \times d$. Наличие в исходных данных по растворимости экспериментальной погрешности негативно сказывается на обусловленности матрицы СЛАУ, что снижает качество результатов расчёта из-за высокой вычислительной погрешности. Применение в этом случае методов регуляризации требует дополнительных вычислений, которые также могут быть эффективно распараллелены.

Для обеспечения параллельной работы нескольких пользователей целесообразно использовать высокопроизводительные вычислительные системы, использующие технологии параллельной и распределённой обработки данных, такие как гибридный суперкомпьютер «ПГУ-Тесла», размещённый в НОЦ ПиРВ ПГНИУ.

Интеграция серверной части сервиса в мультиплатформенную систему для научной визуализации позволит снять ограничение на целевую платформу клиентской части и эффективно распределить рендеринг фазовых диаграмм между Клиентом и Сервером. Представленная на рис. 2. архитектура серверной части сервиса состоит из следующих основных компонентов:

- front-end — управляющий узел суперкомпьютера, предоставляющий внешний интерфейс к высокопроизводительной вычислительной системе;
- PBS (Portable Batch System) – система пакетной обработки заданий; на суперкомпьютере «ПГУ-Тесла» используется PBS Torque версии 2.5.7.
- Сервер — приложение, обрабатывающее запросы Клиента и взаимодействующее с Решателем.

- Решатель — приложение, проводящее непосредственные расчёты на вычислительной системе по запросу Сервера.
- Вычислительное ядро — библиотека классов, абстрагирующая понятия проблемной области, и реализующая базовые вычислительные операции, связанные с построением фазовой диаграммы и решением основных прикладных задач.
- БД по растворимости — пополняемая пользователями база экспериментальных данных по растворимости водно-солевых систем.

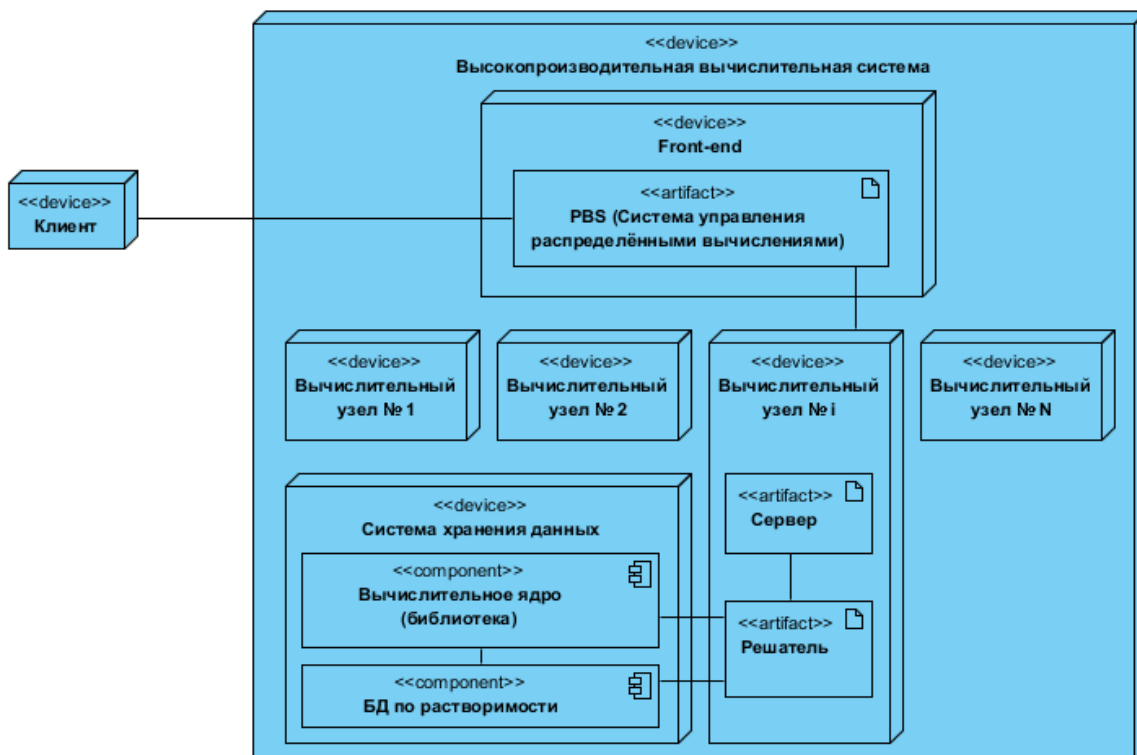


Рис. 2. UML диаграмма развёртывания серверной части сервиса

Предлагаемый принцип работы серверной части сервиса основан на организации удалённого доступа пользователя непосредственно к вычислительным узлам кластера. По запросу Клиента PBS запускает на кластере в качестве задачи Сервер. После того, как PBS зарезервировала за задачей все запрошенные ресурсы (вычислительный узел, оперативная память, графические ускорители, процессорное время) и запустила Сервер на вычислительном узле Клиент получает удалённое управление Сервером, позволяющем проводить расчёты непосредственно на вычислительных узлах суперкомпьютера. Во время исполнения Сервер обрабатывает запросы Клиента и вызывает Решатель, который обращаясь к функциям Вычислительного ядра и данным по растворимости, размещённым на системе хранения данных, проводит вычисления, результат которых передаётся Серверу и затем к Клиенту.

ЗАКЛЮЧЕНИЕ

Обобщение на многомерный случай математической модели изотермы-изобары растворимости трёх- и четырёхкомпонентной водно-солевой системы простого эвтонического типа позволяет создать универсальный и адекватный способ представления многомерных данных по растворимости. Программная реализация многомерной модели в объектно-ориентированной парадигме в виде прикладной библиотеки классов позволяет создать на её основе проблемно-ориентированный сервис, автоматизирующий рутинные и нетривиальные расчёты, связанные с подготовкой и обработкой экспериментальных данных.

Исследование обобщённого многомерного случая приводит к необходимости решения задач, вычислительная сложность которых нелинейно зависит как от объёма исходных данных, так и от размерности фазового пространства. Используемые алгоритмы обладают большим ресурсом распараллеливания, что позволяет значительно сократить время расчётов благодаря применению современных технологий параллельных и распределённых вычислений с использованием суперкомпьютерной техники.

ЛИТЕРАТУРА:

1. Н.С. Курнаков Введение в физико-химический анализ. Л.; М.: Изд-во АН СССР, 1940.

2. М.С. Моисеенков, С.А. Мазунин, В.Л. Чечулин Геометрия фазовых диаграмм четверных водно-солевых систем простого эвтонического типа // Вестн. перм. ун-та. Сер.: Математика. Механика. Информатика / Перм. гос. нац. иссл. ун-т.-Пермь, 2012. №2. С. 36 — 44.
3. С.А. Мазунин Основы физико-химического анализа: учеб. пособие по спецкурсу / Перм. Ун-т. Пермь, 2000. Ч.2. Многокомпонентные водно-солевые системы. 212 с.
4. Ф. Препарата, М. Шеймос Вычислительная геометрия: Введение: Пер. с англ. - М.: Мир, 1989. - 478 с.
5. А.В. Скворцов Триангуляция Делоне и её применение. - Томск: Изд-во Том. Ун-та, 2002. - 128 с.
6. А.Б. Здановский., Е.И. Ляховская, Р.Э. Шлеймович Справочник экспериментальных данных по растворимости много водно-солевых систем. Л.: ГХИ, 1953. Т.1; 1954. Т.2.
7. Обращение прямоугольных матриц. Решение систем линейных алгебраических уравнений методом наименьших квадратов: Метод. Пособие по курсу «Избранные главы линейной алгебры» / Перм. ун-т; Сост. Г.С. Шевцов, О.Г. Крюкова, Б.И. Мызникова — Пермь, 2004. - 91 с.
8. Методы решения некорректных задач. Тихонов А.Н., Арсенин В.Я. М.: Наука. Главная редакция физико-математической литературы, 1979. Изд. 2-е.
9. К.В. Рябинин Разработка мультиплатформенной клиент-серверной системы для научной визуализации // Научно-технический вестник Поволжья. №2 2013 г. - Казань: Научно-технический вестник Поволжья, 2013. - с. 197 — 203.
10. Konstantin Ryabinin, Svetlana Chuprina, Adaptive Scientific Visualization System for Desktop Computers and Mobile Devices // Procedia Computer Science, Volume 18, 2013, Pages 722-731.
11. Texas Advanced Computing Center, The University of Texas at Austin: [Электронный ресурс]. URL: <http://www.tacc.utexas.edu/home> (дата обращения: 01.07.2013).
12. А.Г. Деменев, М.С. Моисеенков, А.В. Поляков Действующий прототип подсистемы удалённой научной визуализации для суперкомпьютера «ПГУ-Тесла» // Электронный многопредметный научный журнал «Университетские исследования», 6 с. [Электронный ресурс]. URL: http://www.uresearch.psu.ru/files/articles/532_55668.doc (дата обращения: 01.07.2013).