数理解析研究所講究録
第 1649 巻 2009 年 160-164

Metropolis Walkのcover timeにおけるタイトな上界

野中 良哲 * 小野 廣隆 † 定兼 邦彦 † 山下 雅史 †
* 九州大学大学院システム情報科学府 † 九州大学大学院システム情報科学研究院

概要

グラフ上のランダムウォークとは，グラフ上の頂点をトークンがランダムに遷移していくモデルである．全ての隣接点へ等確立で遷移する標準ランダムウォークでは，頂点数 n のグラフに対して，hitting timeと cover timeがそれぞれ $O(n^3)$ であることが知られている．本稿では，標準ランダムウォークに Metropolis-Haistings アルゴリズムを適用して得られるランダムウォークである Metropolis walk の hitting time と cover time がそれぞれ $O(fn^2)$ および $O(fn^2 \log n)$ であることを示す．ここで，$f = \max_{u,v\in V} \pi_u/\pi_v$ である．

1 準備

グラフ $G = (V, E)$ 上のランダムウォークとは, グラフ上の頂点をトークンが以下のようなプロセスで遷移するモデルである.

1. ある頂点 $v_0 \in V$ から出発する.
2. v_0 の隣接点 $v_1 \in N(v_0)$ をランダムに選択し, v_1 に移動する.これを 1ステップとする.
3. v_1 に対しても 2と同様に隣接する頂点に移動し, 以下同様に繰り返す.

すべての隣接点に等確率で遷移するランダムウォークを標準ランダムウォークと呼ぶ．標準ランダムウォークの遷移確率行列 $P = (p_{uv})_{u,v\in V}$ は以下のように定義される.

定義 1.1 (標準ランダムウォーク)

$$p_{uv} = \begin{cases} \frac{1}{\deg(u)} & u \in N(u), \\ 0 & otherwise. \end{cases}$$

通常，単にランダムウォークと言えば標準ランダムウォークのことを指す.

すべての頂点 $u \in V$ に対して，

$$\pi_u = \sum_{v\in V} \pi_v p_{vu} \tag{1}$$

を満たす確率分布 $\pi = (\pi_u)_{u\in V}$ を定常分布と呼ぶ.

また，ランダムウォークの速さを表す指標として，hitting time と cover time がある.

定義 1.2 (hitting time) グラフ $G = (V, E)$ におけるランダムウォークで, 頂点 $u \in V$ から出発した粒子が遷移確率行列 P に従ってランダムに移動し, 頂点 $v \in V$ に初めて到達するまでに要する遷移数の期待値を u から v への hitting timeとして $H_G^P(u, v)$ と表す. また, グラフ G の hitting time(H_G^P) を以下のように定義する.

$$H_G^P = \max_{u,v\in V} H_G^P(u, v).$$

定義 1.3 (cover time) グラフ $G = (V, E)$ におけるランダムウォークで, 頂点 $u \in V$ から出発した粒子遷移確率行列 P に従ってランダムに移動し, 頂点集合 $U \subseteq V$ を全て訪問するまでに要する遷移数の期待値を U に対する u からの cover timeとして $C_G^P(U; u)$ と表す. このとき G の cover time(C_G^P) を以下のように定義する.

$$C_G^P = \max_{u\in V} C_G^P(V; u).$$

2 既存研究

2.1 hitting time, cover time に関する結果

標準ランダムウォークについては hitting time および, cover time がともに $O(n^3)$ であることが分かっている [1, 2]. また, この解析は図 1 のような lollipop グラフと呼ばれるグラフにおいてタイトなものであることが知られている [3]. また, hitting time および cover time に関して,

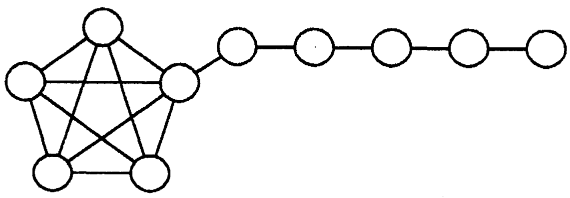

図 1: Lollipop グラフ

以下の定理と補題が知られている.

定理 2.1 [6] 任意のグラフ $G = (V, E)$ と遷移確率行列 P および頂点集合 $U \subseteq V$ について, 以下が成り立つ.

$$C_G^P(U) \geq h(|U|-1) \min_{u,v\in U} H_G^P(u,v),$$
$$C_G^P(U) \leq h(|U|-1) \max_{u,v\in I} H_G^P(u,v).$$

ただし, $h(n)$ は n 番目の調和関数であり, $h(n) = \sum_{i=1}^{n} i^{-1}$ である.

補題 2.2 [5] グラフ $G = (V, E)$ において, 任意の遷移確率行列 P および, 任意の $u \in V$ とその隣接頂点 $v \in N(u)$ に関して $H_G^P(u,v) \leq (p_{v,u}\pi_v)^{-1}$ が成り立つ.

2.2 隣接頂点の次数情報を用いるランダムウォーク

池田らは標準ランダムウォークとは異なる遷移確率を定義することで, hitting time および cover time を小さくする試みとして, 隣接頂点の次数情報を用いるランダムウォークを提案した [5]. 池田らの提案したランダムウォークでは頂点 u からその隣接点 $v \in N(u)$ への遷移確率が以下のように定義される.

$$p_{uv}^{\beta} = \frac{\exp[-\beta U(u,v)]}{\sum_{w\in N(u)} \exp[-\beta U(u,w)]}.$$

ここで β は実数とする. また, $U(u,w)$ は以下のように定義される potential function である.

$$U(u,v) = U(v) = \log \deg(v),$$
$$v \in N(u), u \in V.$$

このランダムウォークでは $\beta = \frac{1}{2}$ のとき, 任意のグラフに対して hitting time が $O(n^2)$, cover time が $O(n^2 \log n)$ であることが示されている.

3 Metropolis walk の hitting time と cover time

3.1 Metropolis walk

Metropolis-Haistings アルゴリズムは, マルコフ連鎖モンテカルロにおいて任意の定常分布を実現する遷移確率行列を設計する手法の一つである [4]. 今, 頂点 u から v への遷移確率を p_{uv} とすると, 任意の 2 頂点について

$$p_{uv}\pi_u = p_{vu}\pi_v \tag{2}$$

を満たすとき, $\pi = (\pi_u)_{u\in V}$ は $P = (p_{uv})_{u,v\in V}$ の定常分布であるしかし, 与えられた P が, 必ずしも求める π について, 式 (2) を満たしているとは限らない. そこで, Metropol-Hastings アルゴリズムでは P を基にして, 新しい遷移確率行列 $P' = (p'_{uv})_{u,v\in V}$ を以下のように定義する.

$$p'_{uv} = \begin{cases} p_{uv} \min\left(\frac{\pi_v p_{vu}}{\pi_u p_{uv}}, 1\right) & v \in N(u), \\ 1 - \sum_{x\in N(u)} p_{ux} & u = v, \\ 0 & otherwise. \end{cases} \tag{3}$$

今，$p_{uv}\pi_w > p_{vu}\pi_v$ の場合を考える．このとき，

$$\begin{aligned}\pi_u p'_{uv} &= \pi_u p_{uv} \min\left(\frac{\pi_v p_{vu}}{\pi_u p_{uv}}, 1\right)\\ &= \pi_v p_{vu}\\ &= \pi_v p'_{vu}\end{aligned}$$

となり，確かに式 (2) の条件を満たしている．逆の場合も同様に式 (2) を満たす．

標準ランダムウォークに対して Metropolis-Hastings アルゴリズムを適用して得られるランダムウォークを Metropolis walk と呼ぶ．その遷移確率行列 $P^* = (p_{uv})_{u,v\in V}$ は以下の式で定義される．

定義 3.1 (Metropolis walk)

$$p'_{uv} = \begin{cases} \frac{1}{\deg(u)} \min\left(\frac{\pi_v \deg(u)}{\pi_u \deg(v)}, 1\right) & v \in N(u),\\ 1 - \sum_{x\in N(u)} p'_{u,w} & v = u,\\ 0 & otherwise. \end{cases}$$

3.2 上界

Metropolis walk の hitting time と cover time の上界を示す．

補題 3.2 [5] 異なる 2 頂点 $u, v \in V$ において,u から v への最短路 $u = v_0, v_1, \cdots, v_l = v$ に対して以下が成り立つ．

$$\sum_{i=0}^{t} \deg(v_i). \le 3n \tag{4}$$

証明　すべての $1 \le i < i+2 < j \le l$ に関して，

$$(N(v_i) \cup \{v_i\}) \cap (N(v_j) \cup \{v_j\}) = \emptyset$$

が成り立つ．仮に頂点 $w \in N(v_i) \cap N(v_j)$ が存在したとすると，パス $v_0, v_1, \cdots, v_i, w, v_j, \cdots, v_l$ がさらに短い $u-v$ パスとなりなり，矛盾する．よって u, v 以外の頂点は，それぞれ高々3 回しかカウントされず，u, v は丁度一回カウントされるので，式 (4) が成立する．　□

定理 3.3 任意のグラフ G および定常分布 π に対して定義される Metropolis walk の遷移確率行列 P^* について，

$$\begin{aligned} H_G^{P^*} &= O(n^2 f),\\ C_G^{P^*} &= O(n^2 f \log n). \end{aligned}$$

である．ただし，$f = \max_{u,v\in V} \pi_u/\pi_v < \infty$.

証明　定義 3.1 より，任意の $u \in V$ と $v \in N(u)$ について

$$p_{vu} = \min\left(\frac{\pi_u}{\pi_v \deg(u)}, \frac{1}{\deg(v)}\right) \tag{5}$$

が成り立つ．これに補題 2.2 用いると，

$$\begin{aligned} H_{uv} &\le \frac{1}{\pi_v} \max\left(\frac{\pi_v \deg(u)}{\pi_u}, \deg(v)\right)\\ &= \max\left(\frac{\deg(u)}{\pi_u}, \frac{\deg(v)}{\pi_v}\right)\\ &\le fn \max(\deg(u), \deg(v)) \end{aligned} \tag{6}$$

となる．ここで，u から v への最短路を $u = v_1, \ldots, v_l = v$ とすると，

$$H(u, v) \le \sum_{i=0}^{l-1} H(v_i, v_{i+1}) \tag{7}$$

である． 式 (6),(7) および補題 3.2 より，

$$H(u, v) \le 6fn^2 \tag{8}$$

となる．また, 式 (8) 及び定理 2.1 から，

$$C_G^Q \le h_{n-1} 6fn^2 \tag{9}$$

が得られる．　□

Metropolis-Hastings アルゴリズムにより，定常分布は任意に設定することができるので，f が定数となるような定常分布を選べば，池田らの提案したランダムウォークと同等の上界を得ることができる．

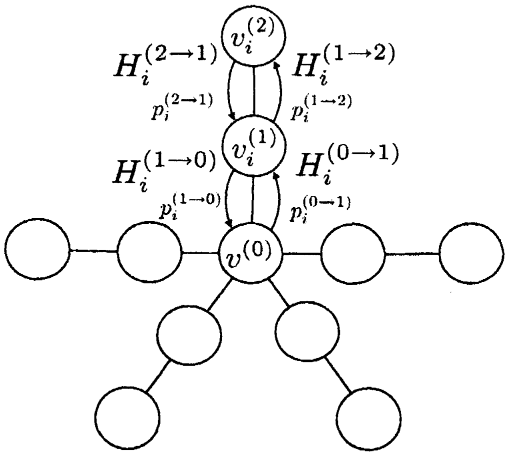

図 2: Glitter star グラフ

3.3　下界

Star グラフの各辺に新たに一つづつ頂点を挿入したグラフを glitter star グラフと呼ぶ.

定理 3.4 *Glitter star* グラフ上での *Metropolis walk* において，*hitting time* と *cover time* がそれぞれ $\Omega(fn^2), \Omega(fn^2\log n)$ となる定常分布 π が存在する．ただし，$f = \max_{u,v} \pi_u/\pi_v < \infty$

証明　Gliter star グラフ S に対する任意の遷移確率行列 P について，遷移確率と hitting time をそれぞれ次のように表す．各 i について隣接頂点への遷移確率及び hitting time を，

$$
\begin{aligned}
p_{v^{(0)}v_i^{(1)}} &= p_i^{(0\to1)}, & H_S^P(v^{(0)}, v_i^{(1)}) &= H_i^{(0\to1)},\\
p_{v_i^{(1)}v^{(0)}} &= p_i^{(0\to1)}, & H_S^P(v_i^{(1)}, v^{(0)}) &= H_i^{(0\to1)},\\
p_{v_i^{(1)}v_i^{(2)}} &= p_i^{(1\to2)}, & H_S^P(v_i^{(1)}, v_i^{(2)}) &= H_i^{(1\to2)},\\
p_{v_i^{(2)}v_i^{(1)}} &= p_i^{(2\to1)}, & H_S^P(v_i^{(2)}, v_i^{(1)}) &= H_i^{(2\to1)},
\end{aligned}
$$

とする．また, 各頂点の定常分布を $\pi_{v^{(0)}} = \pi^{(0)}$, $\pi_{v_i^{(1)}} = \pi^{(1)}$，$\pi_{v_i^{(2)}} = \pi^{(2)}$ とする (図 2 参照)．ここで，$\pi^0 = \pi_{\min}$，$\pi_i^{(1)} = \pi_i^{(2)} = \pi_{\max}$ とすると，それぞれの hitting time は以下の式で表される．

$$
\begin{aligned}
H_i^{(0\to1)} &= \frac{n-1}{2}(f(n-3)+1),\\
H_i^{(1\to0)} &= f(n-1),\\
H_i^{(1\to2)} &= 2\left(n-2+\frac{1}{f}\right),\\
H_i^{(2\to1)} &= 2.
\end{aligned}
$$

よって, 任意の $i \neq j$ について,

$$
\begin{aligned}
&H_S^{P^*}(v_i^{(2)}, v_j^{(2)})\\
&= \frac{n-1}{2}(f(n-3)+1) + f(n-1) + 2\left(n-2+\frac{1}{f}\right) + 2\\
&= \Theta(fn^2)
\end{aligned}
$$

となる．つぎに, $U = V \setminus (\{v^{(0)}\} \cup \{v_i^{(1)} : i = 1, 2, \cdots, \ell\})$ とすれば定理 2.1 より，

$$
C_S^{P^*}(U) \geq h\left(\frac{n-1}{2}\right) H_S^{P^*}(v_i^{(2)}, v_j^{(2)}),
$$

なので，$C_S^{P^*} \geq C_S^{P^*}(U) = \Omega(fn^2\log n)$ が得られる．　□

これは定理 3.3 の上界に対するタイトな例となっている

さて，一般にある定常分布 π を満たす遷移確率行列はいくつも考えらる．しかし，パスグラフ (図 3) については次の定理が成り立つ．

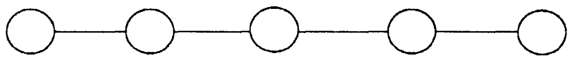

図 3: パスグラフ

定理 3.5 パスグラフ T に対して，ある定常分布 π を満たす遷移確率行列の集合 $\mathcal{P}$ を考える．このとき，任意の $P \in \mathcal{P}$ について $H_T^P = \Omega(fn^2)$, $C_T^P = \Omega(fn^2)$ となる π が存在する．

証明　まず，n 個の頂点を持つパスグラフ T 上でのランダムウォークの *hitting time* に関して以下が成り立つ．

$$
H_T^P \geq \frac{1}{2}\sum_{v\in V} \pi_v^{-1} - n.
$$

また，

$$\sum_{v\in V}\pi_v^{-1}=\sum_{v\in V}\left(\pi_v^{-1}\sum_{w\in V}\pi_w\right)$$
$$=\sum_{(v,w)\in V\times V}\frac{\pi_w}{\pi_v}$$

である．ここで，定常分布が $\lfloor\frac{n}{2}\rfloor$ 個の頂点で $\pi_{\max}$ であり，$\lceil\frac{n}{2}\rceil$ 個の頂点で $\pi_{\min}$ とすると，$2H_T^P\geq C_T^P=\Omega(n^2f)$ が得られる． □

4 まとめ・今後の課題

本稿では，標準ランダムウォークにたMetropol-Hastings アルゴリズムを用いたMetorpolis walk について，hitting time が $O(n^2f)$，cover time が $O(n^2f\log n)$ となることを示した．ここで，$f=\pi_{\max}/\pi_{\min}$ である．もし，f を定数にすれば，標準ランダムウォークよりも高速なランダムウォークが実現できる．また，パスグラフについては，いかなる遷移確率行列を用いても hitting time と cover time が $\Omega(fn^2)$ となる定常分布が存在する．これは,Metropolis walk が hitting time に関しては任意の定常分布を実現るランダムウォークとして最適なものの一つであることを意味する．一方、任意の定常分布を実現するランダムウォークの中で cover time が $O(fn^2)$ となるものが必ず存在するのか, もし存在するならばどのようなランダムウォークであるかといったことは未だ明らかになっていない．

参考文献

[1] R.Aleliunas, R.M Karp, R.J. Lipton, L.Lovaasz, and C.Rackoff, "Random walks, universal traversal sequences, and the complexity of maze problems", Proc. 20th IEEE Ann.Symposium on Foundations of Computer Science, 218-223, 1979.

[2] D.J.Aldous, "On the time taken by random walks on finite groups to visit every state", Z.Wahrsch. verw. Gebiete 62 361-1983.

[3] G. Brightwell and P. Winkler, "Maximum hitting time for random walks on graphs," *Journal of Random Structures and Algorithms,* 3, 263–276, 1990.

[4] W.K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications", *Biometrika,* 57, 97–109,

[5] Satoshi Ikeda, Izumi Kubo, Norihiro Okumoto, Masafumi Yamashita "Impact of Local Topological Information on Random Walks on Finite Graphs", Proceedings of Thirtieth International Colloquium on Automata,Languages and Programming, 1054-1067,2003.

[6] P.Matthews, "Covering Problems for Markov Chain", The Annals of Probability Vol.16, No3, 1215-1228, 1988.