OUKA
Osaka University Knowledge Archive

Title	多次元データの効率的検索と検索結果の高度活用に関する研究
Author(s)	中川, 忠輔
Citation	大阪大学, 2006, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/46637
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

氏名	中川忠輔（なかがわただすけ）
博士の専攻分野の名称	博士（情報科学）
学位記番号	第　20504　号
学位授与年月日	平成18年3月24日
学位授与の要件	学位規則第4条第1項該当 情報科学研究科マルチメディア工学専攻
学位論文名	**多次元データの効率的検索と検索結果の高度活用に関する研究**
論文審査委員	（主査） 教授　薦田　憲久 （副査） 教授　藤原　融　　教授　岸野　文郎　　教授　西尾章治郎 教授　下條　真司　　神戸大学教授　大川　剛直

論文内容の要旨

現代は情報社会と言われているが、それを支えているのは、社会の様々な局面から発生する色々な情報を収集し、蓄積し、そして必要に応じて検索するための技術であり、コンピュータシステムはその中にあって不可欠の役割を果たしている。科学技術計算や企業などで扱われるデータは、複数の属性項目を持つ多次元データであることが多く、大量で複雑なデータを、人が有効利用するためには、データそのものをうまく整理し、効率良く管理しておく必要がある。コンピュータによってこれらを実現しておけば、例えば、生体情報科学の分野のように膨大なデータを利用する際に、迅速かつ正確な検索を行うことが可能となり、また、特に企業では、多次元データの加工・分析によりユーザの意思決定支援が可能となる。

このような背景を踏まえ、本研究では、多次元データの効率的検索方式と検索結果の高度活用方式を提案する。

生体情報科学の分野においては、蛋白質の機能理解のために、立体構造情報だけでなく分子表面上の物性情報が扱われている。大量かつ複雑なデータ群から類似データを検索するには、データベース内のデータ群と比較対象のデータとの間で比較を行い、類似性を評価する必要があるが、データベース内の全てのデータと比較を行うのは非効率的である。また、比較を高速に行うためには対象データの中でも重要な意味を持つ部分だけを抽出し、比較に掛かる時間を削減する必要がある。これらの課題に対し、本論文では、生体情報科学分野を対象として、多次元データ群からの効率的類似データ検索方式および多次元データの高速類似性評価方式を提案する。

一方、多次元データは企業活動においても経営成果を生み出すための重要な役割を果たしている。ここでは、企業における意思決定支援のための情報活用方法として、時系列データと付加情報を用いたリスク管理支援システムの実現という課題に取り組む。ユーザの代わりにサービスサプライヤが機器を保有し、機器の運用に応じて従量制で使用料を得る機器提供サービスでは、サービスサプライヤの利益がユーザの機器運用方法によって変動する。このため、ユーザとの契約成立が不確定な段階でサービスサプライヤが、適切な見込みユーザの組合せを決定する必要がある。さらに契約不成立の可能性を織り込んでリスク量を見積もり価格に反映する必要がある。これらの課題に対し、本論文では、リスク量に着目した企業データベース活用方法を提案する。

論文審査の結果の要旨

科学技術計算や企業などで扱われるデータは、複数の属性項目を持つ多次元データであることが多く、大量で複雑なデータを人が有効活用するためには、データそのものをうまく整理し、効率良く管理しておく必要がある。生体情報科学の分野では、蛋白質の機能理解のために多次元データの効率的な検索が求められている。また、企業では多次元データを加工・分析することで、より高度な意思決定が為される。本論文は、これらの課題を踏まえて、多次元データの高効率検索技術と検索結果の高度活用技術に関する研究成果を纏めたものである。その主要な成果を要約すると次の通りである。

⑴未知の蛋白質の分子表面データと、データベース内の全てのデータとを比較するのは非効率であるという課題に対し、蛋白質間の類似性評価のために蛋白質の分子表面を構成する頂点を対応付けて分子表面データ間の距離を定義し、予めデータベース内のデータ間の距離を計算して距離空間を構築し、データを効率的に検索する方法を提案している。

⑵分子表面データの検索の効率性を高めるために、蛋白質の分子表面上の物性や形状的特徴の類似した領域を特徴領域として抽出し、特徴領域同士を比較し、分子表面データ間の類似性を大まかかつ高速に評価する方法を提案している。

⑶機器提供サービス事業における高度意思決定のためのリスク管理支援システムの実現という課題に対し、リターンが大きく、かつリスク量が小さくなるユーザの組合せを構成し、組合せによるリスク低減効果を価格に反映する方法を提案し、経済産業省のデータを適用して、契約数とリスクプレミアムの関係を数値化している。

以上のように、本論文は多次元データの効率的検索と検索結果の高度活用において成果を挙げた先駆的研究として、情報科学に寄与するところが大きい。よって本論文は博士（情報科学）の学位論文として価値あるものと認める。