

データマイニングによる塩水濃度と水溶液の動的構造に関する研究

(北陸先端大院・知識) ○杉山 歩, 中村 大城, 山口 和宏, 水上 卓,

Dam Hieu Chi, Ho Tu Bao

【序】塩水は自然界に大量に存在する私たちのもっとも身近な水溶液の一つである。例えば、海水は地球上における水分の97%を占めており、また医療や生体化学の面でも重要な溶液となっている。また、工業的な側面からは塩水は摺動部に腐食摩耗が発生し、機器の性能や寿命に悪影響を及ぼすことから腐食摩耗と電気化学的腐食および両者の相互作用に対する研究が進められている。

現在、塩水の研究手法は様々方法が用いられているが、その一つとして分子シミュレーションが挙げられる。分子シミュレーションは方程式を解くことでコンピュータ上に物理特性その特性を示し、解明する手法である。近年の計算機の急速な発達により大規模かつ長時間のシミュレーションを実行することが可能となり、水様液系に至っては数万~の水分子を配した系をマイクロ秒オーダーでの計算も可能となっている。

一方、近年これらの巨大データから統計的手法により帰納的に法則を見つけ出す、データ駆動型アプローチによる研究が複雑系分野を中心に大きなトレンドとなりつつある。そこで、本研究では巨大データの有効的に使用するため帰納法的アプローチにより新しい知識を獲得する手法の開発し、塩水に対するシミュレーションデータから分子ダイナミクスに関する法則を見つけ出すアプローチを採用する。

巨大データの帰納法的アプローチとしてデータマイニングが挙げられる。データマイニングは主にマーケティング等の統計処理で広く用いられており、特にWeb上の巨大な情報が複雑なネットワークの解析手法として成果を上げている。例えば、検索エンジンではビッグデータの解析手法としてデータマイニングを用いて設計されている。これは計算機に多量データから学習させることにより、内在する法則性を導き出す。

このデータマイニングを用いることでシミュレーションにより得られる巨大データを解析することで、従来と異なるアプローチで塩水特性の解明に取り組む。

【計算モデル】本計算はAmber10(amber force field03)を使用し、TIP3P, TIP4PEW水分子モデルを配し、塩水濃度をTable 1に示したNa⁺, Cl⁻イオンの数によりコントロールし、NaCl水様液系及びbulk水系に対し分子動力学計算を行い、ターゲットデータを作成した。解析には先行研究で用いたデータマイニングによるタンパク質水様液系の構造解析手法を採用した[1]。

データマイニングには水分子の座標データの時間経過による変化を追うことで、水分子の振る舞いを定量化し、解析をおこなった。解析方法は以下のプロセスに依る。

1. 水分子の軌道をステップ数で等分する。

Table 1 Water Models

concentration	0.43%	1.64%	4.74%	6.18%	32.30%	40.20%
Total atom	4505	4505	4505	4505	4505	4505
Na+	6	23	68	90	577	774
Cl-	6	23	68	90	577	774
water	4493	4459	4369	4325	3351	2957

2. 等分した軌道の重心を求める。
3. 重心間距離を求める。

等分を細分化する事で一つの分子が 64 等分され、重心間距離が求めるまで繰り返す。ここでそれぞれの分割ごとに得られた重心間距離に対して次の値を求める。

- ・ 平均値
- ・ 標準偏差
- ・ 最大値
- ・ 最小値

上記の計23個の値を用いて水一個の分子挙動を一つの 23 次元のベクトルとして表現し、そのベクトルを主成分分析に正規化したものをプロットする。

【結果】 上記の手法によって計算した各塩分濃度の水様液の全水分子の第一及び第二主成分の解析結果を図 1 に示す。詳細は当日報告する。

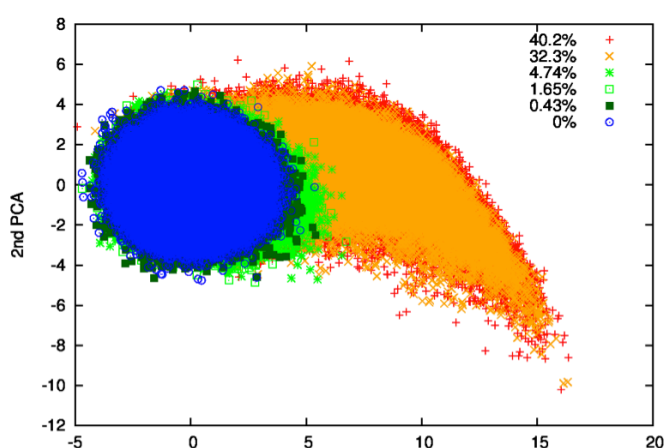


図 1 : Bulk 水と NaCl 溶液中の水分子のダイナミクス分類結果