Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 202 (2007) 144 – 154
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Coalescent random walks on graphs

Jianjun Paul Tian[a,*], Zhenqiu Liu[b]

[a] *Mathematical Biosciences Institute, The Ohio State University, 18th Avenue, Columbus OH 43210, USA*
[b] *Division of Biostatistics, Greenebaum Cancer Center and Department of Epidemiology and Preventive Medicine, University of Maryland Medicine, USA*

Received 2 June 2005; received in revised form 5 October 2005

Abstract

Inspired by coalescent theory in biology, we introduce a stochastic model called "multi-person simple random walks" or "coalescent random walks" on a graph G. There are any finite number of persons distributed randomly at the vertices of G. In each step of this discrete time Markov chain, we randomly pick up a person and move it to a random adjacent vertex. To study this model, we introduce the tensor powers of graphs and the tensor products of Markov processes. Then the coalescent random walk on graph G becomes the simple random walk on a tensor power of G. We give estimates of expected number of steps for these persons to meet all together at a specific vertex. For regular graphs, our estimates are exact.

1. Introduction

Inspired by coalescent theory in population genetics, we consider, in the present paper, a class of models, called coalescent random walks on graphs which is actually an generalization of coalescent theory. Let us recall the basic idea about coalescent theory firstly. Taking a sample with n individuals from a population, we label them as $1, 2, \ldots, n$, and ask a question how long ago the recent common ancestor of the sample lived. Coalescent theory answers this question by running a continuous time Markov chain over the collection of partitions $A_1, A_2, \ldots, A_t$ of $1, 2, \ldots, n$, where each A_i consists of one subset of individuals that have coalesced and hence are identical by descent. To explain this theory, we look at an example that the sample consists of five individuals $1, 2, \ldots, 5$. For the purpose of illustration, we randomly choose partitions at each time when a coalescent event happens. As we work backwards in time, partitions were chosen as the following:

time	0	$\{1\}$	$\{2\}$	$\{3\}$	$\{4\}$	$\{5\}$
	T_4	$\{1, 2\}$	$\{3\}$	$\{4\}$	$\{5\}$	
	T_3	$\{1, 2\}$	$\{3\}$	$\{4, 5\}$		
	T_2	$\{1, 2, 3\}$	$\{4, 5\}$			
	T_1	$\{1, 2, 3, 4, 5\}$.				

Initially, the partition consists of five singletons since there has been no coalescence. After 1 and 2 coalesce at time T_4, they appear in the same set. Then 4 and 5 coalesce at time T_3, etc. Finally, at time T_1 we end up with all the labels

* Corresponding author.
E-mail address: tianjj@mbi.osu.edu (J.P. Tian).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2005.10.039

in one set. After figured out the probabilities of coalescent events, Kingman [5,6] got a continuous time Markov chain asymptotically. That is coalescent theory. If we construct a graph by taking the set of vertices as the set of all partitions of labeled individuals of a sample and the set of edges as the set of coalescent relations, we could see that coalescent theory is a class of continuous time Markov chains over a special class of graphs. The graphs are special, because they have a kind of partial directions from n distinct vertices to the other one vertex which represent the genealogical relations. From a different viewpoint, coalescent theory models how n particles come together under certain conditions. Therefore, we generally consider the following model that we call k-coalescent random walks or multi-person simple random walks.

Given a graph G with n vertices and m edges, and suppose that k persons distribute on n vertices of G. We here allow that several persons can stand together on one vertex and k can be bigger than n or smaller than n. At each time step, one person could randomly move to one of his neighbor vertices with the equal probability of moving any one of his neighbor vertices except the vertex he currently stands on. Then, there arises an interesting question that when these k persons will first time meet together on a specific vertex.

To solve this problem, in the rest of the article, we generalize the concept of the tensor powers of a graph, which are introduced in paper [9]. We recall the tensor products of Markov processes. By using the continuous time Markov chains over the kth tensor powers of the given graph, we turn the k-coalescent random walks on the ground graph G into the simple random walks on its kth tensor powers. This way, we get an estimation of the expectation of the time steps that k persons starting with any distribution on the graph come together at a specific vertex.

For simplicity, we only consider connected simple graphs. These are connected graphs without multiple edges and loops. We adopt the following notations and terminologies for a graph G. The sets of vertices and edges of G are $V(G)$ and $E(G)$, respectively. The *order* n of G is the number of vertices of G, and the *size* m of G is the number of edges of G. Thus, $n=|V(G)|$ and $m=|E(G)|$. For a vertex $x \in V(G)$, $\Gamma(x)$ is the set of vertices which are connected to x by an edge in $E(G)$. The *degree* of a vertex x is $d(x)=|\Gamma(x)|$. We have

$$\sum_{x\in V(G)} d(x)=2m.$$

The adjacent matrix of G is denote by $A(G)$ and the diagonal matrix $D(G)$ has the sequence of degrees at each vertex as its diagonal entries. Finally, we denote

$$d_m=\min\{d(x); x \in V(G)\} \quad \text{and} \quad d_M=\max\{d(x); x \in V(G)\}.$$

We would like to refer the reader to [7,1,9,3] for basic notions and results in the study of simple random walks on graphs.

2. Coalescent random walks and simple random walks on graphs

2.1. The tensor powers of graphs

For a graph G with order n and size m, a nth tensor power of G was introduced in paper [9]. It is not necessary to constrain the order of tensor power to be the order of the graph. Although the motivation of our generalization of tensor powers of a graph is coalescent random walks of any number of persons on the graph, the general tensor powers have their own interesting applications.

Let I_k and I_n be finite sets of cardinalities k and n, respectively. For example, we may have $I_k=\{1,2,\dots,k\}$, and $I_n=\{1,2,\dots,n\}$. We denote the set of all maps from I_k to I_n by $M_{k,n}$. When $k=n$, we have the symmetric group S_n sitting inside of $M_{n,n}$. A map $x \in M_{k,n}$ is called a *generalized permutation of deficiency* j if

$$|x(I_k)|=k-j.$$

We denote $\mathrm{def}(x)=j$.

$M_{k,n}$ is a semigroup under composition when $k\geqslant n$. The symmetrical group S_n is a subgroup of $M_{k,n}$ only when $k=n$. The deficiency determines a grading on $M_{k,n}$ which is compatible with the semigroup product on $M_{k,n}$:

$$\mathrm{def}(x)+\mathrm{def}(y)\geqslant 2\mathrm{def}(x\circ y).$$