

ランダムグラフ上の多種ランダムウォークの全訪問時間

穂坂 祐輔* 山内 由紀子†

来嶋 秀治† 小野 廣隆† 山下 雅史†

* 九州大学システム情報科学府 情報学専攻

† 九州大学システム情報科学研究所 情報学部門

‡ 九州大学経済学研究院

概要

概要 ランダムウォークはインターネットのような巨大なネットワークの探索に有効な手段である。全訪問時間はグラフ上のランダムウォークの重要な指標の一つであり数多くの研究がされている。ネットワークの探索において複数のクローラを使うほうが1台のクローラで探索をするよりも速くネットワーク全体を探索できることは明らかである。Alon ら (2011) は k 個のトークンを使う多重ランダムウォークが完全グラフやランダムグラフなどの特定のグラフにおいて1台のものよりも k 倍速く全探索できることを示したが、同時にサイクルやパスなどのグラフでは $\log k$ 倍にしか高速化できないことも示している。

本研究では k 個のトークンが独立に個々の遷移確率行列に従って遷移する k トークンの多種ランダムウォークを提案する。また単一トークンのランダムウォークにおける Matthews の不等式のように、多種ランダムウォークにおいて全訪問時間の上下界を到達時間から与える不等式を導く。さらに導出した上下界の式が、完全グラフや完全二部グラフ、さらにはランダムグラフにおいてタイトになることを示す。

1 はじめに

ランダムウォークはネットワーク探索の強力な実用的な手法の一つであり、局所的な情報のみから探索を実行できネットワーク全体の情報を必要としないため、特にインターネットのような巨大なネットワークに対して効果的である。有限グラフ上のランダムウォークの指標に到達時間と全訪問時間がある。前者はトークンが任意の頂点に到達するまでの期待遷移数で、後者は全ての頂点

を1回以上訪問するまでの期待遷移数であり、これらに関して多くのことが知られている。標準ランダムウォークとは隣接する頂点に等確率で遷移していくランダムウォークであり、Aleliunas ら [2] らにより、頂点数 n 、辺数 m の任意の (単純) グラフにおいて到達時間と全訪問時間の期待値が $O(mn)$ で抑えられるということが知られている。ロリポップグラフはこの上界のタイトな例になっており、標準ランダムウォークの到達時間、全訪問時間が共に $\Omega(n^3)$ である。到達時間、全訪問時間を短縮するために、Ikeda ら [7] は β ランダムウォークを提案している。 β ランダムウォークの頂点 u からの隣接頂点 v への遷移確率は次のように定義されている。

$$p_{u,v} = \frac{\deg(v)^{-\beta}}{\sum_{v' \in N(u)} \deg(v')^{-\beta}} \quad (1)$$

ここで $N(u)$ は頂点 u の隣接頂点集合、 $\deg(u)$ は頂点 u の次数であり、 β は任意の実数である。また Ikeda らは任意の (単純) グラフにおいて、 $\beta = 0.5$ のときに β ランダムウォークの到達時間の期待値が $O(n^2)$ 、全訪問時間の期待値が $O(n^2 \log n)$ になることを示している。

一方でネットワークの探索を高速化させる別の手段として複数のクローラで探索を並列化することが考えられる。Alon ら [3] は標準ランダムウォークをする k 個の独立なトークン (ただし同期して遷移) からなる多重ランダムウォークについて研究している。多重ランダムウォークでは到達時間を、特定の頂点に1つ以上のトークンが到達するまでの時間としており、全訪問時間を、全ての頂点が1つ以上のトークンに訪問されるまでの時間としている。Alon らは k 個のトークンによる多重ランダムウォークが、完全グラフやランダムグラフなどの特定のグラフにおいて単一 (トークンの) ランダムウォークより k 倍速く全訪問できることを示している。また同時にサイクルやパスのようなグラフに対しては $\log k$ 倍の高

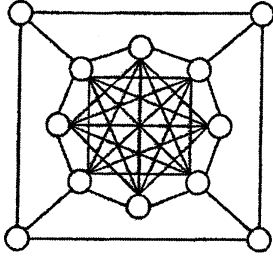


図 1: シミュレーションを行ったグラフ

速化しかできないことも示している。

本研究では多種ランダムウォークという、各トークンが個別の遷移確率行列を持つ新たなランダムウォークのモデルを提案し、その全訪問時間について考える。多種ランダムウォークとは、多重ランダムウォークを一般化したものであり、各トークンに異なる遷移確率行列を採用することで、同一の遷移確率行列の場合より高速にネットワークを探索させることを目指している。

多種ランダムウォークが多重ランダムウォークよりも高速に探索し得ることを端的に示すため計算機によるシミュレーションの結果を紹介する。このシミュレーションでは、異なる遷移確率行列で遷移する 2 トークンが全頂点を訪問するまでのステップ数 (多種ランダムウォークの全訪問時間に相当) を計測している。各トークンはそれぞれ $\beta = 0.5$ と $\beta = 1.0$ の β ランダムウォーク (式 (1)) の遷移確率行列に従う。また比較対象として 2 つのトークンが共に $\beta = 0.5$ 、共に $\beta = 1$ 、共に $\beta = 0$ の遷移確率に従う 2 トークンが全頂点を訪問するまでのステップ数 (いずれも $\beta = 0.5$, $\beta = 1$, $\beta = 0$ の多重ランダムウォークの全訪問時間に相当) を計測した。シミュレーションを行うグラフは頂点数 c のクリークと、頂点数 $n - c$ ($c \geq n/2$) のサイクルからなり、クリークの各頂点からサイクルの頂点への辺をちょうど 1 本もつ。図 1 は $n = 12$, $c = 8$ としたときのシミュレーションのグラフの例である。表 1 は多重/多種ランダムウォークの全訪問時間のシミュレーション結果であり、頂点数 1,000 のグラフにおいて 10,000 回試行を行った平均値である。Ikeda ら [7] の研究によれば、 $\beta = 0.5$ の β ランダムウォークは $\beta = 1.0$ の場合よりも高速なランダムウォークであるが、シミュレーションでは $c = 666, 750$, そして 800 のグラフにおいて、多種ランダムウォークのほうが高速であった。

本研究では多種ランダムウォークの全訪問時間につ

遷移確率行列	$c = 500$	$c = 750$	$c = 800$	$c = 900$
$(\beta = 0.5) \ \& \ (\beta = 1)$	16,455	11,365	9,299	6,924
$(\beta = 0.5) \times 2$	14,309	14,647	14,254	12,701
$(\beta = 1) \times 2$	145,265	147,190	126,715	75,497
$(\beta = 0) \times 2$	377,687	756,096	821,740	907,882

表 1: 2 つの β ランダムウォークに従うトークンからなる多重/多種ランダムウォークの全訪問時間

いて議論する。単一ランダムウォークにおいては全訪問時間の上下界を到達時間によって与える不等式が Matthews [9] により知られており、2.1 節で詳しく述べる。本稿では多種ランダムウォークにおける全訪問時間の上下界を到達時間で抑える不等式を示す。またその上下界が完全グラフ、完全二部グラフ、ランダムグラフなどのグラフにおいてタイトであることを示す。

2 準備

2.1 単一ランダムウォーク

本稿では 1 つのトークンによるランダムウォークを単一ランダムウォークと呼ぶ。 n 頂点のグラフ $G = (V, E)$ が与えられ、 G 上の単一ランダムウォークの遷移確率行列を $P = (p_{uv}) : u, v \in V$ とする。ここで $\{u, v\} \in E$ のときに限り $p_{uv} > 0$ とする。

Definition 2.1. 遷移確率行列 P に従うトークンが頂点 u を出発して v に到達するまでのステップ数の期待値を、 u から v への到達時間と呼び $H_G^P(u, v)$ と記述する。また G に対する到達時間 H_G^P を次のように定義する。

$$H_G^P = \max_{u, v \in V} H_G^P(u, v). \quad (2)$$

Definition 2.2. 遷移確率行列 P に従うトークンが頂点 u を出発して他の全ての頂点に到達するまでのステップ数の期待値を、 u からの全訪問時間と呼び $C_G^P(u)$ と記述する。また G に対する全訪問時間 C_G^P を次のように定義する。

$$C_G^P = \max_{u \in V} C_G^P(u). \quad (3)$$

Matthews [9] により、到達時間と全訪問時間について以下のような関係式が知られている。

$$h_{n-1} \min_{u \neq v \in V} H_G^P(u, v) \leq C_G^P \leq h_{n-1} \max_{u \neq v \in V} H_G^P(u, v) \quad (4)$$

ここで h_n は調和級数で, $h_n = \sum_{i=1}^n \frac{1}{i}$ である.

2.2 多種ランダムウォーク

k トークンの多種ランダムウォークでは, それぞれのトークンが個々の遷移確率に従って独立に遷移する. n 頂点のグラフ $G = (V, E)$ が与えられ, G 上の多種ランダムウォークの各トークンの遷移確率行列を $P_i = (p_i(u, v)) : u, v \in V, 1 \leq i \leq k$ とする. また多種ランダムウォークの現在の状態を $S = (s_1, s_2, \dots, s_k) \in V^k$ とし, 各トークンの滞在している頂点は $s_i \in V$ で表わされたとする. このとき各トークンはそれぞれの遷移確率行列 $p_i(s_i, s'_i)$ に従って遷移し, 多種ランダムウォークの次の状態を $S' = (s'_1, s'_2, \dots, s'_k) \in V^k$ とする. つまりトークンは同期して遷移する.

多種ランダムウォークの到達時間を以下のように定義する.

Definition 2.3. 遷移確率行列集合 $P^k := (P_1, P_2, \dots, P_k)$ に従う多種ランダムウォークが状態 $S = (s_1, s_2, \dots, s_k)$ から遷移を開始して, 頂点 v に少なくとも 1 つのトークンが到達するまでの同期ステップ数の期待値を, 状態 S から v の到達時間と呼び $H_G^{P^k}(S, v)$ と記述する. また G に対する到達時間 $H_G^{P^k}$ を次のように定義する.

$$H_G^{P^k} = \max_{S \in V^k, v \in V} H_G^{P^k}(S, v). \quad (5)$$

Definition 2.4. 遷移確率行列集合 $P^k := (P_1, P_2, \dots, P_k)$ に従う多種ランダムウォークが状態 $S = (s_1, s_2, \dots, s_k)$ から遷移を開始して, 他のすべての頂点を少なくとも 1 つのトークンが訪問するまでの同期ステップ数の期待値を, 状態 S からの全訪問時間と呼び $C_G^{P^k}(S)$ と記述する. また G に対する全訪問時間 $C_G^{P^k}$ を次のように定義する.

$$C_G^{P^k} = \max_{S \in V^k} C_G^{P^k}(S). \quad (6)$$

$G = (V, E)$ を n 頂点の有限グラフとし, $N(u)$ を $u \in V$ の隣接頂点集合とする. u の次数は $\deg(u) = |N(u)|$ で表わす. 多種ランダムウォークはマルコフ過程として記

述できる. マルコフ過程の各状態は k 個の頂点の順序集合で定義し, $\Omega = (V^k)^{\mathbb{N} \cup \{0\}}$ をマルコフ過程の状態の有限順序列の全集合とする. また Ω 上のマルコフ測度空間を $M(\Omega)$ とする. $S = (s_1, s_2, \dots, s_k)$, $S' = (s'_1, s'_2, \dots, s'_k)$ をこのマルコフ過程の状態対とすると, $S \rightarrow S'$ の状態遷移確率 $p(S, S')$ は次のように定義される.

$$p(S, S') = \prod_{i=1}^k p_i(s_i, s'_i).$$

ここで $p_i(u, v)$ は多種ランダムウォークにおける i 番目のトークンの $u \rightarrow v$ への遷移確率である.

$\mu \in M(\Omega)$ を初期状態 $S_0 \in V^k$ を持つマルコフ測度とする, つまり $\omega = (\omega_0, \omega_1, \dots) \in \Omega$ としたとき

$$\mu(X_0(\omega) = S_0) = 1,$$

が成立する. ここで $X_t(\omega)$ は ω の t 番目の要素を表す. また任意の $S \in V^k$ について以下が成立する.

$$\sum_{S' \in (V^k)} p(S, S') = 1.$$

また任意の $S, S', U_0, U_1, \dots, U_i \in V^k$ と $t \in \mathbb{N} \cup \{0\}$ について以下が成立する.

$$\begin{aligned} \mu(X_{i+1}(\omega) = S' \mid X_0(\omega) = U_0, \\ X_1(\omega) = U_1, \dots, X_i(\omega) = U_i = S) \\ = \mu(X_{i+1}(\omega) = S' \mid X_i(\omega) = U_i = S) = p(S, S'), \end{aligned}$$

さらに任意の $u, v, w_0, w_1, \dots, w_i \in V$ と $t \in \mathbb{N} \cup \{0\}$ について以下が成立する.

$$\begin{aligned} \mu(Y_j(X_{i+1}(\omega)) = v \mid Y_j(X_0(\omega)) = w_0, \\ Y_j(X_1(\omega)) = w_1, \dots, Y_j(X_i(\omega)) = w_i = u) \\ = \mu(Y_j(X_{i+1}(\omega)) = v \mid Y_j(X_i(\omega)) = u) = p_j(u, v), \end{aligned}$$

ここで $Y_j(S)$ は $S \in V^k$ の j 番目の要素を表す. そして任意の $1 \leq j \leq k$ に対して上記の条件を満たすマルコフ測度は以下の M^+ により示される.

$$\begin{aligned} M^+(\Omega) = \left\{ \mu \in M^+(\Omega) \mid p(S, T) > 0 \right. \\ \left. \text{only if } \forall j, Y_j(T) \in N(Y_j(S)) \right\}. \end{aligned}$$

3 多種ランダムウォークの全訪問時間の上下界

Theorem 3.1. $G = (V, E)$ を連結グラフとし, P_i ($i = 1, \dots, k$) を G 上の任意の遷移確率行列として, $P^k := (P_1, P_2, \dots, P_k)$ とすると以下が成り立つ.

$$h_{n-1} \min_{S \in V^k, v \in V} (H_G^{P^k}(S, v) - 1) \leq C_G^{P^k} \quad (7)$$

$$\leq h_{n-1} \max_{S \in V^k, v \in V} (H_G^{P^k}(S, v)) \quad (8)$$

ここで $h_n = \sum_{i=1}^n i^{-1}$ である.

Proof. V の順列の全集合を S_V とし, S_V 上の一様測度を ν とする. V の順列 $\pi = (v_1, v_2, \dots, v_n) \in S_V$ の i 番目の要素を $\sigma_i(\pi) = v_i$ とする. 頂点 $u \in V$ と u だけからなる順序集合 $U_0 = (u, u, \dots, u) \in V^k$ について, ν の $\{\pi : \sigma_1(\pi) = u\}$ を満たす条件付き測度を ν_u とする. 同様に μ の $\{\omega : X_0(\omega) = (u, u, \dots, u) = U_0\}$ を満たす条件付き測度を μ_u とする. また μ_u と ν_u の直積測度を P_u とする. 一方で $\Omega = V^{\mathbb{N} \cup \{0\}}$ を頂点の有限順序列の全集合とし, 写像 f を $\omega = (\omega_0, \omega_1, \dots) \in \Omega$ について

$$f(\omega) = \dot{\omega} = Y_1(X_0), Y_2(X_0), \dots, Y_k(X_0), Y_1(X_1), \\ Y_2(X_1), \dots, Y_k(X_1), Y_1(X_2), \dots,$$

つまり

$$X_{ki+j}(\dot{\omega}) = Y_j(X_i(\omega))$$

を満たす $\Omega \rightarrow \dot{\Omega}$ の全単射とする. さらに $\dot{\mu}_u$ を $\{\dot{\omega} : X_i = u \text{ if } 0 \leq i \leq k-1\}$ を満たす $\dot{\omega}$ の条件付き測度とし, $\dot{\mu}_u$ と ν_u の直積測度を $\dot{P}_u$ とする. そして $\tau(\omega, v)$, $\dot{\tau}(\dot{\omega}, v)$, $T_j(\omega, \pi)$, $\dot{T}_j(\dot{\omega}, \pi)$ をそれぞれ次のように定義する.

$$\tau(\omega, v) = \inf\{t \geq 0 : v \in X_t(\omega)\},$$

$$\dot{\tau}(\dot{\omega}, v) = \inf\{t \geq 0 : v = X_t(\dot{\omega})\},$$

$$T_j(\omega, \pi) = \max_{i \leq j} \tau(\omega, \sigma_i(\pi)),$$

$$\dot{T}_j(\dot{\omega}, \pi) = \max_{i \leq j} \dot{\tau}(\dot{\omega}, \sigma_i(\pi)).$$

ここで $X_i(\dot{\omega}) \in V$ は $\dot{\omega}$ の i 番目の要素である. このとき

$$T_{j-1}(\omega, \pi) < T_j(\omega, \pi) \Leftrightarrow T_{j-1}(\omega, \pi) < \tau(\omega, v_j),$$

である. さらに

$$T_{j-1}(\omega, \pi) < \tau(\omega, v_j) \Rightarrow \dot{T}_{j-1}(\dot{\omega}, \pi) < \dot{\tau}(\dot{\omega}, v_j)$$

が成立する. ただし

$$T_{j-1}(\omega, \pi) < \tau(\omega, v_j) \neq \dot{T}_{j-1}(\dot{\omega}, \pi) < \dot{\tau}(\dot{\omega}, v_j),$$

であることに注意しなければならない. なぜなら多種ランダムウォークでは 1 ステップで複数の未到達の頂点に訪問し得るため, 以下の事象が成立する可能性があるためである.

$$(T_{j-1}(\omega, \pi) = \tau(\omega, v_j)) \wedge (\dot{T}_{j-1}(\dot{\omega}, \pi) < \dot{\tau}(\dot{\omega}, v_j)).$$

ゆえに,

$$\begin{aligned} & P_u(T_{j-1}(\omega, \pi) < T_j(\omega, \pi)) \\ & \leq \dot{P}_u(\dot{T}_{j-1}(\dot{\omega}, \pi) < \dot{\tau}(\dot{\omega}, v_j)) \\ & = \dot{P}_u(\dot{\tau}(\dot{\omega}, \sigma_i(\pi)) < \dot{\tau}(\dot{\omega}, \sigma_j(\pi)), 2 \leq i < j) \\ & = \int_{\dot{\omega}} \nu_u(\{\pi : \dot{\tau}(\dot{\omega}, \sigma_i(\pi)) < \dot{\tau}(\dot{\omega}, \sigma_j(\pi)), \\ & \quad 2 \leq i < j\}) d\dot{\mu}_u(\dot{\omega}) \\ & = \int_{\dot{\omega}} \frac{(n-1)!}{(j-1)!(n-j)!} \times \frac{(j-2)!(n-j)!}{(n-1)!} d\dot{\mu}_u(\dot{\omega}) \\ & = \frac{1}{j-1} \end{aligned}$$

任意の $\pi \in S_V$ で, すべての頂点を訪問するまでのステップ数は $T_n(\omega, \pi)$ なので,

$$\begin{aligned} C_G^{P^k} &= E_{P_u} \left[T_n(\omega, \pi) \right] \\ &= \sum_{j=2}^n E_{P_u} \left[T_j(\omega, \pi) - T_{j-1}(\omega, \pi) \right] \\ &= \sum_{j=2}^n E_{P_u} \left[T_j(\omega, \pi) - T_{j-1}(\omega, \pi) : \right. \\ & \quad \left. T_{j-1}(\omega, \pi) \neq T_j(\omega, \pi) \right] \\ &= \sum_{j=2}^n H_G^{P^k}(X_{T_{j-1}(\omega, \pi)}(\omega), v_j) \cdot \\ & \quad P_u(T_{j-1}(\omega, \pi) \neq T_j(\omega, \pi)) \\ &\leq \max\{H_G^{P^k}(S, v) : v \in V, S \in V^k\} \cdot \\ & \quad \sum_{j=2}^n P_u(T_{j-1}(\omega, \pi) < T_j(\omega, \pi)) \\ &\leq \max\{H_G^{P^k}(S, v) : v \in V, S \in V^k\} \sum_{j=2}^n \frac{1}{j-1} \\ &\leq h_{n-1} \max\{H_G^{P^k}(S, v) : v \in V, S \in V^k\}. \end{aligned}$$

よって不等式 (7) の右辺が示された.

不等式 (7) の左辺を示すため以下を利用する.

$$\tau(\omega, v) = \left\lceil \frac{\tau(\omega, v)}{k} \right\rceil,$$

$$T_j(\omega, \pi) = \left\lceil \frac{T_j(\omega, \pi)}{k} \right\rceil.$$

また定義より以下も成立する.

$$\begin{aligned} T_{j-1}(\omega, \pi) \neq T_j(\omega, \pi) &\Leftrightarrow T_{j-1}(\omega, \pi) < T_j(\omega, \pi) \\ &\Leftrightarrow T_{j-1}(\omega, \pi) < \tau(\omega, v_j). \end{aligned}$$

よって,

$$\begin{aligned} &E_{P_u}[T_j(\omega, \pi)] \\ &= E_{P_u}\left[\sum_{j=2}^n (T_j(\omega, \pi) - T_{j-1}(\omega, \pi))\right] \\ &= \sum_{j=2}^n E_{P_u}[T_j(\omega, \pi) - T_{j-1}(\omega, \pi)] \\ &\geq \sum_{j=2}^n \left(E_{P_u}[T_j(\omega, \pi) - T_{j-1}(\omega, \pi)] \cdot \right. \\ &\quad \left.P_u(T_{j-1}(\omega, \pi) \neq T_j(\omega, \pi))\right) \\ &= \sum_{j=2}^n \left(E_{P_u}[\tau(\omega, v_j) - \tau(\omega, X_{T_{j-1}(\omega, \pi)})] \cdot \right. \\ &\quad \left.P_u(T_{j-1}(\omega, \pi) < T_j(\omega, \pi))\right) \\ &\geq \min_{s, v \in V} \left\{ E_{P_u}[\tau(\omega, v) - \tau(\omega, s) : \tau(\omega, s) < \tau(\omega, v)] \right\} \cdot 4 \\ &\quad \sum_{j=2}^n P_u(T_{j-1}(\omega, \pi) < \tau(\omega, v_j)) \\ &= h_{n-1} \min_{s, v \in V} \left\{ E_{P_u}[\tau(\omega, v) - \tau(\omega, s) : \tau(\omega, s) < \tau(\omega, v)] \right\}. \end{aligned}$$

また

$$E_{P_u}[\tau(\omega, v) - \tau(\omega, s) : \tau(\omega, s) < \tau(\omega, v)] \geq \min_{S \in V^k} H_G^{P^k}(S, S),$$

と合わせて,

$$\begin{aligned} C_G^{P^k} &= E_{P_u}[T_n(\omega, \pi)] \\ &= E_{P_u}\left[\left\lceil \frac{T_n(\omega, \pi)}{k} \right\rceil\right] \\ &\geq \frac{E_{P_u}[T_n(\omega, \pi)]}{k} \end{aligned}$$

$$\begin{aligned} &\geq \frac{h_{n-1}}{k} \min_{s, v \in V} \left\{ E_{P_u}[\tau(\omega, v) - \tau(\omega, s) : \right. \\ &\quad \left. \tau(\omega, s) < \tau(\omega, v)] \right\} \\ &\geq h_{n-1} \min_{s, v \in V} \left\{ E_{P_u}\left[\frac{\tau(\omega, v)}{k} - \frac{\tau(\omega, s)}{k} : \right. \right. \\ &\quad \left. \left. \tau(\omega, s) < \tau(\omega, v)\right] \right\} \\ &\geq h_{n-1} \min_{s, v \in V} \left\{ E_{P_u}\left[\frac{\tau(\omega, v)}{k} - \left\lceil \frac{\tau(\omega, s)}{k} \right\rceil : \right. \right. \\ &\quad \left. \left. \tau(\omega, s) < \tau(\omega, v)\right] \right\} \\ &\geq h_{n-1} \min_{s, v \in V} \left\{ E_{P_u}\left[\left\lceil \frac{\tau(\omega, v)}{k} \right\rceil - \left\lceil \frac{\tau(\omega, s)}{k} \right\rceil - 1 : \right. \right. \\ &\quad \left. \left. \tau(\omega, s) < \tau(\omega, v)\right] \right\} \\ &= h_{n-1} \min_{s, v \in V} \{E_{P_u}[\tau(\omega, v) - \tau(\omega, s)] - 1\} \\ &\geq h_{n-1} \min_{v \in V, S \in V^k} \{H_G^{P^k}(S, v) - 1\}, \end{aligned}$$

不等式 (7) の左辺も示された.

□

最後にもう一つ別の下界を示す. これは全訪問時間の定義より自明に成立する.

Proposition 3.1. $G = (V, E)$ を連結グラフとし, P_i ($i = 1, \dots, k$) を G 上の任意の遷移確率行列として, $P^k := (P_1, P_2, \dots, P_k)$ とすと以下が成り立つ.

$$\max_{S \in V^k, v \in V} H_G^{P^k}(S, v) \leq C_G^{P^k}. \quad (9)$$

上下界のタイトな例

この節では定理 3.1 の多重標準ランダムウォークによるタイトな例を示す.

4.1 完全グラフ K_n

任意の開始状態 $S \in V^k$ と任意の目的頂点 $v \in V$ ($v \notin S$) について, グラフの対称性より到達時間を次のように計算できる.

$$H_{K_n}^{P^k}(S, v) = \frac{1}{1 - \left(1 - \frac{1}{n-1}\right)^k} \quad (10)$$

トークン数 k について $k \leq n$ とすると, 以下の不等式が得られる.

$$1 - \frac{k}{n-1} \leq \left(1 - \frac{1}{n-1}\right)^k \leq 1 - \frac{k}{2(n-1)}$$

これを変形して,

$$\frac{k}{2(n-1)} \leq 1 - \left(1 - \frac{1}{n-1}\right)^k \leq \frac{k}{n-1}.$$

よって

$$\frac{n-1}{k} \leq \frac{1}{1 - \left(1 - \frac{1}{n-1}\right)^k} \leq \frac{2(n-1)}{k}. \quad (11)$$

式 (10) および (11) より

$$\min_{S \in V^k, v \in V} H_{K_n}^{P^k}(S, v) = \max_{S \in V^k, v \in V} H_{K_n}^{P^k}(S, v) = \Theta\left(\frac{n}{k}\right).$$

さらに定理 3.1 より以下を得る.

$$C_{K_n}^{P^k} = O\left(\frac{n}{k} \log n\right),$$

またこの上下界はタイトである.

4.2 完全二部グラフ $K_{m,n}$

U, W を完全二部グラフ $K_{m,n}$ の極大独立集合とし, $|U| = m, |W| = n$ とする. また一般性を失わず, $m \leq n$ と仮定できる. 全てのトークンが同一の頂点から遷移を始めるので, $K_{m,n}$ で到達時間 $H_{K_{m,n}}^{P^k}(S, v)$ が最小になるのは開始状態 S が W に含まれる頂点からなり, 目的頂点が $v \in U$ のときである. 同様に最大になるのは, 開始状態 S が W に含まれる頂点からなり, 目的頂点が $v \in W$ のときである. よって

$$\begin{aligned} \min_{S \in V^k, v \in V} H_{K_{m,n}}^{P^k}(S, v) &= \frac{2}{1 - \left(1 - \frac{1}{m}\right)^k} - 1, \\ \max_{S \in V^k, v \in V} H_{K_{m,n}}^{P^k}(S, v) &= \frac{2}{1 - \left(1 - \frac{1}{n}\right)^k}. \end{aligned}$$

完全グラフ上の多重標準ランダムウォークと同様の解析により

$$\begin{aligned} \min_{S \in V^k, v \in V} H_{K_{m,n}}^{P^k}(S, v) &\geq \frac{2m-1}{k} \quad (k < m) \\ \max_{S \in V^k, v \in V} H_{K_{m,n}}^{P^k}(S, v) &\leq \frac{4n}{k} \quad (k < n). \end{aligned}$$

よって以下の不等式が得られる.

$$\frac{2m-1}{k} \log(m+n) \leq C_{K_{m,n}}^{P^k} \leq \frac{4n}{k} \log(m+n).$$

この場合, 定理 3.1 の上下界は $\frac{n}{m} = \Theta(1)$ のときタイトになる.

4.3 ランダムグラフ $G_{n,p}$

単一ランダムウォークにおいて頂点 u から開始したランダムウォークがステップ s で v を一度も訪問していない事象を $A_s(u, v)$ とすると, 任意の $t > 0$ について以下が成り立つ.

$$H(u, v) = \sum_{s=1}^{\infty} \Pr(A_s(u, v)) \leq t + \sum_{s=t+1}^{\infty} \Pr(A_s(u, v))$$

多重ランダムウォークにおいても各トークンは独立に遷移することから以下が成り立つ.

$$\begin{aligned} H(S, v) &\leq t + \sum_{s=t+1}^{\infty} \Pr(\cap_{u \in S} A_s(u, v)) \\ &= t + \sum_{s=t+1}^{\infty} \prod_{u \in S} \Pr(A_s(u, v)). \end{aligned}$$

また π_v を v の定常確率, t_m を単一ランダムウォークの混合時間として以下の式を得る [4].

$$\Pr(A_s(u, v) | A_{s-1}(u, v)) = 1 - \pi_v,$$

さらに標準ランダムウォークであれば, 次数を用いて定常確率が計算出来て, 以下の式を得る.

$$\Pr(A_s(u, v) | A_{s-1}(u, v)) = 1 - \frac{\deg(v)}{2|E|}.$$

また任意の $t < s$ について, 以下が成り立つ.

$$\Pr(A_s(u, v)) = \Pr(A_t(u, v)) \prod_{t'=t+1}^s \Pr(A_{t'}(u, v) | A_{t'-1}(u, v)).$$

また多重ランダムウォークのトークン数を $k = |S|$ とし, 不等式 (11) を用いて,

$$\begin{aligned} H(S, v) &\leq t_m + \sum_{s>t} \prod_{u \in S} \left(1 - \frac{\deg(v)}{2|E|}\right)^{s-t} \\ &= t_m + \sum_{s>t} \left(1 - \frac{\deg(v)}{2|E|}\right)^{k(s-t)} \\ &= t_m + \frac{1}{1 - \left(1 - \frac{\deg(v)}{2|E|}\right)^k} \\ &\leq t_m + \frac{4|E|}{k \cdot \deg(v)}. \end{aligned}$$

ここで $G(n, p)$ において p が十分大きく, 高い確率で $\min_{v \in V} \deg(v) = \Theta(n)$ を満たすとする, $G_{n,p}$ の混合時間が $O(\log n)$ である [11] ことと, 辺数について

$2|E| = \sum \deg(v) = \Theta(n^2)$ となることを合わせて以下のような到達時間の上界を得る.

$$\max_{S \in V^k, v \in V} H(S, v) \leq \log n + O\left(\frac{n}{k}\right),$$

定理 3.1 を適応して以下を得る.

$$C_G^{P^k} = O\left(\frac{n}{k} \log n\right)$$

一方 Alon ら [3] により密なランダムグラフ $G(n, p)$ 上の k トークン多重標準ランダムウォークの全訪問時間は $\Theta(\frac{n}{k} \log n)$ であることが知られている. このことから定理 3.1 の上界は密なランダムグラフ上の多重ランダムウォークにおいてタイトであることが分かる.

5 まとめ

本稿では多重ランダムウォークを一般化した多種ランダムウォークを提案し, 多種ランダムウォークの全訪問時間の到達時間による上下界の式をいくつかのタイトな例とともに示した. しかしながら一般のグラフにおいての高速な多種ランダムウォークを構成する遷移確率行列の組み合わせ方と, 各遷移確率行列の設計方法, さらに出来上がった多種ランダムウォークが単一ランダムウォークと比べてどのくらい速いのかは到達時間, 全訪問時間のどちらにおいても未解決で, 挑戦的な課題である.

参考文献

- [1] D. J. Aldous: On the time taken by random walks on finite groups to visit every state, *Z. Wahrsh. verw. Gebiet*, 62(1983), 361–393.
- [2] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovasz, and C. Rackoff: Random walks, universal traversal sequences, and the complexity of maze problems, *Proc. 20th Ann. Symposium on Foundations of Computer Science 1979*, 218–223.
- [3] N. Alon, C. Avin, M. Koucky, G. Kozma, Z. Lotker, and M. R. Tuttle: Many random walks are faster than one, *Combinatorics, Probability and Computing* 20 (2011), 481–502.
- [4] C. Cooper and A. Frieze: The cover time of sparse random graphs, *Random structures and Algorithms*, 30 (2007) 1–16.
- [5] C. Cooper and A. Frieze: Random walks on random graphs. *NANO-NET, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 3 (2009), 95–106.
- [6] K. Efremenko and O. Reingold: How well do random walk parallelize, *Lecture Notes in Computer Science*, 5687 (2009) 476–489.
- [7] S. Ikeda, I. Kubo, and M. Yamashita: The hitting and cover times of random walks on finite graphs using local degree information, *Theoretical Computer Science*, 410 (2009), 94–100.
- [8] J. Jonasson: On the cover time of random walks on random graphs, *Combinatorics, Probability and Computing*, 7 (1998), 265–279.
- [9] P. Matthews: Covering problems for Markov chain, *The annals of probability*, 16 (1988), 1215–1228.
- [10] Y. Nonaka, H. Ono, S. Kijima, and M. Yamashita: How slow, or fast, are standard random walks? — analysis of hitting and cover times on tree, *CR-PIT*, 119 (2011), 63–68.
- [11] A. Sinclair: Improved bounds for mixing rates of Markov chains and multicommodity flow, *Lecture Notes in Computer Science*, 583 (1992), 474–487.