Theoretical Computer Science 524 (2014) 103–125

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Constructions of independent sets in random intersection graphs

Katarzyna Rybarczyk

Adam Mickiewicz University, Poznań, Poland

ARTICLE INFO

Article history:
Received 27 February 2013
Received in revised form 21 November 2013
Accepted 8 January 2014
Communicated by J. Díaz

Keywords:
Random intersection graph
Independent (stable) set
Greedy algorithm

ABSTRACT

This paper concerns constructing independent sets in a random intersection graph. We concentrate on two cases of the model: a binomial and a uniform random intersection graph. For both models we analyse two greedy algorithms and prove that they find asymptotically almost optimal independent sets. We provide detailed analysis of the presented algorithms and give tight bounds on the independence number for the studied models. Moreover we determine the range of parameters for which greedy algorithms give better results for a random intersection graph than this is in the case of an Erdős–Rényi random graph $G(n, \hat{p})$.

1. Introduction

The random intersection graph model was introduced by Karoński, Scheinerman and Singer-Cohen [16] and generalised by Godehardt and Jaworski [14]. We state the general definition after [14]. Let $\mathcal{V} = \{v_1, \dots, v_n\}$ and $\mathcal{W} = \{w_1, \dots, w_m\}$ be two disjoint sets. We call $\mathcal{V}$ the vertex set and $\mathcal{W}$ the feature set. Let moreover $\mathcal{P}_{(m)} = \{P_0, \dots, P_m\}$ be a probability distribution. A random intersection graph $\mathcal{G}(n, m, \mathcal{P}_{(m)})$ is a graph with the vertex set $\mathcal{V}$ in which each vertex $v \in \mathcal{V}$ is assigned a subset of features $\mathcal{W}(v) \subseteq \mathcal{W}$ independently from all other vertices. The subset of features $\mathcal{W}(v) \subseteq \mathcal{W}$ is chosen uniformly at random from all d-element subsets of $\mathcal{W}$, where the cardinality d is determined according to $\mathcal{P}_{(m)}$. Two vertices v_1 and v_2 are adjacent in a random intersection graph if $\mathcal{W}(v_1) \cap \mathcal{W}(v_2) \neq \emptyset$. For any $v \in \mathcal{V}$ and $w \in \mathcal{W}$, if $w \in \mathcal{W}(v)$ we say that v *has chosen* w or that w *is chosen* by v.

We will concentrate on analysing two most natural random intersection graph models. Let $d = d(n)$ be a sequence of positive integers and $p = p(n) \in (0, 1)$. A uniform random intersection graph $\mathcal{G}_u(n, m, d)$ is a random intersection graph $\mathcal{G}(n, m, \mathcal{P}_{(m)})$ in which $\mathcal{P}_{(m)}$ is the distribution with point mass in d, i.e. in $\mathcal{G}_u(n, m, d)$ for all $v \in \mathcal{V}$ the set $\mathcal{W}(v)$ is chosen uniformly at random from all d-element subsets of $\mathcal{W}$. A binomial random intersection graph $\mathcal{G}_{bin}(n, m, p)$ is a random intersection graph $\mathcal{G}(n, m, \mathcal{P}_{(m)})$ with $\mathcal{P}_{(m)}$ – the binomial distribution $\mathrm{Bin}(m, p)$, i.e. for each pair $(v, w) \in \mathcal{V} \times \mathcal{W}$, w is chosen by v (added to $\mathcal{W}(v)$) with probability $p = p(n)$ independently of all other pairs.

By an independent set in a graph we mean any subset of vertices inducing no edge. The size of the maximum independent set of G we call the independence number of G and denote by $\alpha(G)$. In the article we analyse two algorithms, which construct independent sets in $\mathcal{G}_u(n, m, d)$ and $\mathcal{G}_{bin}(n, m, p)$. Moreover we find asymptotic bounds on $\alpha(\mathcal{G}_u(n, m, d))$ and $\alpha(\mathcal{G}_{bin}(n, m, p))$ and compare them with the size of the independent set generated by the algorithms.

The aim of this paper is, above all, to complement and extend the results presented in the articles of Nikoletseas, Raptopoulos and Spirakis [20,19]. First of all we give tight bounds on the value of the independence number for an important range of parameters. Moreover we provide a detailed analysis of the performance of the greedy algorithm on random intersection graphs $\mathcal{G}_{bin}(n, m, p)$ and $\mathcal{G}_u(n, m, d)$. Last but not least we propose a new greedy algorithm, which rely heavily on the structure of intersection graphs and in many cases performs far better than the classical one.

We analyse asymptotic properties of random graphs. We say that a random graph on n vertices has a property with high probability if probability that it has the property is tending to 1 as $n \to \infty$. Throughout the paper we use standard

0304-3975/$ – see front matter
http://dx.doi.org/10.1016/j.tcs.2014.01.006

asymptotic notations $o(\cdot)$, $O(\cdot)$, $\Omega(\cdot)$, $\Theta(\cdot)$, $\ll$, $\gg$, $\sim$, defined as in [15]. All inequalities hold for large n. For clarity of notation we usually omit $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$, when it does not affect the proof. In the statements of the theorems all limits as well as asymptotic notation are taken as $n \to \infty$. Moreover by $\mathrm{Bin}(n, p)$ and $\mathrm{Po}(\lambda)$ we denote the binomial distribution with parameters n and p and the Poisson distribution with a parameter λ, respectively. For any $V \subseteq \mathcal{V}$ we set $\mathcal{W}(V) = \bigcup_{v\in V} \mathcal{W}(v)$.

1.1. Motivation

Since appearance of the two seminal papers of Erdős and Rényi [8,9] random graphs have become one of the major tools in theoretical computer science. The classical so-called Erdős–Rényi graph $G(n, \hat{p})$ introduced by Gilbert [13] has a vertex set $\mathcal{V} = \{v_1, \ldots, v_n\}$ and each pair of vertices $\{v_i, v_j\}$ is connected by an edge in $G(n, \hat{p})$ independently with the same probability $\hat{p}$. $G(n, \hat{p})$ exhibits no edge appearance dependency, therefore it does not fit to many real life applications. Many more suitable models have been introduced (see for instance Chapter 4 in [1]). One of the most promising models is a random intersection graph. It has already shown to be a good tool to study many problems in theoretical computer science, for example: the problem of gate matrix layout for VLSI design (see [16]), designing wireless sensor networks with random key predistribution (see for example [21]), modelling of complex networks such as internet network on internet router level, internet network on autonomous systems level or World Wide Web (see for example [3,4,7,24]).

The problem of determining the value of the independence number for any random graph model is interesting on its own. However in the case of a random intersection graph the issue of finding the optimal independent set is even more attractive, due to possible applications. In this context we should stress those connected to the networks modelling. In [21] it is suggested that $\mathcal{G}_u(n, m, d)$ is a good model of wireless sensor networks with random key predistribution. Namely vertices of $\mathcal{G}_u(n, m, d)$ may represent sensors and for every $v \in \mathcal{V}$, $\mathcal{W}(v)$ may be the set of keys that sensor v has. Two sensors communicate in the wireless sensor network if they share at least one key used to decode information, i.e. when vertices representing them have chosen at least one common feature. Last but not least the analysis of the structure of $\mathcal{G}_{bin}(n, m, p)$ gives a flavour of how the algorithms work in any other random intersection graph model $\mathcal{G}(n, m, \mathcal{P}_{(m)})$, thus also in the case of the complex networks (see for example [3,4,7,24]).

1.2. Related work

The problem of finding the independence number for random intersection graphs have been studied in the Master thesis of Ueckerdt [25] and in two papers of Nikoletseas, Raptopoulos and Spirakis [20,19]. The results from [25] give very rough bounds on the independence number only in some particular cases. In [20] Nikoletseas, Raptopoulos and Spirakis propose three greedy algorithms for finding independent sets and analyse roughly their behaviour on $\mathcal{G}_{bin}(n, m, p)$ for $mp = \Theta(\ln n)$. In [19] the independence number of $\mathcal{G}_u(n, m, d)$ is studied but with no relation to the algorithms presented in [20].

Finding the asymptotic value of the independence number for Erdős–Rényi random graph $G(n, \hat{p})$ has been a long lasting open problem. Even though in the dense case it was resolved by Bollobás and Erdős [6] and Matula [18], it took long time to extend the result to the sparse case by Frieze [12]. In particular it has been shown that for $1/n \ll \hat{p} \ll 1$ with high probability

$$\alpha\big(G(n, \hat{p})\big) \sim \frac{2\ln n\hat{p}}{\hat{p}} \quad \text{and} \quad \alpha_1\big(G(n, \hat{p})\big) \sim \frac{\alpha(G(n, \hat{p}))}{2}, \tag{1}$$

where α_1 denotes the size of the independent set in $G(n, \hat{p})$ constructed by the greedy algorithm. We should stress here that the phase transition of $G(n, \hat{p})$ occurs when $\hat{p} \sim 1/n$, thus the stated results concerns a random graph which with high probability has the giant component.

2. Main results

The main results presented in the paper concern analysing greedy algorithms constructing independent sets in $\mathcal{G}_u(n, m, d)$ and $\mathcal{G}_{bin}(n, m, p)$. The first of the algorithms is the standard greedy one, however we write it for clarity of notation. Its efficiency was studied in [20] in the special case of $\mathcal{G}_{bin}(n, m, p)$ with $mp = \Theta(\ln n)$, however only some rough bounds were given. The second algorithm is new and it is defined only for intersection graphs.

For any graph G and its vertex v, denote by $N_G(v)$ the set of all neighbours of v in G. For any family of sets $\mathbb{W} = \{\mathcal{W}(v) \subseteq \mathcal{W}\colon\ v \in \mathcal{V}\}$, by the intersection graph generated by $\mathbb{W}$ we mean a graph with the vertex set $\mathcal{V}$ in which v_1 and v_2 are connected by an edge if $\mathcal{W}(v_1) \cap \mathcal{W}(v_2) \neq \emptyset$. In the following reasoning we assume that in the ordering constructed in Algorithm 2 two vertices with feature sets of the same cardinality are ordered according to the Input ordering.

For any intersection graph G, let $\alpha_1(G)$ and $\alpha_2(G)$ be the size of the independent sets $\mathcal{S}_1(G)$ and $\mathcal{S}_2(G)$, respectively. By definition, for any intersection graph G

$$\alpha_1(G) \leqslant \alpha(G) \quad \text{and} \quad \alpha_2(G) \leqslant \alpha(G).$$

In the theorems below we will not mention the bounds on α, which follow by those relations.