

乱歩モデルに基づく施設配置問題 Facility Location Problem based on Random Walk Model

武藤 伸明^{*} 齊藤 和巳^{*}
大久保 誠也^{*} 池田 哲夫^{*}

Nobuaki MUTOH Kazumi SAITO
Seiya OKUBO Tetsuo IKEDA

本論文では、乱歩モデルに基づき行動するユーザを想定し、与えられた期間内で、配置された複数施設の少なくともどれか一つにそのユーザが到達する確率を考え、この確率を最大にするよう施設配置する問題を扱う。まず、本問題を定式化し、この問題が NP-困難クラスに属すること、その目的関数がサブモジュラ関数となることに言及する。次に、本目的関数の効率的な計算方法を示すとともに、遅延評価と呼ばれる手法を導入した貪欲法に基づく解法を提案する。電力網と現実道路網のデータを用いた評価実験では、代表的な従来手法の本施設問題への適用と比較して、解品質と処理効率の両面で提案法の有効性を確認する。

We address the problem of locating a number of facilities by considering a user behavior model based on random walk, so as to maximize the total probability that the user visits at least one facility during a given time span. In this paper, after formulating our facility location problem, we mention that the problem belongs to NP-hard class, and the objective function has a submodular property. Then, after deriving an efficient way of calculating the objective function, we present a method based on a greedy algorithm equipped with so-called lazy evaluation. In our experiments using a power grid and a real road network, we evaluate the performance of our proposed method in terms of solution quality and processing efficiency, in comparison to those of some representative existing methods.

1. はじめに

工場や倉庫、公共施設、無線基地局などの施設の適切な配置は、多様な経済活動を支える重要な問題である。特に、容量制限なし施設配置問題の研究には長い歴史があり、最近でも、精度保証付きの近似アルゴリズムが幾種も提案されるなど高度に発展を続けている[4]。しかしながら、これらの問題設定では、ユーザは距離最小の施設に向かうなどと仮定され、ユーザ行動についてはあまり考慮されていない。現実には、精緻なユーザ行動モデルが重要な役割を果たす応用問題も存在し得る。

本論文では、乱歩モデルに基づき行動するユーザを想定し、与えられた期間内で、設置された複数施設のどれかに

ユーザが少なくとも一度は到達する確率を考え、この確率を最大にするよう施設配置する問題を扱う。本問題の応用として、インターネット上では、ウェブページへの広告配置が典型例となり、実世界では防犯カメラの設置や防災時の避難施設の配置などが想定できる。特に、インターネット上での広告配置問題では、ユーザ行動モデルとして、ウェブページのランキングで有効性が広く知られているページランク[1]のサーフィンモデルを採用する枠組みとなる。一方、本研究は、ユーザ行動モデルに基づきユーザがネットワーク上を移動する枠組みなので、ネットワークのノードとユーザを同一視するロコミなどによる情報拡散モデル[2, 3]とは基本設定が異なる。

本施設配置問題を離散最適化の観点で捉えれば、NP-困難クラスに属する。よって、大規模になれば妥当な計算時間で厳密解を求めるのが一般的に困難になる。ただし、本問題はサブモジュラ[7]と呼ばれる構造を持つことが示せる。サブモジュラ関数最大化問題は潜在的に幅広い応用が存在し、有限ピボット集合に基づくクラスタリング問題[9]、変数選択問題[5]、設置コスト付き非線形施設配置問題[10]などがその例に挙げられる。また、サブモジュラ性の望ましい性質として、貪欲法で効率よく求まる近似解の最悪ケース解品質を理論的に保証することができる[8]。さらに、遅延評価と呼ばれる手法の導入で貪欲法をさらに効率化することができ[6]、非常に大規模かつ複雑な問題でも妥当な計算時間で精度保証付き近似解を得ることが可能になる。

本論文の構成は以下となる。まず、本施設問題を定式化し、この問題がNP-困難クラスに属し、その目的関数がサブモジュラ関数となることを述べる。次に、本目的関数の効率的な計算方法を示すとともに、遅延評価と呼ばれる手法を導入した貪欲法に基づく解法を提案する。次いで、電力網と現実道路網のデータを用いた実験評価により、提案法の有効性を検証する。最後に、本論文の結果をまとめる。

2. 問題の定式化

施設配置可能な候補地の集合を \mathcal{V} とし、その要素(候補地)数を $V = |\mathcal{V}|$ とする。ここでは、候補地群の隣接関係で構成されるネットワークを考え、施設はノードに配置されるとするので、 \mathcal{V} をノード集合と呼ぶ。また、各ノード $v \in \mathcal{V}$ を整数に対応させて同一視し、 $\mathcal{V} = \{1, \dots, V\}$ とする。

乱歩モデルに基づくユーザはまず、 V -次元の初期確率ベクトル \mathbf{b} に従い、確率 b_v でノード $v \in \mathcal{V}$ を選ぶ($b_1 + \dots + b_V = 1$)。その後、 $(V \times V)$ -次元の推移確率行列 \mathbf{G} に従い(任意の行 v で $G_{v,1} + \dots + G_{v,V} = 1$)、確率 $G_{v,w}$ でノード $w \in \mathcal{V}$ に移動する。

いま、 $v(j)$ を j ステップ目の移動先とし、このような移動を M 回繰り返せば、移動経路のノード列 $\mathbf{n} = (v(0), v(1), \dots, v(M)) \in \mathcal{V}^{M+1}$ を得る。明らかに、一般の初期確率ベクトル \mathbf{b} と推移確率行列 \mathbf{G} では、可能なノード列の総数は V^{M+1} となる。また、各ノード列 \mathbf{n} の生起確率 $p(\mathbf{n})$ は以下となる。

$$p(\mathbf{n}) = b_{v(0)} \prod_{j=1}^M G_{v(j-1), v(j)} \quad (1)$$

以降では、ノード列 \mathbf{n} に出現したノード集合を $S(\mathbf{n}) \subseteq \mathcal{V}$

^{*} 正会員 静岡県立大学 {muto, k-saito, s-okubo, t-ikeda}@u-shizuoka-ken.ac.jp

で表す.

K 個の施設の配置先ノード集合を $\mathcal{R} \subseteq \mathcal{V}$ とする ($K = |\mathcal{R}|$). また, ノード v を含むノード列の集合を $\mathcal{N}(v) = \{n | v \in \mathcal{S}(n)\}$ とし, ノード列集合 $\mathcal{N}(\mathcal{R}) = \bigcup_{v \in \mathcal{R}} \mathcal{N}(v)$ を, 施設配置先ノード集合 \mathcal{R} の要素を少なくとも1つ含むノード列の集合とする. このとき, ユーザの M 回の移動で, 少なくとも一度は \mathcal{R} を訪れる確率 $F(\mathcal{R}; M)$ は以下となる.

$$F(\mathcal{R}; M) = \sum_{n \in \mathcal{N}(\mathcal{R})} p(n), \quad (2)$$

すなわち, 本研究での施設配置問題は, 与えられた施設数 K と移動回数 M に対して式(2)を目的関数とし, これを最大にする施設配置先ノード集合 \mathcal{R}^* を求める問題として定式化される.

式(2)で定義した目的関数は, Leskovecらの定式化した施設配置問題の枠組みでは, DL (Detection Likelihood) 問題に対応する[6]. よって, 本問題はNP-困難クラスに属し, 目的関数 $F(\mathcal{R}; M)$ はサブモジュラ関数となり, 包含関係 $Q \subset \mathcal{R}$ を満たす任意の集合に対して, 以下のサブモジュラ不等式を満たす.

$$F(Q \cup \{v\}; M) - F(Q; M) \geq F(\mathcal{R} \cup \{v\}; M) - F(\mathcal{R}; M).$$

ここで, Leskovecらの研究[6]と対比した本研究の特色は, 式(1)で定義した確率を, Leskovecらはモデルを使わず観測データの経験分布で置き換えるのに対し, 本研究では, 乱歩モデルを土台として解析的に計算する点である. すなわち, 実問題では観測データの入手が困難な場合も想定されるが, 乱歩モデルがユーザ行動の妥当な近似と見なせるならば, 観測データなしでも, 本論文の枠組みならば妥当な施設配置が実現可能となる. また, 移動回数 M をパラメータとしてユーザに指定させ, その大小で長期的と短期的なユーザ行動を分けて考えることも可能になる. 一方, ユーザ移動回数と施設配置数の設定に関しては, 施設への広告などアイテムの配置を想定し, 単位時間あたり一つのアイテムを設置するとすれば, ユーザ移動回数 M は経過時間と見なせるので, K アイテムを設置する全期間での総費用は MKC となる. よって, 本実験では, 総費用一定の下での施設配置問題を想定し, MK が一定の条件の下で実験評価する. なお, 乱歩モデルに基づく問題設定のため, 施設配置先ノード集合はネットワーク上に均等配置される傾向になるが, 距離に基づく均等配置を目的関数とするK-Median法 (詳細は後述) との比較では, 例えば道路網における交通の要所のように, ある種の特徴的な配置先が提案法で求まることも実験で示す.

3. 目的関数の計算法

既に述べたように, ノード列集合の要素数は V^{M+1} となり, V や M が大きくなると単純に目的関数(2)を計算できない. 以下では, 目的関数 $F(\mathcal{R}; M)$ の効率的な計算法について述べる. まず, $M = 0$ ならば, 初期状態で集合 \mathcal{R} の要素が選ばれるときだけなので, 目的関数は次式で計算できる.

$$F(\mathcal{R}; 0) = \sum_{v \in \mathcal{R}} b_v = \mathbf{b}^T \sum_{v \in \mathcal{R}} \mathbf{e}_v \quad (3)$$

ここで, \mathbf{e}_v は第 v 要素の値のみ1で他が0の V -次元ベクトルを表し, \mathbf{e}_v^T は \mathbf{e}_v の転置を意味する. 次に, $M = 1$ のときは, 式(3)の確率 $F(\mathcal{R}; 0)$ とともに, $M = 1$ での移動により初めて \mathcal{R}

の要素に到達する確率 $(\mathbf{b} - \sum_{v \in \mathcal{R}} \mathbf{e}_v \mathbf{e}_v^T \mathbf{b})^T \mathbf{G} \sum_{v \in \mathcal{R}} \mathbf{e}_v$ を加えるので, 目的関数は以下のように計算できる.

$$\begin{aligned} F(\mathcal{R}; 1) &= F(\mathcal{R}; 0) + \left(\mathbf{b} - \sum_{v \in \mathcal{R}} \mathbf{e}_v \mathbf{e}_v^T \mathbf{b} \right)^T \mathbf{G} \sum_{v \in \mathcal{R}} \mathbf{e}_v \\ &= \mathbf{b}^T \left(\sum_{v \in \mathcal{R}} \mathbf{e}_v \mathbf{e}_v^T + \mathbf{G} - \sum_{v \in \mathcal{R}} \mathbf{e}_v \mathbf{e}_v^T \mathbf{G} \right) \sum_{v \in \mathcal{R}} \mathbf{e}_v \\ &= \mathbf{b}^T \left(\mathbf{G} + \sum_{v \in \mathcal{R}} \mathbf{e}_v (\mathbf{e}_v^T - \mathbf{e}_v^T \mathbf{G}) \right) \sum_{v \in \mathcal{R}} \mathbf{e}_v \end{aligned} \quad (4)$$

ここで, 式(4)の最終行の右辺は, ノード $v \in \mathcal{R}$ のそれぞれに対し, 推移確率行列 \mathbf{G} の第 v 行において, 元の $\mathbf{e}_v^T \mathbf{G}$ を単位ベクトル \mathbf{e}_v へ置き換えた推移確率行列による乱歩モデルと見なせる.

式(4)より, \mathbf{b} は初期確率ベクトル, $(\mathbf{G} + \sum_{v \in \mathcal{R}} \mathbf{e}_v (\mathbf{e}_v^T - \mathbf{e}_v^T \mathbf{G}))$ は上述した推移確率行列である. また, $\sum_{v \in \mathcal{R}} \mathbf{e}_v$ は \mathcal{R} に含まれる v のみが1となる単位ベクトルの和であり, \mathcal{R} に含まれる v に到達する確率を足し合わせるためのものである. よって, 式(4)を M ステップに拡張すれば以下を得る.

$$F(\mathcal{R}; M) = \mathbf{b}^T \left(\mathbf{G} + \sum_{v \in \mathcal{R}} \mathbf{e}_v (\mathbf{e}_v^T - \mathbf{e}_v^T \mathbf{G}) \right)^M \sum_{v \in \mathcal{R}} \mathbf{e}_v \quad (5)$$

すなわち, 本問題の目的関数は式(5)で計算可能であり, V 次元正方形行列と V 次元ベクトルの積を M 回繰り返すので, 計算量は高々 $O(MV^2)$ となる. 明らかに, 式(5)で規定する乱歩モデルにおいては, $v \in \mathcal{R}$ に到達したユーザは v にしか移動できない状況となり, これにより一度でも $v \in \mathcal{R}$ を経由したノード列 n の確率 $p(n)$ を加算する仕組みとなっている. すなわち, ノード集合 \mathcal{R} は乱歩モデルの吸収状態となり, M を大きくすれば目的関数 $F(\mathcal{R}; M)$ は1に漸近する.

4. 貪欲法に基づく解法

本問題の目的関数 $F(\mathcal{R}; M)$ はサブモジュラ関数なので, 最適解に対する貪欲解での目的関数値の比は $(1 - 1/e)$ 以上となることが保証される[8]. ここで, e は自然対数の底である. また, 遅延評価(lazy evaluation)の導入によって単純貪欲法の探索効率を向上させることができる[6, 9, 10]. 以下では, 本問題に対応したアルゴリズムと計算量について述べる.

単純貪欲(simple greedy)法とは, 与えられたノード数 $K = |\mathcal{R}|$ に対して, $k = 1, \dots, K$ で既に選定したノード集合を固定し, 目的関数値を最大にするノード v を V から求めて追加することで結果のノード集合 \mathcal{R} を求める方法である. 以下に, 単純貪欲法のアルゴリズムを示す.

- SG1. $k \leftarrow 1, \mathcal{R}_0 \leftarrow \phi$ と初期化する;
- SG2. $\hat{v}_k = \operatorname{argmax}_{v \in \mathcal{V} \setminus \mathcal{R}_{k-1}} \{F(\mathcal{R}_{k-1} \cup \{v\}; M)\}$ を求め, $\mathcal{R}_k \leftarrow \mathcal{R}_{k-1} \cup \{\hat{v}_k\}$ とする;
- SG3. $k = K$ ならば $\mathcal{R} = \{\hat{v}_1, \dots, \hat{v}_K\}$ を出力し終了する;
- SG4. $k \leftarrow k + 1$ とし, ステップ SG2. へ戻る.

ここで, ステップ SG2. における最良追加要素のタイブレークを一意に設定すれば, 単純貪欲法で求まる結果(貪欲解)のノード集合 $\hat{\mathcal{R}}_K$ は常に等しくなる. 単純貪欲法の主たる計算量は, ステップ SG2. を K 回繰り返すことにより,

$O(KMV^3)$ となる。以下では、この方法を SG (simple greedy) 法と呼ぶ。

次に、遅延評価(lazy evaluation)の導入法について述べる。まず、ステップ SG2. の評価式を次式で書き換える。

$$H(v; \mathcal{R}_{k-1}) = F(\mathcal{R}_{k-1} \cup \{v\}; M) - F(\mathcal{R}_{k-1}; M) \quad (6)$$

前述のとおり、目的関数 $F(\mathcal{R}; M)$ はサブモジュラ関数であるため、式(6)は逆単調性が成り立つ。すなわち、 $\mathcal{R}_0 = \emptyset$ とし、単純貪欲法の各反復 k で求まる解を $\mathcal{R}_1, \mathcal{R}_2, \dots$ とすれば、 $H(v; \mathcal{R}_k)$ は $k \geq 1$ で非増加である。したがって、 $1 \leq j \leq k$ の任意の j において、 $H(v; \mathcal{R}_j) \geq H(v; \mathcal{R}_k)$ である。この関係式を利用して探索効率を向上させるのが遅延評価付き貪欲法(greedy with lazy evaluation)である。いま、ノード $v \in \mathcal{V}$ に対する評価上限値を $\zeta(v)$ とし、 $\zeta(v)$ で \mathcal{V} を降順にソートしたリストの上位 i 番目のノードを $r_i \in \mathcal{V}$ とすると、遅延評価付き貪欲法アルゴリズムは以下となる。

- GL1. $k \leftarrow 1, \mathcal{R}_0 \leftarrow \emptyset$ とし、各ノード $v \in \mathcal{V}$ に対し、 $\zeta(v) \leftarrow \infty$ と初期化し、 r_i を任意に定める;
- GL2. $Z \leftarrow 0$ とし、 $i = 1, \dots, V$ の順で GL2-1 から GL2-2 の処理を繰り返す;
 - GL2-1. $\zeta(r_i) \leq Z$ なら、 $\zeta(\hat{v}_k) \leftarrow 0, \mathcal{R}_k \leftarrow \mathcal{R}_{k-1} \cup \{\hat{v}_k\}$ とし、ステップ GL3. に進む;
 - GL2-2. $\zeta(r_i) > Z$ なら $\hat{v}_k \leftarrow r_i, Z \leftarrow \zeta(r_i)$ とする;
- GL3. $k = K$ ならば $\mathcal{R} = \{\hat{v}_1, \dots, \hat{v}_K\}$ を出力し終了する;
- GL4. $\zeta(v)$ でノードを降順にソートしたものを、 r_1, \dots, r_V とし、 $k = k + 1$ としステップ GL2. へ戻る。

ここで、 Z は処理過程の時点で最良評価値を格納する変数である。遅延評価付き貪欲法では、ステップ GL2. の第 k 反復目の最良ノード選定において、 $\zeta(v)$ を用いて各ノード v を降順にソートしたリストを利用する。このリストの先頭から順に探索する過程において、あるノード $v \in \mathcal{V}$ での実際の改善値と Z を比較し、探索リスト上での未探索の先頭要素での上限値が小さくなれば、その時点で最良ノード \hat{v}_k が求まる。明らかに、遅延評価付き貪欲法の結果は単純貪欲法の結果と一致する。以下では、この方法を GL (greedy with lazy evaluation) 法と呼ぶ。

GL 法の主たる計算量について述べる。ステップ GL2. での平均反復回数を α ($1 \leq \alpha \leq V$)とすれば、ステップ GL2. の総計算量は $O(\alpha MV^2)$ となる。一方、GL4. では、評価上限値による $(V - K)$ 個の施設をソートするので $O((V - K) \log(V - K))$ の計算量が必要となる。よって、GL2. からGL4. を K 回反復するので、遅延評価付き貪欲法の主たる計算量は $O(\alpha K MV^2)$ となる。すなわち、遅延評価付き貪欲法の計算効率は α に依存するが、この量を解析的に求めることは困難である。

5. 評価実験

5.1 評価データと実験設定

実験データには、幅広く利用可能なベンチマーク実ネットワークとして、Watts らがスモールワールドに関する論文で用いた電力網[11] (ノード数 4,941, リンク数 6,594) と、現実的な応用を想定したネットワークとして、全国デジタル道路地図データベース(CD) 静岡市版[12] をもとに道路網データを抽出・作成した静岡市の道路網 (ノード数 42,716, リンク数 59,713) を用いた。ここで、道路網のノードとは任意の交差点に対応し、それらを結ぶ道をリンクと見なした。以下の評価実験では、道路網の作成処理簡便化などのため、これらネットワークを無向グラフとして扱った。ただし、本論文の問題設定や解法は任意の有向グラフに適用可能な枠組みである。

実験では、ユーザは各ノードにおいて隣接ノードをランダムに一つ選択して次のノードに移動するとして推移確率行列 \mathbf{G} を設定した。初期確率ベクトル \mathbf{b} に関しては、すべての確率を等しく設定した ($b_1 = \dots = b_V = 1/V$)。なお、各 b_v を $(0, 1)$ 範囲の一樣乱数で設定し、これらの総和が 1 となるように正規化して設定しても、ほぼ同様の実験結果が得られることを確認している。ユーザ移動回数と施設配置数の設定に関しては、前述しているように、施設への広告などアイテムの配置を想定し、総費用一定の下での施設配置問題を考えるので、 MK が一定の条件の下で実験評価した。具体的には、電力網と道路網の双方で、 $M = 1$ のケースでも、 K が全ノード数 V と比較し、ある程度小さくなるよう、基本的には $MK = 1,200$ と設定して評価した。

GL 法の性能評価のための比較法として、その簡易版で遅延評価なしの SG 法、ネットワーク上の K -メディアン問題として施設設置する KM (K-Median) 法、各ノードでの目的関数値 $F(\{v\}; M)$ を降順にソートして上位 K ノードに施設配置する FI (Finite Iteration) 法、推移確率行列 \mathbf{G} より求まる定常確率ベクトル要素を降順にソートして上位 K 要素のノードに施設配置する PRK (PageRank) 法、及び、ベースライン性能を検証するために、ノード集合より K 個をランダムに重複を許さず選定して施設配置する RS (Random Selection) 法を採用した。ここで KM 法の目的関数は $F_{KM}(\mathcal{R}; M) = \sum_{u \in \mathcal{V}} \min_{v \in \mathcal{R}} d(u, v)$ であり、 $d(u, v)$ はネットワーク上での最短パス長とする。なお、本実験では無向グラフを対象とするので、PRK 法をノード度数でのランキング法に帰着させ、度数の等しいノード群の順位は任意に定めた。さらに、性能上限を分析するため、FI 法で求まる施設集合 $\hat{\mathcal{R}}$ より、目的関数値上限を $\bar{F}(M) = \sum_{v \in \hat{\mathcal{R}}} F(\{v\}; M)$ で求めた。 $\bar{F}(M)$ が上限となることは、目的関数のサブモジュラ性より、任意の施設集合 \mathcal{R} に対して、 $F(\mathcal{R}; M) \leq \sum_{v \in \mathcal{R}} F(\{v\}; M) \leq \bar{F}(M)$ より確認できる。

5.2 評価データと実験設定

図 1 には、解品質として、配置施設数 K を変化させたときに、各手法で得られた解(施設配置結果)での目的関数値を示す。ここで図 1 (a)と(b)は、それぞれ電力網と道路網の結果である。なお、SG 法の結果は、GL 法と完全に一致するので省略した。図より、GL 法の結果は、他の手法より両データとも任意の K で優れたことが分かる。詳細には、GL 法の結果が最良で、次いで FI 法または KM 法となった。 K が小さい範囲では KM 法が FI 法より優り、 K が大きくなるとその逆になった。これらの次に PRK 法となるが、いずれの手法でもベースラインの RS 法より優る結果となった。目的関数上限値 $\bar{F}(M)$ との比較では、GL 法については、電力網では平均で目的関数値上限の 87.7%、道路網では上限の 99.0%の解品質を達成している。すなわち、道路網では $\bar{F}(M) \cong \bar{F}(\hat{\mathcal{R}})$ であり、GL 法で厳密に近い解が得られたことが分かる。これに対して電力網では、 $\bar{F}(M)$ と $\bar{F}(\hat{\mathcal{R}})$ の差が GL 法でも比較的に大きいことが見て取れる。

人工ネットワークを用い、問題設定を変化させ、提案 GL

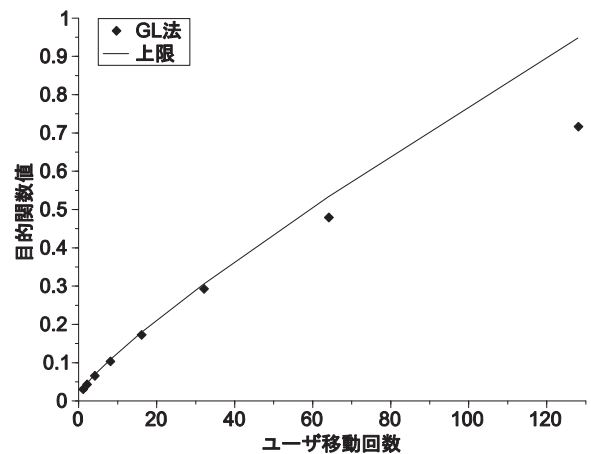
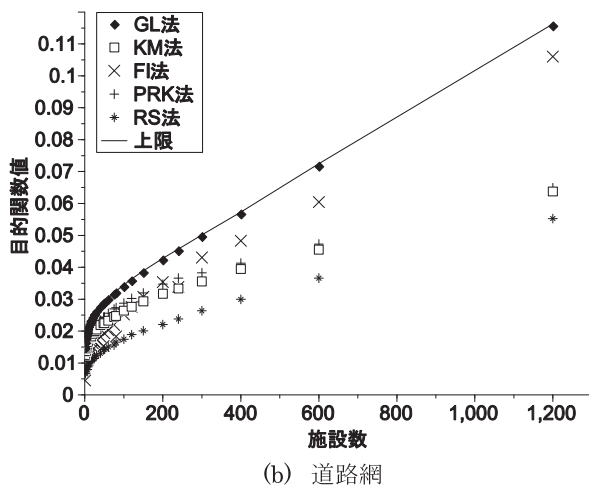
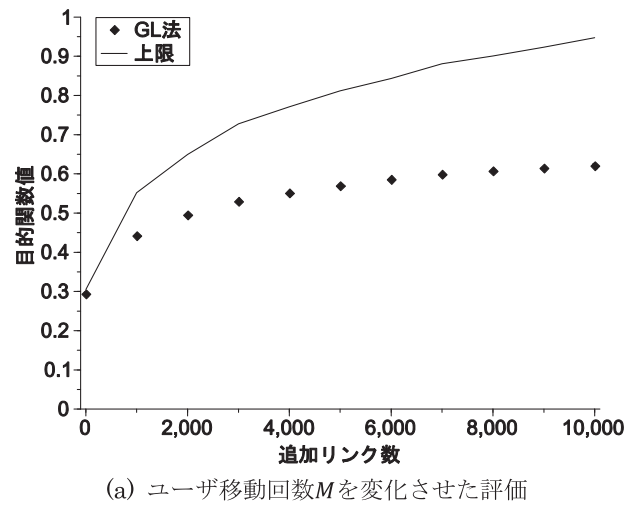
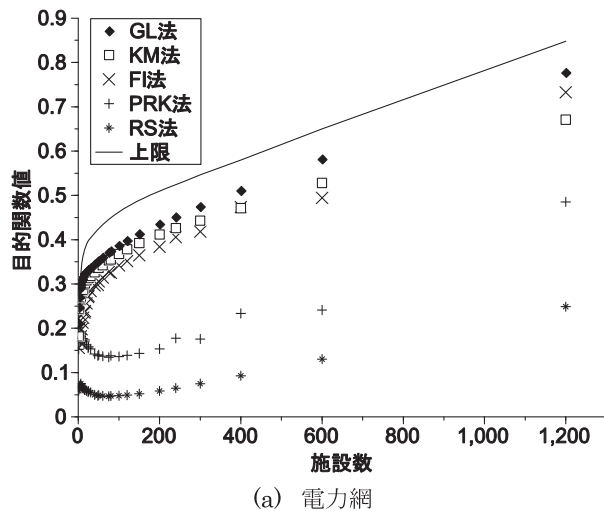


図 1 目的関数値による各手法の解品質比較

図 2 目的関数値による各手法の解品質比較

Fig. 1 Solution quality evaluation by objective function

Fig. 2 Solution quality evaluation using synthetic networks

法で求まる施設集合 $\hat{\mathcal{R}}$ での目的関数値 $F(\hat{\mathcal{R}}; M)$ と上限値 $\bar{F}(M)$ の関係について探求する. まず, 70×70 正方格子の各格子をノードとし, 最隣接ノード間をリンクで結合させた人工ネットワーク(ノード数 4,900 ($=70 \times 70$), リンク数 9,660)を考える. まず, 図 2 (a) には, $K = 75$ に設定し, 移動回数を $M = 2^m$ で変化させた結果を示す. ただし, $m = 0, 1, \dots, 7$. 図より, 移動回数 M が大きくなるに従い, $F(\hat{\mathcal{R}}; M)$ と $\bar{F}(M)$ の差も大きくなるのが分かる. 次に, 図 2 (b) には, 70×70 正方格子網に対し, $K = 75$ かつ $M = 32$ に固定し ($MK = 1,200$), 優先選択[13]に基づきリンクを追加して成長させたネットワークでの結果を示す. ここで, 追加リンク数 L は $L = 0, 1,000, \dots, 10,000$ と変化させ, 図には 10 回の試行での平均値を示している. 図より, 追加リンク数 L が大きくなるに従い, $F(\hat{\mathcal{R}}; M)$ と $\bar{F}(M)$ の差は大きくなるのが分かる. これらの結果より, ネットワーク規模と比較した移動回数 M の大きさ, 優先選択[13]で実現するような次数の高いノードの存在が目的関数値 $F(\hat{\mathcal{R}}; M)$ と上限値 $\bar{F}(M)$ との差に影響すると想定できる. 実際, ノード数 42,716 の道路網において, 次

数は $1, \dots, 6$ の 6 通りで, 最大次数は 6 となり, それぞれの次数となるノード(交差点)数は 1 から順に 7,500, 2,904, 23,378, 8,701, 217, 16 となる. これに対し, ノード数 4,941 の電力網では, 次数 6 以上のノード数は 346 と全ノード数の 7%程度を占め, 最大次数 19 のノードも存在する. よって, 電力網の方が道路網と比較して, 相対的に, M の大きな設定となり, 次数の高いノードも多く存在する.

各手法での施設配置の傾向を直感的に理解するため, 可視化により評価した. 図 3 は, $K = 75, M = 16$ の場合で, 静岡市道路網における各手法での施設配置のうち, 西は安倍川, 南は久能, 東は三保, 清水に至る範囲を図示した. 図 3 より, GL 法は, 例えば, 四角で囲んだ橋(安倍川大橋)のたもと(ポイント 1), 三角で囲んだ半島(三保半島)への入り口(ポイント 2)など通行の要所や交通の流れの集約交差点に施設を配置しているのが分かる. 一方, 他手法では, このよう交差点への施設配置は困難なことが示唆され, それぞれの特徴は以下となる. KM 法は施設群が互いに離れるように配置

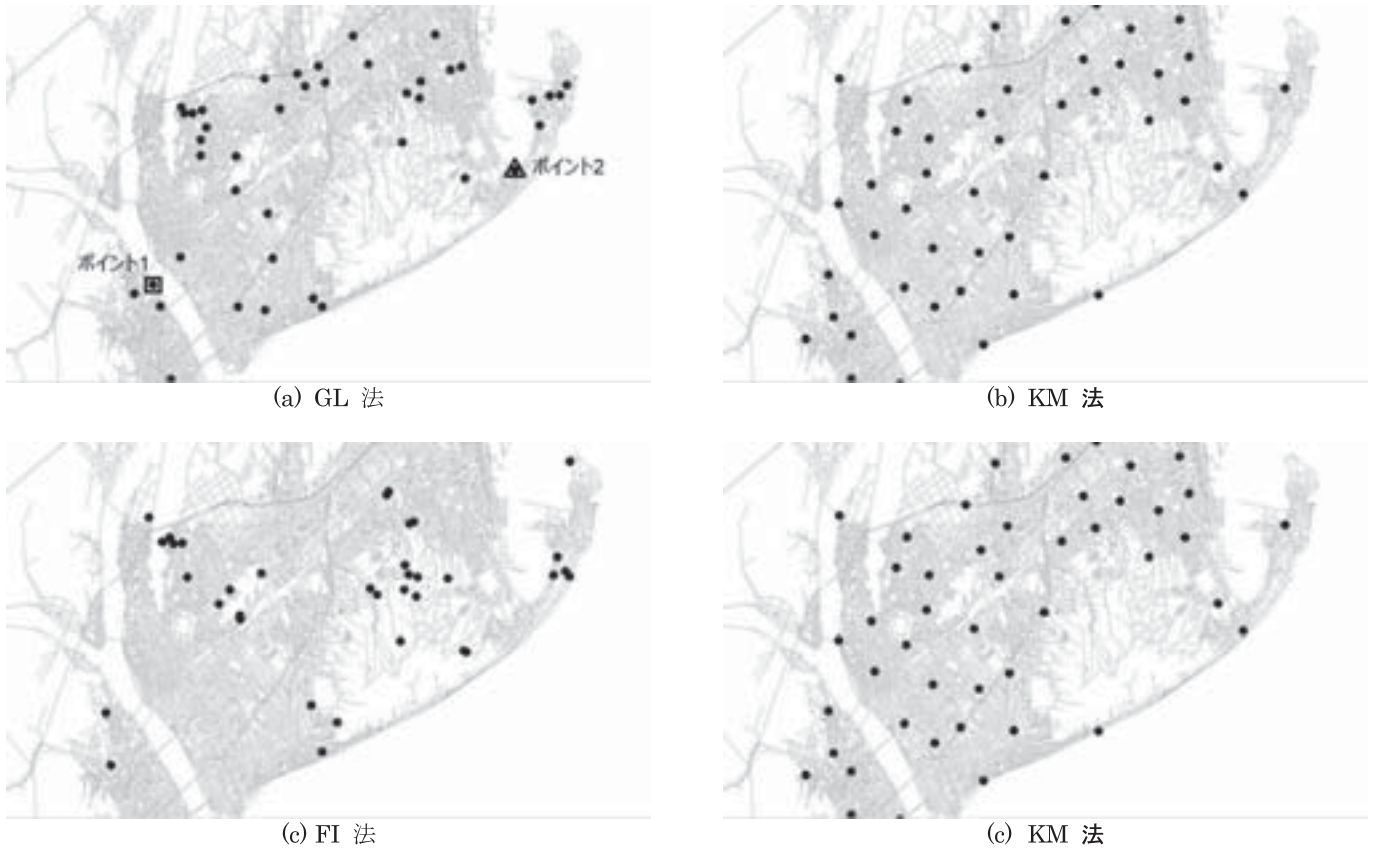


図 3 道路網における各手法の施設配置結果

Fig. 3 Facility location results for road network

している. FI 法は一部の施設群を互いに非常に近くに配置している. PRK 法は, 交差点の次数が大きい順に施設を配置することになり, 図示した $K = 75$ の場合, 前述した次数分布に従い, 次数 6 の交差点がすべてと, 次数 5 の交差点からランダムに選択した配置となる.

図 4 には, 各手法で結果を得るまでに要した計算時間(sec.) を比較する. ただし, 組み合わせ最適化問題を解かない RS 法, PRK 法, および FI 法は比較対象から除外した. 図 4 (a) と (b) は, それぞれ電力網と道路網の結果であり, 各手法のプログラムは C++ 言語で実装し, 実験には Intel Core2 Duo E8500 3.16 GHz CPU の計算機を用いた. 図より, どちらのネットワークへの適用でも, $K = 1$ のケースを除き, 処理効率の面で GL 法は SG 法よりも顕著に優れていることが分かる. すなわち, 逆単調性に基づく遅延評価導入の有効性が実証されたと考える. ここで, 4 節で述べたように, GL 法の主たる計算量は $O(\alpha KMV^2)$ であり, 遅延評価付き貪欲法の効率を定める α は K の増加に伴って減少すると考えられる. 実際に実験結果でも, MK が一定の条件下で, GL 法の計算量は K の増加に伴って大きく減少していることが分かる. 一方, SG 法の主たる計算量は $O(KMV^3)$ であり, MK が一定の条件下なら常に計算時間は一定となっている. 一方, GL 法と KM 法の比較では, 施設配置数 K が多くなるに従い GL 法の計算効率が大幅に優るようになることが分かる. すなわち, 多数施設の配置問題では, 計算効率の面でも GL 法が KM 法より優ることが示唆される.

実験結果をまとめると, GL 法は他の手法より高品質な解を出力することが分かり, 遅延評価を導入した GL 法の計算

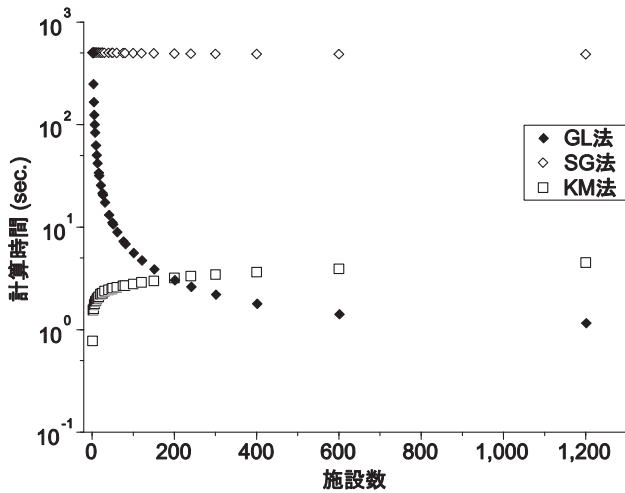
効率は K が大きくなるに従い高くなっている. したがって, 乱歩モデルに基づいた施設配置問題において, GL 法は適切な施設配置を効率良く実現できると考えられ, このような問題を解くための基本法として GL 法は重要な役割を果たすことが期待できる.

6. おわりに

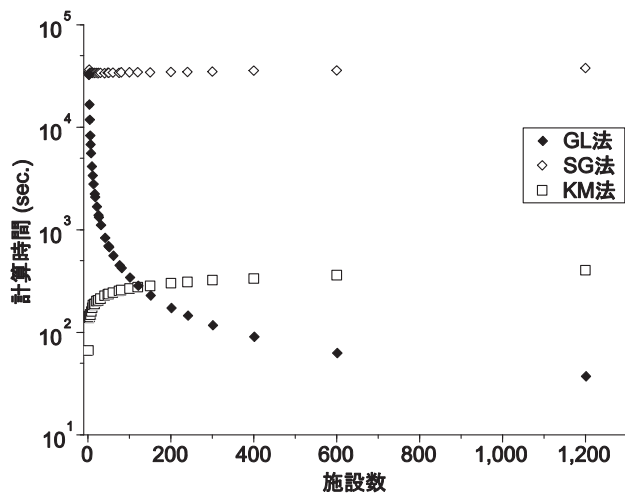
本論文では, 乱歩モデルに基づき行動するユーザを想定した施設配置問題を扱った. すなわち, 与えられた期間内で, 配置された複数施設のどれかにそのユーザが到達する確率を考えて最大化する問題である. 本施設配置問題が NP-困難クラスに属することと, その目的関数がサブモジュラ関数となることに言及し, 目的関数の計算法を示し, 遅延評価を導入した貪欲法に基づく解法を提案した. 電力網と現実道路網のデータを用いた評価実験では, 代表的な従来手法の本施設問題への適用と比較して, 解品質の面で提案法の有効性を確認するとともに, 遅延評価の導入で処理効率も大幅に改善できることを示した. このことより, さらに多様なデータでの評価実験は必要であるものの, 施設配置問題において提案法の有望性が示唆されたと考える. 今後は, より精緻なユーザ行動モデルを探求し, 提案法の有効性を評価する.

[謝辞]

本研究は株式会社豊田中央研究所との共同研究および, 科研費補助金基盤研究(C) (No. 22500133) の補助を受けた.



(a) 電力網



(b) 道路網

図4 処理時間による各手法の効率比較

Fig. 4 Efficiency evaluation by processing time

[文献]

[1] S. Brin and L. Page, "The anatomy of a large scale hypertextual Web search engine," Proc. WWW'98, pp.107-117, (1998).
 [2] D. Kempe and J. Kleinberg and E. Tardos, "Maximizing the spread of influence through a social network," Proc. KDD'03, pp.137-146, (2003).
 [3] M. Kimura, K. Saito, R. Nakano, and H. Motoda, "Extracting Influential Nodes on a Social Network for Information Diffusion," Data Mining and Knowledge Discovery, 20:70-97, (2010).
 [4] ベルンハルトコルテ, イェンスフィーゲン, "組合せ最適化第2版 - 論理とアルゴリズム", シュプリンガー・ジャパン, (2009).
 [5] A. Krause and C. Guestrin, "Near-optimal Nonmyopic Value of Information in Graphical Models," Proc. UAI'05, pp.324-331, (2005).

[6] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," Proc. KDD'07, pp.420-429, (2007).
 [7] 室田一雄, "離散凸解析の考えかた最適化における離散と連続の数理", 共立出版, (2007).
 [8] G. Nemhauser, L. Wolsey, M. Fisher, "An analysis of the approximations for maximizing submodular set functions," Mathematical Programming, 14:265-294, (1978).
 [9] 齊藤和巳, 武藤伸明, 池田哲夫, 入月卓也, 永田大, 伊藤かの子, "遅延評価導入による局所改善クラスタリング法の高速度化," 情報処理学会論文誌, 数理モデル化と応用, 3:62-72, (2010).
 [10] 武藤伸明, 齊藤和巳, 池田哲夫, 永田大, 伏見卓恭, "非線形施設配置問題における貪欲法の遅延評価による高速化 - 複数施設が確率的に貢献する場合 -, " 情報処理学会論文誌, 数理モデル化と応用, 4:1-10, (2011).
 [11] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world' networks," Nature, 393:440-442, (1998).
 [12] 全国デジタル道路地図データベース静岡市版, (2011).
 [13] A-L. Barabasi, R. Albert, "Emergence of Scaling in Random Networks", Science 286:509-512, (1999).

武藤 伸明 Nobuaki MUTOH

静岡県立大学経営情報学部准教授。1995年横浜国立大学大学院工学研究科博士課程後期修了。博士(工学)。電子情報通信学会他会員。離散幾何学, 組合せ最適化に従事。

齊藤 和巳 Kazumi SAITO

静岡県立大学経営情報学部教授。1985年慶応義塾大学理工学部数理科学科卒業。博士(工学)。電子情報通信学会他会員。知能情報学, 発見科学, 複雑ネットワークの科学に従事。

大久保 誠也 Seiya OKUBO

静岡県立大学経営情報学部助教。2005年電気通信大学大学院電気通信学研究科電子情報学専攻博士後期課程修了。博士(工学)。電子情報通信学会他会員。情報セキュリティ, 量子計算に従事。

池田 哲夫 Tetsuo IKEDA

静岡県立大学経営情報学部教授。1981年東京大学大学院理学系研究科情報科学専攻修士課程修了。博士(工学)。電子情報通信学会他会員。データ工学, 情報検索, GIS(地理情報システム)に従事。