HOKKAIDO UNIVERSITY

Title	オンラインユニットクラスタリング問題の競合比の自動解析
Author(s)	川原, 純; 小林, 浩二
Citation	2010年度科学技術振興機構ERATO湊離散構造処理系プロジェクト講究録. p.380.
Issue Date	2011-06
Doc URL	http://hdl.handle.net/2115/48363
Type	conference presentation
Note	ERATO 湊離散構造処理系プロジェクトシンポジウム（第1回）：第9回情報科学技術フォーラム(FIT2010)イベント企画セッション. 2010年9月8日（水）. 九州大学伊都キャンパス.
File Information	08.FIT_kawahara.pdf

Instructions for use

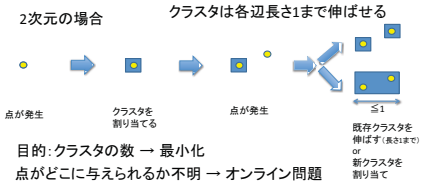
Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

オンラインユニットクラスタリング問題の競合比の自動解析

JST ERATO研究員 川原 純　　joint work with 小林浩二

オンラインユニットクラスタリング問題

2次元の場合　　クラスタは各辺長さ1まで伸ばせる



目的：クラスタの数 → 最小化

点がどこに与えられるか不明 → オンライン問題

アルゴリズムの評価法

アルゴリズムの競合比 = $\max_{\sigma} \frac{ALG(\sigma)}{OPT(\sigma)}$　入力

アルゴリズム ALG = 3

最適な配置 OPT = 2

最悪時評価

競合比 = $\frac{ALG}{OPT} = \frac{3}{2}$

応用例　施設配置計画問題

既存結果と本研究の結果

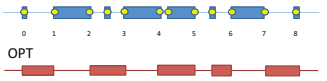
1次元オンラインクラスタ問題

競合比			
決定性	上限	7/4 = 1.75	[Epstein 08]
	下限	8/5 = 1.6	
乱択	上限	7/4 = 1.75	
	下限	3/2 = 1.5	

今回は1次元のみを扱う

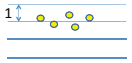
下限を与える入力　[Epstein 08]

3, 4, 5, 6, 2, 1, 0, 2.5, 7, 4.1, 5.5, 8



競合比 = 8/5

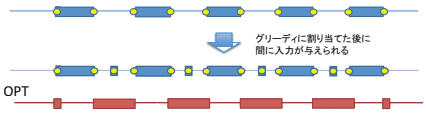
2次元以上はほぼ自明な結果のみ [Chan 04], [Epstein 08]



線を引いてそれぞれの領域で独立に1次元のアルゴリズムを実行

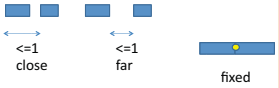
競合比は1次元の2倍（d次元で 2^{d-1}倍）

グリーディアルゴリズム



競合比は2

既存研究（決定性競合比= 7/4）[Epstein 08]



アルゴリズムは上記2種類の状態を回避するように動く

本研究のアルゴリズム → 回避しない。競合比を悪くしないように配置

本研究の結果

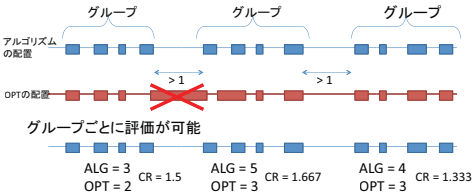
グルーピングアルゴリズムの開発

競合比の上限 5/3 = 1.667 を証明

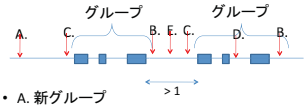
（1.75 → 1.667　に改良）

提案アルゴリズム

グルーピングアルゴリズム（GA）

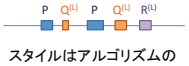


- GAの動作は、点の与えられる場所に応じて次の5つの場合に分けられる



- A. 新グループ
- B. グループの右端(1以内)
- C. グループの左端（B.と対称）
- D. グループの中
- E. グループの結合

- GA はアルゴリズムの進行中、各クラスタに6種類のラベルを振る（スタイルと呼ぶ）
- {P, $Q^{(L)}$, $Q^{(R)}$, $R^{(L)}$, $R^{(R)}$, S}

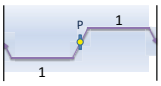


- スタイルはアルゴリズムの進行と共に変化する
- 下記の性質をもつことが期待される

(1) 新しいクラスタを割り当て、ラベル 'P' を振る

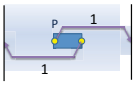
スタイルと呼ぶ

(2) ラベルや位置関係に応じて、守備範囲が決まる



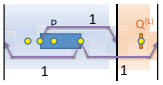
P の守備範囲は基本的には"伸ばせる限り"

(3) 基本的に守備範囲内にあるクラスタが担当



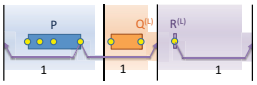
(4) 守備範囲外なら、新しく $Q^{(L)}$ を作成

$Q^{(L)}$の守備範囲は　左："ひっつく"　右：左Pの右から+1



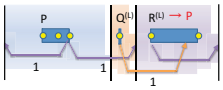
(5) $R^{(L)}$ の左守備範囲は「左Pの右＋1」

右守備範囲は伸ばせる限り

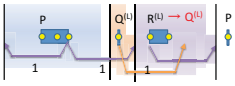


アルゴリズムの記述（一部）

(6-1) R は P の性質が保てると確定した時点で P になる



(6-2) それ以外で守備範囲外に与えられたときはQ(L)になる



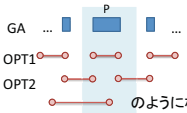
計算機を援用したアルゴリズムの解析

- OPT の評価はOPTのクラスタ配置の端の点を数える

OPT = ○ の数 / 2

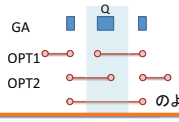
○ を（OPTの）得点と呼ぶ

- P は少なくとも2点の得点が期待されるクラスタ



のようになることはない

- Q は少なくとも1点の得点が期待されるクラスタ



のようになることはない

- Pを基準にする

R Q P Q Q Q P Q Q S Q P Q Q P Q S P Q Q Q P Q Q Q Q P Q R

- 証明すること
  - P は必ず2点になっている
  - P 以外は必ず1点になっている
  - PとPの間にはクラスタが高々4個しか存在しない

ALG = 5

OPT = 得点/2 = 3

競合比 = ALG/OPT <= 5/3

・計算機を用いて起こり得る"状態"をすべて列挙する

最初は点1つから始める

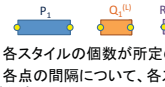
点が発生する場合が各クラスタの守備範囲の内か外かで場合分けを行う（計算機で自動的に可能）

場合分けを行ったときに成り立つべき点間の条件を不等式として記憶

・状態遷移図を作成

アルゴリズムの設計から状態は有限個になる

・競合比が目標値を下回ることを全状態でチェック



$Q_1^{(L)}$.left − P_1.left > 1

$R_1^{(L)}$.left − P_1.right > 1 ...

- 各スタイルの個数が所定の関係を満たすかチェック
- 各点の間隔について、各スタイルが所定の条件を満たすかチェック

・条件式は点のペア数だけ存在し、それぞれの条件式が成立するかしないか2通りの場合が生じる。

→ ZDD を用いて組合せ集合を保持