

アニーリング法による乗り合いのマッチング最適化

理学専攻・情報科学コース 1940647 佐藤 由佳

1 はじめに

近年、移動利便性の向上や移動手段確保の観点から、乗合タクシーやオンデマンドバスといった乗合型の交通サービスが全国的に検討、普及し始めている。本研究では都市部における乗合タクシーに着目し、同じ出発地から異なる目的地に移動するタクシーユーザー（以下、ユーザー）が多くいるようなシチュエーションでの乗り合いを考えた。

乗り合いは、タクシーの乗車定員などの制約条件を満たしつつ、乗り合う複数ユーザーのマッチング、すなわち組合せを求める問題として捉えることができる。このように、制約条件のもとで目的関数を最小化もしくは最大化する問題は最適化問題と呼ばれ、特に解の集合が組合せ的な構造を持つ問題を組合せ最適化問題という [1]。本研究では、この組合せ最適化問題に対し、アニーリング法によるクラスタリングを適用し、ユーザー集団のグループ分けを行った。

2 乗り合いにおけるクラスタリング

2.1 乗り合いのモデル

乗り合いの数理モデルは、以下である。

$$H = \sum_{k=1}^K \sum_{i<j}^I d_{ij} q_{ik} q_{jk} + H_{p1} + H_{p2} \quad (1)$$

d_{ij} はデータ間距離、 I はユーザー数、 K はタクシー数を示す。変数 q_{ik} はユーザー i におけるタクシー k への乗車の有無を表しており、乗る場合は 1、乗らない場合は 0 となる。 H_{p1} 、 H_{p2} は制約項である。ユーザー i 、 j が同じタクシーに乗った場合のみ、データ間距離が加算される。同じタクシーに乗ったユーザーのデータ間距離が短いほど、式 (1) の第 1 項が小さくなり、最適な組合せが求められる。

本研究では、特徴量として各ユーザー間の目的地の位置情報を与え、データ間距離を算出した。

$$d_{ij} = \sqrt{(\text{lat}_i - \text{lat}_j)^2 + (\text{lng}_i - \text{lng}_j)^2} \quad (2)$$

lat 、 lng はそれぞれ平均 0、分散 1 となるよう標準化された緯度・経度である。標準化は以下のように行なった。

$$f = \frac{f_{\text{org}} - \mu}{\sigma} \quad (3)$$

f_{org} が元データ、 μ は平均、 σ は標準偏差を表す。

また、式 (1) の制約項 H_{p1} 、 H_{p2} はそれぞれ「ユーザー 1 人は 1 台のタクシーにしか乗れない」、「タクシーには L 人以下しか乗れない」という制約条件を表し、次式で示される。

$$H_{p1} = \alpha \sum_{i=1}^I \left(\sum_{k=1}^K q_{ik} - 1 \right)^2 \quad (4a)$$

$$H_{p2} = \beta \sum_{k=1}^K \left\{ \left(\sum_{i=1}^I q_{ik} - \sum_{l=0}^L l y_{lk} \right)^2 + \left(\sum_{l=0}^L y_{lk} - 1 \right)^2 \right\} \quad (4b)$$

L はタクシーの乗車定員、 l は乗車人数である。本研究では、 $L = 4$ とした。変数 y_{lk} はタクシー k の乗車人数が l 人かどうかを決定する。 l 人であれば 1、そうでなければ 0 をとる。式 (4a)、(4b) は各項のペナルティ係数 α 、 β を別々に調整する必要がある。

アニーリング法を利用するため、式 (1) を以下の QUBO に変換する。

$$H = \sum_{a<b} J_{ab} x_a x_b + \sum_a h_a x_a \quad (5)$$

J_{ab} 、 h_a は 2 次、1 次の係数である。式 (1) の変数 q_{ik} 、 y_{lk} は $x_a \in \{q_{11}, q_{12}, \dots, q_{IK}, y_{01}, \dots, y_{LK}\}$ と変換される。

2.2 実行データ

駅から多くのユーザーが移動するような状況を想定した仮想データを作成した。図 1 に作成データの一例を示す。ユーザー数は 50 人とし、3 方面に偏った状態で分布している。

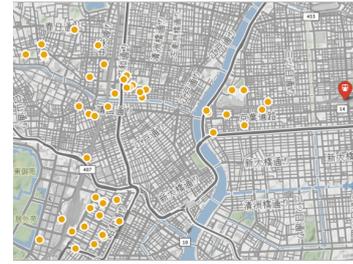


図 1: 作成データの一例。赤ピンが出発地、オレンジの点が各ユーザーの目的地を示す。

2.3 アニーリング結果

50 人のユーザー集団を 15 台のタクシーに乗り合わせることにした。シミュレーテッド・アニーリング法（以下、SA 法）を用いて実行した。図 2 がクラスタリング結果を地図にマッピングしたものである。この結果は、2 つの制約条件を満たし、かつペナルティ係数の小さい解である。目的地が近いユーザー同士が同じクラスになるよう、クラスタリングされていることが確認できた。

3 大規模ユーザー集団分割におけるクラスタリング

本研究ではユーザー集団が大規模になると、求める乗り合いのグループ規模が最大 4 人と細かく分割する必要があるため、1 つの目的関数で良い近似解を求めることが難しい。そのため、前章のクラスタリング処理を行う前処理として、ユーザー集団の大まかな分割をアニーリング法を用いて行った。

3.1 大まかな分割を行うモデル

3 つの分割手法を検討し、比較した。どの手法も数理モデルの第 1 項の最小化をアニーリング法によって



図 2: クラスタリング結果 (タクシー 15 台). 赤ピンが出発地, 同色の点と同じクラスに属するユーザーの目的地である.

求める.

● 2 クラスタ分割手法

$$H = \sum_{i < j} d_{ij} (2c_i - 1)(2c_j - 1) \quad (6)$$

データ間距離 d_{ij} は式 (2) で与えられている. 変数 $c_i \in \{0, 1\}$ はユーザー i がどちらのクラスに所属するかを示す. 同じクラスに属するユーザー間のデータ間距離は正の値で加算され, 異なるクラスに属するユーザー間のデータ間距離は負の値で加算される. そのため, データ間距離の離れたユーザーは異なるクラスに所属した方が式 (6) のコストは小さくなる.

● K クラスタ分割手法

$$H = \sum_{k=1}^K \sum_{i < j} d_{ij} q_{ik} q_{jk} + \alpha \sum_{i=1}^I \left(\sum_{k=1}^K q_{ik} - 1 \right)^2 \quad (7)$$

データ間距離 d_{ij} は式 (2) である. 変数 $q_{ik} \in \{0, 1\}$ はユーザー i がクラス k に所属するか否かを示す. データ間距離の離れたユーザーは異なるクラスに所属した方が式 (7) の第 1 項は小さくなる. 第 2 項はユーザーがちょうど 1 つのクラスに所属するようにする制約項であり, ペナルティ係数 α でその制約の大きさを調整する.

● コミュニティ検出を応用した K クラスタ分割手法 [2]

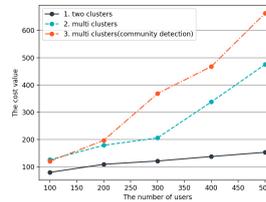
$$H = - \sum_{k=1}^{K_{\max}} \sum_{i < j} \exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right) q_{ik} q_{jk} + \alpha \sum_{i=1}^I \left(\sum_{k=1}^{K_{\max}} q_{ik} - 1 \right)^2 + \gamma \sum_{k=1}^{K_{\max}} \sum_{i < j} q_{ik} q_{jk} \quad (8)$$

第 1 項はデータ間距離 d_{ij} を正規分布に基づいて負数としたものを係数としており, 変数 $q_{ik} \in \{0, 1\}$ はユーザー i がクラス k に所属するか否かを示

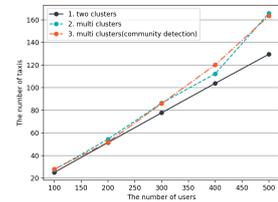
す. 第 2 項は式 (7) と同様である. 第 3 項はクラスタ数を増やそうとする制約であり, クラスタ数が増えると和が小さくなる. これは, 第 1 項がクラスタ数を減らす傾向を持つため, 導入している.

3.2 比較結果

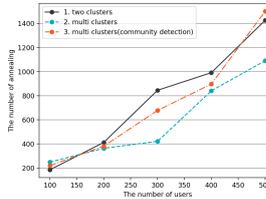
SA 法を用いて第 2.2 節と同様の分布を持つ 100~500 人の仮想データに対し, 手法の検証を行った. 同じタクシーに乗るユーザーの目的地間の距離の総和であるコスト, タクシー数, アニーリング回数を比較した. その結果を以下の図で示す. コストは 2 クラスタ分割が最も小さくなり, 複雑な手法を取るほど大きくなる傾向にあることが確認できた. タクシー数は分割手法による大きな差は生じなかった. またアニーリングの回数はユーザー集団の規模によってばらついているが, 大体 K クラスタ分割, 2 クラスタ分割, コミュニティ検出を応用した K クラスタ分割の順に増えていることが分かる.



(a) コスト.



(b) タクシー数.



(c) アニーリング回数.

図 3: 分割手法の比較. 紺が 2 クラスタ分割, 青が K クラスタ分割, オレンジがコミュニティ検出を応用した K クラスタ分割の結果である.

4 まとめ

本研究では, 2 段階のクラスタリングにより乗り合いのマッチング処理を行った. 大規模ユーザー集団の分割におけるクラスタリングにおいては 3 つの手法を検討し, それぞれの特性を観察した. その結果, 時間 (アニーリング回数) と解の良さがトレードオフの関係になっており, 良い解を求めるには 2 クラスタ分割を用いればよく, それなりの解を高速に得たい場合には K クラスタ分割を用いれば良いことがわかった.

参考文献

- [1] 梅澤俊治, 「組合せ最適化入門: 線形計画から整数計画まで」, 自然言語処理, Vol.21, 1059-1090 (2014).
- [2] 増田直紀, 今野紀雄, 「複雑ネットワーク 基礎から応用まで」, 近代科学社, (2010).