日本オペレーションズ・リサーチ学会
2021年 秋季研究発表会

制御可能な不確実性を含むStochastic Matchingの確率的挙動の最適化

05000754 日本電信電話株式会社 *引間友也 HIKIMA Yuya
日本電信電話株式会社 赤木 康紀 AKAGI Yasunori
日本電信電話株式会社 金 秀明 KIM Hideaki
日本電信電話株式会社 幸島 匡宏 KOHJIMA Masahiro
日本電信電話株式会社 倉島 健 KURASHIMA Takeshi
日本電信電話株式会社 戸田 浩之 TODA Hiroyuki

1. はじめに

不確実性を含むグラフにおいて最適なマッチングを発見する研究は，広告の割当 [1] や腎臓交換 [2] などの応用をもつため，数多くなされてきた．これらの研究では，グラフに含まれる不確実性に関するパラメータは所与のものとして与えられることが多い．例えば，広告割当では，web 閲覧者に該当するノードは所与の確率分布から次々と生じるものとして扱われる．しかしながら，多くの応用では不確実性のパラメータは制御可能である．本研究では，この設定のもと，得られるマッチングの重みの期待値を最大化する新しい最適化問題を定式化し，その近似解法を与える．提案する最適化問題と解法は，クラウドソーシングや交通サービスにおける価格最適化など，複数の重要な応用をもつ．

2. 最適化問題の定式化

本研究では，次のような状況を考える (図 1)．参加者の集合を U，資源の集合を V とし，それらをノードとしてもつ重み付き二部グラフ $G=(U,V,E)$ を考える．ここで，$(u,v)\in E$ であるとき，u と v はマッチング可能である．また，w_{uv} を枝 (u,v) の重みであるとする．これらの設定の下，サービス提供者は各参加者に対して x_u を設定する．それに対して，各参加者 u は確率 $p_u(x_u)$ で市場にそのまま参加し，$1-p_u(x_u)$ の確率で市場から去る．グラフ $G'=(U',V,E')$ を，去った参加者に該当するノードを取り除いた二部グラフであるとする．このとき，サービス提供者は以下の最適化問題を解くことで，マッチング $M\subset E'$ を決定し，各 $(u,v)\in M$ について利益 $w_{uv}+x_u$ を得る．

$$
\begin{aligned}
(\mathrm{P_{sub}})\quad \max_{\boldsymbol{z}}\quad & \sum\nolimits_{(u,v)\in E}(w_{uv}+x_u)\cdot z_{uv} \\
\text{s.t.}\quad & \sum\nolimits_{v\in\delta(u)} z_{uv}\le a_u,\quad \forall u\in U \\
& \sum\nolimits_{u\in\delta(v)} z_{uv}\le 1,\quad \forall v\in V \\
& z_{uv}\in\{0,1\},\quad \forall (u,v)\in E
\end{aligned}
$$

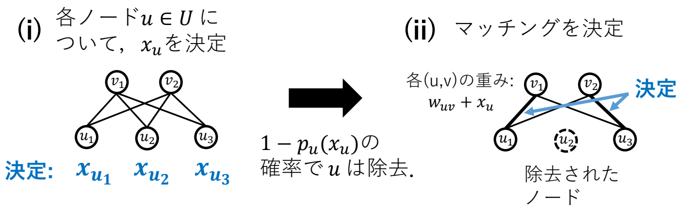

図 1: マッチングのプロセス

但し，$\delta(\xi)$ はノード ξ に隣接するノード集合を指す．また，$a_u\in\{0,1\}$ は $a_u=0$ のとき参加者 u が市場に不参加，$a_u=1$ のとき参加者 u が市場に参加したことを表す変数である．決定変数 $z_{uv}\in\{0,1\}$ は $z_{uv}=1$ のとき，(u,v) をマッチングすることを表す．

上記のような状況設定は，クラウドソーシングなどにおいて，各ワーカの賃金（x_u）を決定したうえで，その賃金を承諾したワーカ ($\{u\in U\mid a_u=1\}$) に対してタスク (v) を割り当てる場合などが当てはまる．

このとき，本研究で取り組む最適化問題は以下である．

$$
(\mathrm{P})\qquad \max_{\boldsymbol{x}\in\mathbb{R}^{|U|}}\ \mathbb{E}_{\boldsymbol{a}\sim D(\boldsymbol{x})}\Big[\max_{\boldsymbol{z}\in \boldsymbol{Z}(\boldsymbol{a})} f(\boldsymbol{x},\boldsymbol{z})\Big]
$$

但し，$f(\boldsymbol{x},\boldsymbol{z})$ と $Z(\boldsymbol{a})$ は，それぞれ $(\mathrm{P_{sub}})$ の目的関数値と実行可能領域である．$D(\boldsymbol{x})$ は $\{0,1\}^{|U|}$ 上の確率分布であり，その確率質量関数は $\Pr(\boldsymbol{a}\mid\boldsymbol{x})=\prod_{u\in U}\left\{p_u(x_u)^{a_u}(1-p_u(x_u))^{(1-a_u)}\right\}$ である．この最適化問題は，最終的に得られるマッチングの重みの期待値を $\boldsymbol{x}$ について最大化するものである．

(P) は応用上重要であるものの，以下の理由で効率的に解くことが難しい．

(i) 確率により生起する承諾結果 $\boldsymbol{a}$ の場合の数が 2^n 個あり，目的関数値や勾配を正確に計算するには 2^n 個の重み付き二部グラフマッチングを解く必要がある．

(ii) 応用上，関数 $p_u(x)$ として非凸な関数を用いることが多く，その場合目的関数も非凸になってしまう．

3. 関数 p_u に対する仮定

本研究では以下の仮定を関数 p_u に対しておく．

仮定 1. $p_u(x)$ は x に関して全単射かつ単調減少である．また，Monotone hazard rate function [3] である．すなわち， $-p'_u(x)/p_u(x)$ が x に関して単調非減少である．さらに，$p_u(x)=0$ or $\lim_{x\to\infty} p_u(x)=0$ となる x が存在する．

上記の仮定は，機械学習でよく用いられる確率分布である正規分布や指数分布の相補累積分布関数を含むような緩い仮定である．

4. 提案手法

以下の最適化問題の最適値を $\hat{f}(\boldsymbol{x})$ とする．

$$
\begin{aligned}
\max_{\boldsymbol{z}} \quad & \textstyle\sum_{(u,v)\in E}(x_u + w_{uv})\cdot z_{uv} \\
\text{s.t.} \quad & \textstyle\sum_{v\in\delta(u)} z_{uv} \le p_u(x_u), \quad \forall u \in U \\
& \textstyle\sum_{u\in\delta(v)} z_{uv} \le 1, \quad \forall v \in V \\
& 0 \le z_{uv} \le 1, \quad \forall (u,v)\in E
\end{aligned}
$$

このとき，以下の定理が成り立つ．

定理 1 任意の $\boldsymbol{x}$ について，以下が成り立つ．

$$1/3\cdot\hat{f}(\boldsymbol{x}) \le \mathbb{E}_{\boldsymbol{a}\sim D}[\max_{\boldsymbol{z}\in\boldsymbol{Z}(\boldsymbol{a})} f(\boldsymbol{x},\boldsymbol{z})] \le \hat{f}(\boldsymbol{x})$$

ここで，(P) の目的関数 $\mathbb{E}_{\boldsymbol{a}\sim D}[\max_{\boldsymbol{z}\in\boldsymbol{Z}(\boldsymbol{a})} f(\boldsymbol{x},\boldsymbol{z})]$ を $\hat{f}(\boldsymbol{x})$ に置き換えると以下の最適化問題となる．

$$
\begin{aligned}
\text{(PA)} \quad \max_{\boldsymbol{x},\boldsymbol{z}} \quad & \textstyle\sum_{(u,v)\in E}(x_u + w_{uv})\cdot z_{uv} \\
\text{s.t.} \quad & \textstyle\sum_{v\in\delta(u)} z_{uv} \le p_u(x_u), \quad \forall u \in U \\
& \textstyle\sum_{u\in\delta(v)} z_{uv} \le 1, \quad \forall v \in V \\
& 0 \le z_{uv} \le 1, \quad \forall (u,v)\in E \\
& \boldsymbol{x} \in \mathbb{R}^{|U|}
\end{aligned}
$$

(PA) の最適解は定理 1 より (P) の $\frac{1}{3}$-近似解となる．但し，(PA) には制約に非凸関数 $p_u(x_u)$ が含まれており，内点法などを用いて直接解くことはできない．

そこで，(PA) を仮定 1 の下で最小凸費用流問題に帰着することを考える．以下の最適化問題を与える．

$$
\begin{aligned}
\text{(PA}'\text{)} \quad \max_{\boldsymbol{z}} \quad & \sum_{u\in U} p_u^{-1}\Big(\sum_{v\in\delta(u)} z_{uv}\Big)\sum_{v\in\delta(u)} z_{uv} + \sum_{(u,v)\in E} w_{uv} z_{uv} \\
\text{s.t.} \quad & \sum_{v\in\delta(u)} z_{uv} \in S_u, \quad \forall u\in U \\
& \sum_{u\in\delta(v)} z_{uv} \le 1, \quad \forall v \in V \\
& 0 \le z_{uv} \le 1, \quad \forall (u,v)\in E
\end{aligned}
$$

但し，S_u は p_u の値域である．この最適化問題は (PA) の決定変数 $\boldsymbol{x}$ を，$x_u := p_u^{-1}(\sum_{v\in\delta(u)} z_{uv})$ の等式によって消去したものである．

このとき，以下の定理 2 が成り立つ．

定理 2 仮定 1 が成り立つとする．(PA$'$) の最適解を $\boldsymbol{z}^*$ とし，各 $u\in U$ について $x_u^* := p_u^{-1}(\sum_{v\in\delta(u)} z_{uv}^*)$ とする．このとき，$(\boldsymbol{x}^*,\boldsymbol{z}^*)$ は (PA) の最適解である．

定理 2 より，(PA$'$) を解けば (PA) の解が求まる．

ここで，新たな添字 s と t を導入する．各 $u\in U$ について，$z_{su} := \sum_{v\in\delta(u)} z_{uv}$ とし，各 $v\in V$ について，$z_{vt} := \sum_{u\in\delta(v)} z_{uv}$ とする．さらに，スラック変数 z_{st} を導入し，$n := \min\{|U|,\ |V|\}$ とする．このとき，(PA$'$) は以下のように書ける．

$$
\begin{aligned}
\text{(FP)} \quad \min_{\boldsymbol{z}} \quad & \sum_{u\in U} -p_u^{-1}(z_{su})\cdot z_{su} - \sum_{(u,v)\in E} w_{uv}\cdot z_{uv} \\
\text{s.t.} \quad & \sum_{u\in U} z_{su} + z_{st} = n, \ \sum_{v\in V} z_{vt} + z_{st} = n \\
& z_{su} - \sum_{v\in\delta(u)} z_{uv} = 0, \quad \forall u\in U \\
& \sum_{u\in\delta(v)} z_{uv} - z_{vt} = 0, \quad \forall v\in V \\
& z_{su}\in S_u, \quad \forall u \in U \\
& 0 \le z_{vt} \le 1, \quad \forall v\in V \\
& 0 \le z_{uv} \le 1, \quad \forall (u,v)\in E \\
& 0 \le z_{st} \le n
\end{aligned}
$$

これは $U\cup V\cup\{s,t\}$ をノードとしてもつ最小費用流問題であり，仮定 1 の下では最小凸費用流問題となる．よって既存解法 [4, 5] を用いて解くことができる．

5. おわりに

本研究では，不確実性のパラメータを制御することができる設定のもと，得られるマッチングの重みの期待値を最大化する新しい最適化問題を定式化した上で，その $\frac{1}{3}$-近似解法を与えた．定理の証明や数値実験の結果等については，当日の発表において説明を行う．

参考文献

[1] Aranyak Mehta. Online matching and ad allocation. *Theoretical Computer Science*, Vol. 8, No. 4, p. 265–368, 2012.

[2] Ning Chen, Nicole Immorlica, Anna R Karlin, Mohammad Mahdian, and Atri Rudra. Approximating matches made in heaven. In *ICALP*, pp. 266–278, 2009.

[3] Richard E Barlow, Albert W Marshall, and Frank Proschan. Properties of probability distributions with monotone hazard rate. *The Annals of Mathematical Statistics*, Vol. 34, No. 2, pp. 375–389, 1963.

[4] Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. *Network Flows: Theory, Algorithms, and Applications.* Prentice-Hall, 1993.

[5] László A Végh. A strongly polynomial algorithm for a class of minimum-cost flow problems with separable convex objectives. *SIAM Journal on Computing*, Vol. 45, No. 5, pp. 1729–1761, 2016.