Discrete Mathematics 311 (2011) 2496–2511

Contents lists available at SciVerse ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Random induced subgraphs of Cayley graphs induced by transpositions

Emma Yu Jin [a], Christian M. Reidys [b,*]

[a] Department of Computer Science, University of Kaiserslautern, 67663 Kaiserslautern, Germany

[b] Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

ARTICLE INFO

Article history:
Received 22 September 2009
Received in revised form 19 July 2011
Accepted 20 July 2011
Available online 17 August 2011

Keywords:
Random graph
Permutation
Transposition
Giant component
Vertex boundary

ABSTRACT

In this paper we study random induced subgraphs of Cayley graphs of the symmetric group induced by an arbitrary minimal generating set of transpositions. A random induced subgraph of this Cayley graph is obtained by selecting permutations with independent probability, λ_n. Our main result is that for any minimal generating set of transpositions, for probabilities $\lambda_n = \frac{1+\epsilon_n}{n-1}$ where $n^{-\frac{1}{3}+\delta} \leq \epsilon_n < 1$ and $\delta > 0$, a random induced subgraph has a.s. a unique largest component of size $(1+o(1)) \cdot x(\epsilon_n) \cdot \frac{1+\epsilon_n}{n-1} \cdot n!$. Here $x(\epsilon_n)$ is the survival probability of a Poisson branching process with parameter $\lambda = 1 + \epsilon_n$.

1. Introduction

One central problem arising in parallel computing is to determine an optimal linkage of a given collection of processors. A particular class of processor linkages with point-to-point communication links are static interconnection networks. The latter are widely used for message-passing architectures. A static interconnection network can be represented as a graph. The binary n-cubes, Q_2^n, [1,35] are a particularly well-studied class of interconnection networks [15,20,21,40].

Akers et al. [2] observed the deficiencies of n-cubes as models for interconnection networks and proposed an alternative: the Cayley graph of the permutation group induced by the $(n-1)$ star-transpositions $(1\,i)$, which was denoted by $\Gamma(S_n, P_n)$. Pak [36] studied minimal decompositions of a particular permutation via star-transpositions and Irving and Ratton [29] extended his results. The star-graph $\Gamma(S_n, P_n)$ is in many aspects superior to n-cubes [1,35]. Some properties of star-graphs studied in [26–28,25,30,34] were cycle-embeddings and path-embeddings. The diameter and the fault diameter of star-graphs were computed by Akers et al. [2], Latifi [32], Rouskov et al. [39] and Lin et al. [33] analyzed diagnosability. An alternative to n-cubes as interconnection networks are the bubble-sort graphs [3], studied by Tchuente [41]. The bubble-sort graph is the Cayley graph of the permutation group induced by all $n-1$ canonical transpositions $(i\,i+1)$, denoted by $\Gamma(S_n, B_n)$.

Recently, Araki [5] brought the attention to a generalization of star- and bubble-sort graphs, the Cayley graph generated by all transpositions [13]. The latter has direct connections to a problem of interest in computational biology: the evolutionary distances between species based on their genome order in the Cayley graph of signed permutations generated by reversals. A reversal is a special permutation that acts by flipping the order as well as the signs of a segment of genes. Hannenhalli and Pevzner [22] presented an algorithm computing minimal number of reversals needed to transform one sequence of distinct genes into a given signed permutation. For distant genomes, however, it is well-known that the true evolutionary distance is generally much greater than the shortest distance [43,12,11,7]. In order to obtain a more realistic estimate of the true evolutionary distance, the expected reversal distance was shifted into focus. Its computation, however,

* Corresponding author.
E-mail addresses: jin@cs.uni-kl.de (E.Y. Jin), duck@imada.sdu.dk, duck@santafe.edu (C.M. Reidys).

0012-365X/$ – see front matter
doi:10.1016/j.disc.2011.07.027

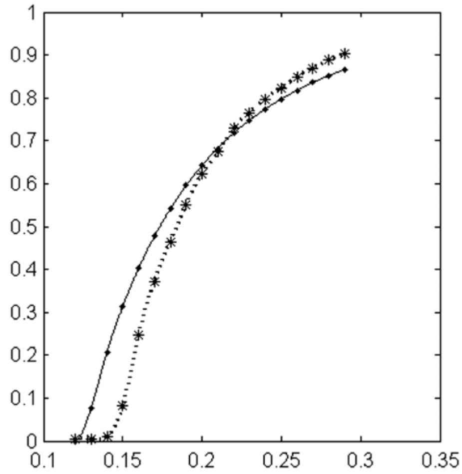

Fig. 1. The evolution of the giant component in random induced subgraphs of $\Gamma(S_9, P_9)$. We display the relative size of the giant component $\frac{|C_9^{(1)}|}{|\Gamma_9|}$ as a function of $\lambda_9 = (1+\epsilon)/8$ as data-curve $(* \cdots *)$ versus the growth predicted by Theorem 1 (solid line with dots).

has proved to be hard and motivated models better suited for computation. The case in point is the work of Eriksen and Hultman [19] where the authors derive a closed formula for the expected transposition distance and subsequently show how to use it as an approximation of the expected reversal distance. Berestycki and Durrett [8] studied the shortest distance of random walks over Cayley graphs generated by all transpositions and canonical transpositions, respectively, and compared the shortest distance with the expected distance [19].

The theory of random graphs was pioneered by Erdös and Rényi in the late 1950s [17,18], who analyzed the phase transition of $G(n, p_n)$, the random graph containing n vertices in which an edge $\{i, j\}$ is selected with independent probability p_n. For $p_n = \frac{c}{n}$ and $c < 1$, the largest component in $G(n, p_n)$ is a.s. of size $O(\log n)$. For $p_n = \frac{1+\theta \cdot n^{-\frac{1}{3}}}{n}$, where $\theta > 0$, a.s. a largest component of size $O\left(n^{\frac{2}{3}}\right)$ emerges. For $p_n = \frac{c}{n}$ and $c > 1$, we have a.s. a unique largest component of size $O(n)$ and all other components are smaller than $O(\log n)$. Erdös and Rényi's construction of the giant component [17,18] has motivated Lemma 3, which assures the existence of certain subtrees of size $\left\lfloor \frac{1}{4} n^{\frac{2}{3}} \right\rfloor$. For a review of Erdös–Rényi random graph theory, see [16] or [42].

In this paper we study a subgraph of the Cayley graph generated by all transpositions, the Cayley graph $\Gamma(S_n, T_n)$, where T_n is a minimal generating set of transpositions. Setting $T_n = P_n$ and $T_n = B_n$ we can recover the star- and the bubble-sort graph as particular instances. We study structural properties of $\Gamma(S_n, T_n)$ in terms of the random graph obtained by selecting permutations with independent probability (see Fig. 1 for the conclusion of Theorem 1 at $n = 9$). The main result of this paper is the following theorem.

Theorem 1. *Let* $\lambda_n = \frac{1+\epsilon_n}{n-1}$, *where* $n^{-\frac{1}{3}+\delta} \leq \epsilon_n < 1$ *and* $\delta > 0$. *Let* T_n *be a minimal generating set of transpositions and let* Γ_n *denote the random induced subgraph of* $\Gamma(S_n, T_n)$, *obtained by independently selecting each permutation with probability* λ_n. *Then* Γ_n *has a.s. a unique giant component,* $C_n^{(1)}$, *whose size is given by*

$$|C_n^{(1)}| = (1 + o(1)) \cdot x(\epsilon_n) \cdot \frac{1+\epsilon_n}{n-1} \cdot n!, \tag{1.1}$$

where $x(\epsilon_n) > 0$ *is the survival probability of a Poisson branching process with parameter* $\lambda = 1+\epsilon_n$ *and also the unique positive root of* $e^{-(1+\epsilon_n)y} = 1 - y$. *Particularly, if* $n^{-\frac{1}{3}+\delta} \leq \epsilon_n = o(1)$, *then we have* $x(\epsilon_n) = (2 + o(1))\epsilon_n$.

In contrast to vertex-induced random graphs, edge-induced random graphs have been studied quite extensively. Random induced subgraphs of n-cubes [9,37], as well as $G(n, p_n)$ and random induced subgraphs of $\Gamma(S_n, T_n)$ exhibit a giant component for very small vertex selection probabilities. One might speculate that the critical probability $p_n = \frac{1+\theta \cdot n^{-\frac{1}{3}}}{n}$ is determined by the size of the generator set. Note that $|T_n| = n-1$ holds for any minimal generating set of transpositions and the size of the generator set for n-cube is n. Specific properties of n-cubes, like for instance, the isoperimetric inequality [23], do not play a key role for establishing the existence of the giant component. The isoperimetric inequality depends on an inductive argument using particular properties of a linear ordering of the vertices of an n-cube. This induction cannot be carried out for Cayley graphs over canonical transpositions. In this paper any argument involving (vertex) boundaries follows from a generic estimate of the vertex boundary in Cayley graphs due to Aldous and Diaconis [4], Babai [6].

The paper is organized as follows: after introducing in Section 2 our notation and some basic facts about branching processes, we analyze in Section 3 vertices contained in polynomial size subcomponents. The strategy is similar to that