データベースシステム 116－19
（1998. 7. 8）

意味的連想処理機構を用いた大量データ分析のための動的クラスタリング方式

吉田 尚史† 清木 康†† 北川 高嗣†

† 筑波大学 電子・情報工学系
†† 慶應義塾大学 環境情報学部

要旨

本稿では，データの意味的解釈を伴うデータマイニングを行うための基礎となる意味的動的クラスタリング方式を提案する．提案方式の特徴は，データの意味を考慮した文脈に応じた動的クラスタリングを実現する点にある．本方式により，同一解析対象のデータに対して文脈に応じて多数の意味的解析結果を得ることが可能となる．ここでは，実験結果を示し，提案方式の有効性を確認する．

キーワード: 動的クラスタリング，意味的連想処理，データマイニング

A Dynamic Clustering Method with Semantic Associative Processing for Large Data Analysis

Naofumi Yoshida,† Yasushi Kiyoki†† and Takashi Kitagawa†

† Institute of Information Sciences and Electronics,
University of Tsukuba

†† Faculty of Environmental Information, Keio University

abstract

In this paper, we presents a semantic dynamic clustering method as the basis of the data mining with semantic recognition of data. The main feature of the method is to make clustering for raw data semantically according to a given context. By using this method, we can obtain a set of semantic clusters according to a given context from a set of raw data. We clarify the feasibility of the method by several experimental results.

keywods: dynamic clustering, semantic associative processing, data mining

1 はじめに

近年，データベースにおける知識獲得 (KDD: Knowledge Discovery in Databases) の研究が注目されている [1, 9, 11]．KDD の一プロセスであり，その重要な役割を担うデータマイニングに関する技術が注目されている [2, 3, 5]．データマイニングとは，データベースとして格納されている大量データの中から，そこに潜在する知識や事実を発見する手法である．

データマイニングの本質は，分析対象である大量の生データ群に内在する知識を抽出することにある．そのためには，多量のデータを対象として意味的に近いデータ群をグループ化し，クラスタ分析することが有効であると考えられる．

本稿では，データの意味的解釈を伴うデータマイニングを実現するための基礎となる意味的動的クラスタリング方式を提案する．提案方式の特徴は，データの意味を考慮しつつ，文脈に応じて動的にクラスタリングを行う点にある．本方式により，データの意味を扱い，解析対象のデータに対して，文脈に応じた解析結果を得ることが可能となる．

クラスタリングについては，多変量解析の分野

[12] において多くの方式が提案されている．そうした従来の統計的解析法との比較において，本方式の特徴は，文脈に応じて動的に解析結果を求めることができる点にある．すなわち，本方式は，一つの分析対象について静的に解析結果を得るだけでなく，文脈や状況に応じて動的に解析結果を得ることを可能とする．

本方式は，意味の数学モデル [6, 7, 13] の意味的連想処理機構を用いて実現される．意味の数学モデルでは，データ間の意味的な同一性，差異性は，静的な関係によって決定されるのではなく，文脈や状況に応じて動的に変化するものと考える．このモデルはデータ間の意味的な関係を文脈に応じて動的に計算する体系を与えている．本方式は，意味の数学モデルの文脈理解機能を応用し，文脈に応じた意味的なクラスタリングを行うことを可能とする．

意味の数学モデルでは，直交空間における部分空間の選択を行う演算を定義し，その演算によりデータの意味を文脈に応じて動的に解釈する機構を実現している．意味的動的クラスタリング方式は，この部分空間の選択の機構を用いて，文脈を反映した部分空間上にデータ群のマッピングを行った後に，それらのマッピングされたデータ群を対象としたクラスタリングを行うことにより，文脈に応じた動的なクラスタリングを実現する．この方式では，部分空間の選択後に，クラスタリングのアルゴリズムを適用する．分析対象に応じて，自由にクラスタリングのアルゴリズムを選択可能である．

2 意味的動的クラスタリング方式の概要

本節では，意味的動的クラスタリング方式の概要を示す．

2.1 概要

本方式は，次の 4 ステップにより実現される．

Step-1 ：正規直交空間の生成
分析対象アイテム群を特徴づける特徴量群を抽出し，正規直交空間を生成する．

Step-2 ：分析対象アイテム群の意味空間へのマッピング
分析対象アイテム群を抽出した特徴量群で特徴づけ，**Step-1** で生成した正規直交空間にマッピングする．

Step-3 ：問合せに応じた部分空間選択
意味的連想検索方式の応用により，問合せ（文脈語列）に応じて部分空間選択を行う．

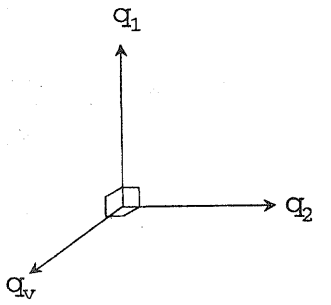

図 1: Step-1: 正規直交空間の生成 ($q_1 \sim q_v$: 正規直交軸)

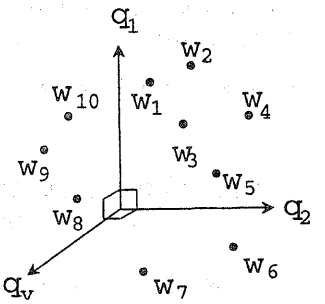

図 2: Step-2: 分析対象アイテム群の意味空間へのマッピング ($w_1 \sim w_{10}$: 分析対象アイテム)

Step-4 ：部分空間上での分析対象アイテム群のクラスタリング
Step-3 で選択された部分空間上において，分析対象アイテム群をクラスタリングする．

2.2 正規直交空間の生成

まず，全ての分析対象アイテム群を特徴づけることができる特徴量群を抽出する．それを用いて，相関量を計算する場となる正規直交空間を生成する（図 1）．

2.3 分析対象アイテム群の意味空間へのマッピング

全ての分析対象アイテム群を，前項で抽出した特徴量群で特徴づける．それを用いて，生成した正規直交空間に分析対象アイテム群をマッピングする（図 2）．

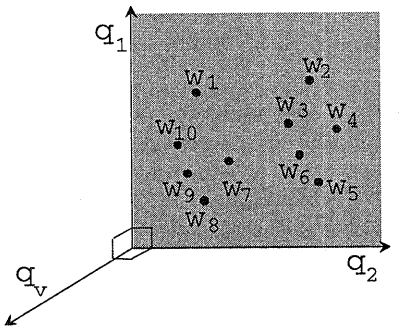

図 3: Step-3: 問合せに応じた部分空間選択

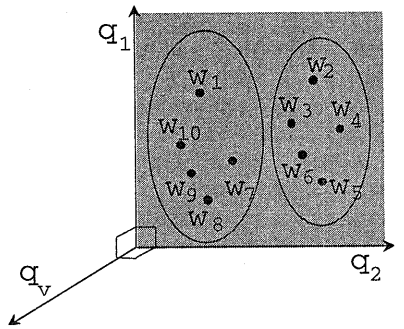

図 4: Step-4: 部分空間上での分析対象アイテム群のクラスタリング

2.4 問合せに応じた部分空間選択

意味の数学モデル [6, 7, 13] の特徴である部分空間選択の方式を用いて，問合せに応じて，生成した正規直交空間の部分空間を動的に選択する（図 3）．全ての分析対象群は，選択された部分空間にマッピングされる．

2.5 部分空間上での分析対象アイテム群のクラスタリング

前項で選択された部分空間上において，分析対象アイテム群をクラスタリングする（図 4）．すなわち，文脈に応じた意味的解釈を伴う動的なクラスタリングを行う．以上の手続きにより，分析者の多様な視点に動的に対応することが可能である．

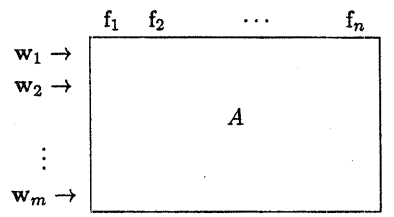

図 5: データ行列 A の構成

3 動的クラスタリング方式の定式化

ここでは，意味の数学モデル [6, 7] に基づいた意味的動的クラスタリングの定式化について述べる．

本方式では，次の 3 種類の特徴付ベクトル群が与えられていることを前提とする．第 1 は，イメージ空間を生成するための特徴付ベクトル群である．第 2 は，文脈語列（問合わせ）ためのの特徴付ベクトル群である．最後は，分析対象アイテム群に対応する特徴付ベクトル群である．

3.1 イメージ空間 $\mathcal{I}$ の設定

ここでは，m 個の単語について各々 n 個の特徴 $(f_1, f_2, \cdots, f_n)$ を列挙した各単語に対する特徴付ベクトル $\mathbf{w}_i (i = 1, \cdots, m)$ が与えられているものとし，そのベクトルを並べた m 行 n 列のデータ行列を A とする（図 5）．

1. データ行列 A の相関行列 A^TA を作る．

2. A^TA を固有値分解する．

$$A^TA = Q \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_\nu & & \\ & & & 0 & \\ & & & & \ddots 0 \end{pmatrix} Q^T,$$

$$0 \leq \nu \leq n.$$

ここで行列 Q は，

$$Q = (\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n)^T$$

である．この $\mathbf{q}_i$ は，相関行列の固有ベクトル，つまり意味素である．

3. このとき，イメージ空間 $\mathcal{I}$ を以下のように定義する．

$$\mathcal{I} := span(\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_\nu).$$

$(\mathbf{q}_1, \cdots, \mathbf{q}_\nu)$ は $\mathcal{I}$ の正規直交基底である．

3.2 意味射影集合 Π_ν の設定

P_{λ_i} を次の様に定義する．

$P_{\lambda_i} \overset{\mathrm{d}}{\Longleftrightarrow} \lambda_i$ に対応する固有空間への射影，

i.e. $P_{\lambda_i} : \mathcal{I} \to span(\mathbf{q}_i)$.

意味射影の集合 Π_ν を次のように定義する．

$$\begin{aligned}\Pi_\nu := \{\ & 0, P_{\lambda_1}, P_{\lambda_2}, \cdots, P_{\lambda_\nu}, \\ & P_{\lambda_1}+P_{\lambda_2}, P_{\lambda_1}+P_{\lambda_3}, \cdots, P_{\lambda_{\nu-1}}+P_{\lambda_\nu}, \\ & \vdots \\ & P_{\lambda_1}+P_{\lambda_2}+\cdots+P_{\lambda_\nu}\}.\end{aligned}$$

Π_ν の要素の個数は 2^ν 個であり，これは 2^ν 通りの意味の様相表現ができることを示している．

3.3 意味解釈オペレータ S_p の構成

文脈ベクトル

$$s_\ell = (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_\ell)$$

と，しきい値 $\varepsilon_s (0 \leq \varepsilon_s < 1)$ が与えられたとき，意味解釈オペレータ S_p は，その文脈ベクトル s_ℓ に応じて，意味射影 $P_{\varepsilon_s}(s_\ell)$ を決定する．すなわち，$s_\ell \in T_\ell$（ここで T_ℓ は ℓ 語によって構成される語群シーケンスのすべての集合である．），$\Pi_\nu \ni P_{\varepsilon_s}(s_\ell)$ とすると，意味解釈オペレータ S_p は，T_ℓ から Π_ν への作用素として定義される．また，$\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_\ell\}$ は，特徴付ベクトルであり，データ行列 A の特徴と同一の特徴を用いている．オペレータ S_p は次のように定義される．

1. $\mathbf{u}_i (i = 1, 2, \cdots, \ell)$ をフーリエ展開する．$\mathbf{u}_i$ と $\mathbf{q}_j$ の内積を u_{ij} とする．

   $$u_{ij} := (\mathbf{u}_i, \mathbf{q}_j)\ ,\ j = 1, 2, \cdots, \nu.$$

   ベクトル $\hat{\mathbf{u}}_i \in \mathcal{I}$ を次のように定める．

   $$\hat{\mathbf{u}}_i := (u_{i1}, u_{i2}, \cdots, u_{i\nu}).$$

   これは，単語 $\mathbf{u}_i$ をイメージ空間 $\mathcal{I}$ に写像したものである．

2. 文脈ベクトル s_ℓ の意味重心 $\mathbf{G}^+(s_\ell)$ を求める．*

   $$\mathbf{G}^+(s_\ell) := \frac{\left(\sum_{i=1}^{\ell} u_{i1}, \sum_{i=1}^{\ell} u_{i2}, \cdots, \sum_{i=1}^{\ell} u_{i\nu}\right)}{\| \left(\sum_{i=1}^{\ell} u_{i1}, \sum_{i=1}^{\ell} u_{i2}, \cdots, \sum_{i=1}^{\ell} u_{i\nu}\right) \|_\infty}$$

   この $\|\cdot\|_\infty$ は，無限大ノルムを示す．

3. 意味射影 $P_{\varepsilon_s}(s_\ell)$ を決定し，部分空間（以下，意味空間とよぶ）を選択する．

   $$P_{\varepsilon_s}(s_\ell) := \sum_{i \in \Lambda_{\varepsilon_s}} P_{\lambda_i} \quad \in \Pi_\nu.$$

   但し $\Lambda_{\varepsilon_s} := \{\, i \mid (\mathbf{G}^+(s_\ell))_i > \varepsilon_s \}$ とする．

3.4 意味空間における距離の定義

文脈語ベクトル s_ℓ が与えられたとする．また，分析対象アイテム x と分析対象アイテム y の特徴つきベクトルを，イメージ空間に写像したベクトルを $\mathbf{x} \in \mathcal{I}$，$\mathbf{y} \in \mathcal{I}$ とする．このデータ間の距離 $\rho(\mathbf{x}, \mathbf{y}; s_\ell)$ を次のように定める．

$$\rho(\mathbf{x}, \mathbf{y}; s_\ell) = \sqrt{\sum_{j \in \Lambda_{\varepsilon_s}} \{c_j(s_\ell)(x_j - y_j)\}^2},$$

ここで，$c_j(s_\ell)$ は，文脈ベクトル s_ℓ に依存して決まる重みであり，次のように定義する．

$$c_j(s_\ell) := \frac{\sum_{i=1}^{\ell} u_{ij}}{\| \left(\sum_{i=1}^{\ell} u_{i1}, \cdots, \sum_{i=1}^{\ell} u_{i\nu}\right) \|_\infty},$$

$$j \in \Lambda_{\varepsilon_s}.$$

このように，距離計算において，イメージ空間を構成する各意味素（固有ベクトル）に重みづけ $(c_j(s_\ell))$ を行うことにより，ε_s の値が小さい場合，すなわち，3.3節 (3) において意味空間を構成するために選択される固有ベクトルの数が多くなる場合においても，文脈の認識に関する ε_s の値の影響を小さくしている．

3.5 意味空間におけるクラスタリング方式

本方式は，文脈を反映した部分空間上にデータ群のマッピングを行った後に，それらのマッピングされたデータ群を対象としたクラスタリングを行うことにより，文脈に応じた動的なクラスタリングを実現する方式である．

ここでは，部分空間上のクラスタリングのアルゴリズムとして融合法 [12] を用いた方式について述べる．融合法は，以下のように記述される．

1. k 分析対象アイテムについて，全ての分析対象アイテムから全ての分析対象アイテムへの距離を求める．すなわち，3.4 節で定義した距離計算を $k(k-1)/2$ 回行う．

2. k 分析対象アイテムを，k 個のクラスタとみなす．各々のクラスタは，意味空間上の座標として，各々のクラスタを構成する 1 つの分析対象アイテムの意味空間上の座標を持つ．

3. 最小距離を持つ一組の分析対象アイテムを一つのクラスタとする．生成されたクラスタを意味空間上の 1 点で代表させる．

4. (3) の操作を，分析対象アイテム群が指定された個数のクラスタになるまで繰り返す．

ここで，(3) における，個々のクラスタを意味空間上の 1 点に代表させる方法については，次の 4 方法がある．

(a) クラスタを構成するある分析対象アイテムの座標を用いる．

(b) クラスタの重心の座標を用いる．

(c) クラスタ間の最小距離を求める毎に，クラスタを構成する分析対象アイテムのうち，距離が最小となる分析対象アイテムの座標を用いる．

(d) クラスタ間の最小距離を求める毎に，クラスタを構成する分析対象アイテムのうち，距離が最大となる分析対象アイテムの座標を用いる．

4 提案方式の実現

4.1 距離計算のためのメタデータの生成

本方式における，意味空間上での距離計算に用いられる各メタデータについては，次に示す方法によって生成した．

4.1.1 イメージ空間生成用のメタデータの生成

空間生成用メタデータの生成，すなわち，データ行列 A の生成を行うために，“Longman Dictionary of Contemporary English[10]” (以下，“LD” と呼ぶ) の英英辞典を使用した．この LD は，約 2,000 語の基本語だけを用いて約 56,000 語の見出し語を説明している．約 2,000 語の基本語をデータ行列 A の列，すなわち，特徴とした．

次の操作を行うことにより，3.1節におけるイメージ空間の作成に使用するデータ行列 A を自動生成した．空間生成用メタデータの各単語 (2,148 語) について，各単語の説明語として LD から取り出した基本語群を用いて，2,148 行 2,148 列の行列を作成した．その単語を説明する基本語が肯定の意味に用いられていた場合 “1”，否定の場合 “-1”，使用されていない場合 “0” とし，見出し語自身が特徴である場合その特徴の要素を “1” として自動生成する．その操作後に，列ごとに 2 ノルムで正規化する．

3.1節における固有値分解の際の固有値の数，すなわちイメージ空間の次元数は，2,133 であった．

4.1.2 分析対象アイテム群のメタデータの生成

イメージ空間へ写像する分析対象アイテム群のメタデータ生成については，各分析対象アイテム群を説明する LD の基本語群からの行ベクトルの生成おいて，その単語を説明する基本語が肯定の意味に用いられていた場合，その特徴の要素を “1”，否定の場合 “-1”，使用されていない場合 “0” とし，文脈語自身が特徴語である場合，その特徴の要素を “1” として自動生成する．

4.1.3 文脈語列 (問合わせ) メタデータの生成

イメージ空間へ写像する文脈語列のメタデータ生成については，各文脈語を説明する BD の基本語群からの行ベクトルの生成おいて，その単語を説明する基本語が肯定の意味に用いられていた場合，その特徴の要素を “1”，否定の場合 “-1”，使用されていない場合 “0” とし，文脈語自身が特徴語である場合，その特徴の要素を “1” として自動生成する．

4.2 クラスタリングのアルゴリズム

意味空間上でのクラスタリングのアルゴリズムは，3.5節で述べた．ここで，個々のクラスタを意味空間上の 1 点に代表させる方法については，そのクラスタの重心の座標を用いる方法を採用した．この方法については，3.5節の (b) として述べた．これは，(c) および (d) と比較して (b) は計算量が比較的少なく，さらに，(a) と比較して (b) は，そのクラスタを代表させるために最も客観的な方法であると考えられる．

5 実験

本節では，実験により，意味の数学モデルを用いた意味的動的クラスタリング方式の実現可能性を検証する．

5.1 実験環境

4節で述べた方法により，実験システムを構成した．

分析対象アイテム群として，“Longman Dictionary of Contemporary English[10]” (以下，“LD”

```
...
about1 [prep] *on1 *the1 *subject1 *of :
about2 [adv] *here1 *and *there1 ;
       in1 all1 directions[n] or places1 :
about3 [adj] *out1 *of bed1 ; active1 :
above2 [adv] in1 or to1 a higher1 place1 :
above1 [prep] higher2 than2 ; over1 :
abroad [adv] to1 or in1 another country1 or
       countries :
absence [n] the1 state1 or a period of
        being2 away1 :
absent1 [adj] not -present3 :
accept [v] to3 take1 or receive , ( something
       offered or given ) ,
       [esp.] willingly :
acceptable [adj] good1 enough1 ; satisfactory :
...
correct1 [adj] *based *on1 or *in1 *accordance
         *with the1 truth or the1 facts ;
         right3 ; without -mistakes2 :
...
experience1 [n] ( the1 gaining of ) { knowledge
            or skill } which comes from practice
            in1 { an activity or doing something }
            for1 a { long1 time1 }  , rather[adv]
            than2 from books1- :
experience2 [v] to3 { feel1 , | suffer , or learn }
            by1 ( an ) experience1 :
...
```

図 6: 実験に使用した分析対象アイテム群（部分）

と呼ぶ）の英英辞典による単語データを用いた．LD の約 2,000 語の基本語から，同音異義語を別の単語として扱い，2,945 語を抽出した．この 2,945 語を分析対象アイテム群に設定した．単語データの ID については，同音異義語を考慮して，about1，about2 のように単語データに数字を付加したものを ID とした．意味が一意に定まる語については，単語データのみを ID とした．この分析対象アイテム群の一部を図 6に示す．

5.2 実験方法

分析対象アイテム群に様々な文脈語列（問合わせ）を実験システムに与え，解析結果を得る．1 分析対象アイテム群から文脈に応じた解析結果が得られることを確認する．

文脈語列（問合わせ）には，次の 3 種類を与えた．与えた検索語列は，“colour”，“light,bright”，“serious,grave” である．

また，クラスタリングの際のパラメータとして，

```
...
cluster 568:
copper1 metal1 tin1
cluster 576:
correct1 think1 opinion judgment right3
cluster 579:
cotton1
...
cluster 880:
expensive
cluster 881:
experience1
cluster 918:
farm1
...
```

図 7: 実験結果 1（文脈: colour，部分）

```
...
cluster 568:
copper1 metal1 tin1
cluster 576:
correct1 opinion judgment formal
cluster 579:
cotton1
...
...
```

図 8: 実験結果 2（文脈: light,bright，部分）

クラスタ数を 300 に設定した．これは，分析対象アイテム群の数の約 1/10 の数である．すなわち，1 クラスタを構成する分析対象アイテムの数は平均約 10 である．

5.3 実験結果

実験結果を次のように示す．3 文脈語列 “colour”，“light,bright”，“serious,grave” に対応する実験結果は，それぞれ，図 7，図 8，図 9である．どの実験結果においても，与えられた文脈語列に無関係な巨大なクラスタ（以下，Zero クラスタと呼ぶ）が得られた（Zero クラスタの内容については，ここでは省略する）．

5.4 考察

実験結果より，互いに意味的に相関の強い単語が同一クラスタを構成していることが確認できる．さらに，与えた文脈語列により，クラスタ構成の様子が変化しているのが確認できる．

```
...
cluster 568:
copper1 tin1 metal1
cluster 576:
correct1 opinion right3 formal
cluster 579:
cotton1 flag5 plant1 vegetable1 root1 crop1
...
cluster 863:
examine
cluster 881:
experience1 skill learn knowledge
cluster 902:
faint1
...
```

図 9: 実験結果 3 (文脈: serious, grave , 部分)

例えば, 分析対象アイテムとして correct1 に注目した場合, 次のように分析できる. 文脈語列 (問合せ) として "colour" を与えた場合には correct1, think1, opinion, judgment, right3 の 5 単語がひとつのクラスタを形成した. 文脈語列として "light, bright" を与えた場合には correct1, opinion, judgment, formal の 4 単語がひとつのクラスタを形成した. 文脈が "serious, grave" の場合には correct1, opinion, right3, formal の 4 単語が一つのクラスタを形成した.

分析対象アイテムとして experience1 に注目した場合は, 次のように分析できる. 文脈語列 (問合せ) として "colour" を与えた場合には experience1 のみがひとつのクラスタを形成した. 文脈語列として "light, bright" を与えた場合には experience1 は, Zero クラスタの構成要素となった. 文脈が "serious, grave" の場合には experience1, skill, learn, knowledge の 4 単語が一つのクラスタを形成した. この例では, ある分析対象アイテムに相関の強い文脈のもとでは, その単語に意味的に相関の強い分析対象アイテム (同義語など) 群がクラスタを形成することが分かる.

また, Zero クラスタについては, 次のように考察できる. 本方式は, 文脈を与え, その文脈のもとで分析対象アイテム群を意味的な近さによりクラスタリングする方式と位置付けられる. よって, その文脈に意味的に相関のない分析対象アイテム群は, 文脈によって選ばれた部分空間内で一つの巨大なクラスタを形成すると予想される. 事実, 実験によりその文脈に無関係な分析対象アイテムだけの集合である巨大なクラスタを確認することができた.

この実験結果は, 文脈依存の意味的動的クラスタリング方式の実現可能性を示している.

6 結論

本稿では, データの意味的な解釈を伴うデータマイニングを行うための基礎となる意味的動的クラスタリング方式を提案した. 本方式は, 文脈に依存した意味的な近さに応じて動的にクラスタリングができる点が特徴である. さらに, 既存のクラスタリングのアルゴリズムを自由に組み合わせる可能である. 本方式により, 解析対象のデータに対して, 文脈に応じて動的に意味的解析結果を得ることが可能となった. また, 実験により, 本方式の実現可能性を確認した.

今後は, 本方式の高速化, 本方式のデータマイニングシステムへの適用, 分析対象アイテム群の特徴量抽出方式の確立, および, 本方式の各種メディアへの適用を行う予定である.

謝辞

本研究の一部は, 文部省科学研究費重点領域研究「高度データベース」によっています. ここに記して感謝致します.

参考文献

[1] Brachman, R.J., Khabaza, T., Kloesgen, W., Piatetsky-Shapiro, G. and Simoudis, E., "Mining Business Databases," Communications of the ACM, Vol.39, No.11, pp. 41-48, Nov. 1996.

[2] Fukuda, T., Morimoto, Y., Morishita, S., Tokuyama, T., "Data Mining using Two-dimensional Optimized Accociation Rules: Scheme, Algorithms, and Visualization," Proceedings of ACM SIGMOD Conference on Management of Data, pp. 13-23, June 1996.

[3] 福田剛志, 森本康彦, 森下真一, 徳山豪, "データマイニングの最新動向 – 巨大データからの知識発見技術 –," 情報処理, Vol. 37, No. 7, pp. 597-603, 1996.

[4] Kitagawa, T. and Kiyoki, Y.: The mathematical model of meaning and its application to multidatabase systems, Proceedings of 3rd IEEE International Workshop on Research Issues on Data Engineering: Interoperability

in Multidatabase Systems, pp. 130-135, April 1993.

[5] 喜連川優,"データマイニングにおける相関ルール抽出技法," 人工知能学会誌, Vol. 12, No. 4, pp. 513-520, Jul. 1997.

[6] 清木 康, 金子 昌史, 北川 高嗣: 意味の数学モデルによる画像データベース探索方式とその学習機構, 電子情報通信学会論文誌,D-II,Vol.J79-D-II,No. 4,pp. 509-519, 1996.

[7] Kiyoki, Y., Kitagawa, T. and Hayama, T.: A metadatabase system for semantic image search by a mathematical model of meaning, ACM SIGMOD Record, vol. 23, no. 4, pp. 34-41, 1994.

[8] Kiyoki, Y.,Kitagawa, T. and Hitomi, Y.: A fundamental framework for realizing semantic interoperability in a multidatabase environment, Journal of Integrated Computer-Aided Engineering, Vol.2, No.1, pp. 3-20, John Wiley & Sons, Jan. 1995.

[9] 河野浩之,"データベースからの知識発見の現状と動向," 人工知能学会誌, Vol.12, No.4, pp. 497-504, Jul. 1997.

[10] Longman Dictionary of Contemporary English, Longman, 1987.

[11] 西尾章治郎,"大規模データベースにおける知識獲得," 情報処理, Vol. 34, No. 3, pp. 343-350, 1993.

[12] 塩谷實,"多変量解析概論," 朝倉書店, 1990.

[13] 吉田尚史, 清木康, 北川高嗣, "意味的連想検索機能を持つメディア情報検索システムの実現方式," 情報処理学会論文誌, Vol. 39, No. 4, pp. 911-922, 1998.