Discrete Mathematics 341 (2018) 1275–1281

Contents lists available at ScienceDirect

Discrete Mathematics

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

Geodesic cycles in random graphs

Yuanzhi Li, Lingsheng Shi *

Institute for Interdisciplinary Information Sciences and Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Article history:
Received 17 September 2015
Received in revised form 12 January 2018
Accepted 15 January 2018

Keywords:
Geodesic cycle
Hyperbolic
Random graph

ABSTRACT

A cycle in a graph is geodesic if the distance of each pair of nodes on the cycle coincides with their distance restricted on the cycle. In this article, we prove that a random graph in $\mathcal{G}(n, p)$ has a geodesic cycle of length at least $(2-\epsilon)\log_d n$ with probability tending to one as $n \to \infty$ for $d > 1+\epsilon$ with $d = np$ and each small positive constant ϵ. This lower bound on the length of the longest geodesic cycle is almost tight since the diameter of the giant component in the random graph is asymptotically almost surely within $(1 \pm \epsilon)\log_d n$ for sufficiently large d. Taking four nodes that split the geodesic cycle into four paths of approximately the same length implies that the giant component in the random graph is a.a.s. not $(1/2 - \epsilon)\log_d n$-hyperbolic. This bound on the hyperbolicity improves a super-constant bound of Narayan–Saniee–Tucci and also comes close to its exact value for $d \gg \ln^5 n/(\ln\ln n)^2$ which is obtained by Mitsche and Prałat.

1. Introduction

Random graphs were first introduced by Erdős and Rényi [9] where the authors considered a graph chosen uniformly at random from all (undirected) graphs with n nodes and m edges. A closely related model of random graphs denoted by $\mathcal{G}(n, p)$ due to Gilbert [10] is defined on a node set of size n with each pair of nodes appearing as an edge independently with probability p. For more information on random models, we refer to the monographs [6,12]. For decades, the random graph remains a hot topic. In recent years, many different models of random graphs have been introduced to depict networks with some particular features, for instance, the power law distribution [1,21] and small world phenomenon [13,20]. However, research often goes back to the Erdős–Rényi model when studying new properties of random graphs.

For two functions $f(n)$ and $g(n)$, we say that $f(n) = O(g(n))$ or $g(n) = \Omega(f(n))$ if there exists an integer n_0 and a constant $c > 0$ such that $|f(n)| < c|g(n)|$ for all $n > n_0$, and also $f(n) = o(g(n))$ or $g(n) \gg f(n)$ if $|g(n)|/|f(n)|$ tends to infinity together with n. Throughout this article, $\ln n$ always denotes the natural logarithm of n.

The longest cycles in random graphs were studied by Bollobás et al. [5,7] and Łuczak [14]. The longest induced cycles in random graphs were investigated by Łuczak [15,16]. In this article, we study the longest geodesic cycles of random graphs in $\mathcal{G}(n, p)$. We say that a property holds *asymptotically almost surely* (a.a.s. for short) in $\mathcal{G}(n, p)$ if its probability tends to one as n tends to infinity. A cycle in a graph G is *geodesic* if the distance of each pair of nodes on the cycle coincides with their distance restricted on the cycle. Formally, suppose that u and v are two nodes of G and let $d_G(u, v)$ denote the distance (the length of a shortest path) of u and v in G. A geodesic cycle C in G is a cycle with the property that $d_G(u, v) = d_C(u, v)$ for all pairs of nodes u and v in C. Geodesic cycles were first analyzed by Negami and Xu [19] when studying self-centered graphs. A shortest cycle is obviously geodesic, while a longest one is not in general. Benjamini et al. [2] were the first to explicitly study geodesics in random graphs. In [2] it is proven that a.a.s. every pair of nodes in a random d-regular graph lies in an

* Corresponding author.
E-mail address: lshi@math.tsinghua.edu.cn (L. Shi).

https://doi.org/10.1016/j.disc.2018.01.014

almost geodesic cycle of length $2\log_{d-1} n \pm o(\ln n)$. It is well known that the model $\mathcal{G}(n, p)$ of sparse random graphs has a.a.s. different component structure in different regimes for different values of p:

1. If $p = o(1/n)$, then a.a.s the random graph consists only of many small trees.
2. If $p = d/n$ for a constant $d \in (0, 1)$, then a.a.s. the random graph consists of trees and unicyclic components, where the largest component is a tree of order $O(\ln n)$.
3. If $p = d/n$ for a constant $d > 1$, then a.a.s. a giant component emerges on cn nodes with c depending only on d, and the random graph is a union of the giant component, small unicyclic components and small trees. There are at most $f(n)$ nodes on the unicyclic components, where $f(n)$ is any slow function tending to infinity as $n \to \infty$.
4. If $p = c \ln n/n$ with a constant $c > 1$, then the random graph is a.a.s. connected.

We show that a.a.s. there is a geodesic cycle of length at least $(2 - \epsilon)\log_d n$ in the (giant component of) random graph from $\mathcal{G}(n, d/n)$ for $d > 1 + \epsilon$ with each small positive constant ϵ. This result is almost tight in view of the diameter in the random graph. Bollobás [4] showed that the random graph in $\mathcal{G}(n, d/n)$ is a.a.s. of diameter concentrated on at most two values about $\log_d n + \log_d \log_d n + O(1)$ for $d \gg \log^3 n$. Let $f(n, d) = \log_d n + 2\log_{1/d_*} n$ with $d_* < 1$ and $d_* e^{-d_*} = de^{-d}$. Riordan and Wormald [22] proved that the giant component of the random graph in $\mathcal{G}(n, d/n)$ is a.a.s. of diameter $f(n, d) + O(1)$ for fixed $d > 1$ and of diameter concentrated on at most two values about $f(n, d) + O(1)$ for $d = d(n) \to \infty$ and $d \leq n^{1/1000}$. These results imply that the random graph in $\mathcal{G}(n, d/n)$ is a.a.s. of diameter within $(1\pm\epsilon)\log_d n$ for sufficiently large d and thus it cannot have a geodesic cycle of length greater than $(2+\epsilon)\log_d n$. Moreover, different from the approach of Benjamini et al., our proof directly applies the method of second moments and branching process, and it lends itself to future improvements. More precisely, we obtain the following result.

Theorem 1.1. *Let $n \in \mathbb{N}$ and ϵ be a small positive constant. There exists a.a.s. a geodesic cycle of length at least $(2 - \epsilon)\log_d n$ in the random graph from $\mathcal{G}(n, d/n)$ for $d \in (1 + \epsilon, n)$.*

As pointed out by Benjamini et al. [2], there are connections between geodesic cycles and other geometric properties of graphs, among which is the hyperbolicity. Following Gromov [11], the *hyperbolicity* of a graph G is the smallest nonnegative number δ such that for every four nodes u, v, w, x in G, each pair of values in the set $\{d(u, v) + d(w, x), d(u, w) + d(v, x), d(u, x) + d(v, w)\}$ differs by at most 2δ. The Gromov hyperbolicity is used as a soft measure of the tree-likeness of a graph since a tree is clearly 0-hyperbolic by definition. By taking four nodes that split the cycle into four paths of approximately the same length, one can easily verify that Theorem 1.1 implies the giant component in the random graph is a.a.s. not $(1/2-\epsilon)\log_d n$-hyperbolic. This bound considerably improves the super-constant hyperbolicity of Narayan et al. [18] and also comes close to its exact value for $d \gg \ln^5 n/(\ln\ln n)^2$ which is obtained by Mitsche and Prałat [17]. In practice, there are several empirical studies indicating hyperbolicity may be a factor for decentralized search in complex networks [3]. Since no result is known about how to efficiently decentralize a random graph in $\mathcal{G}(n, p)$, our result about the existence of long geodesic cycles also gives an explanation for the hardness of finding short paths.

The remainder of the article is organized as follows. In Section 2 we give a lower bound on the distance between a fixed pair of nodes in a random graph based on the standard Galton–Watson branching process. In Section 3 we show that a fixed cycle is geodesic with high probability and extend the Erdős–Rényi method of showing the existence of balanced graphs of constant order to a more involved analysis about the property of (geodesic) cycles. We then use the standard method of second moments to bound the probability of the existence of geodesic cycles.

2. Distance of a pair of nodes

In this section, we show that with high probability, a fixed pair of nodes in a random graph $\mathcal{G}(n, d/n)$ is at distance at least $(1 - \epsilon)\log_d n$. The upper bound on the distance is also studied by Chung and Lu [8] and by Riordan and Worwald [22] for bounding the diameter of random graphs from above.

Theorem 2.1. *Let $n \in \mathbb{N}$ and $a = a(\epsilon) = e^{-\epsilon}(1 + \epsilon)^{1+\epsilon} > 1$ for a constant $\epsilon \in (0, 1)$, and let u and v be two different nodes of a random graph in $\mathcal{G}(n, d/n)$ with $d \in (1, n)$. For every integer $t \geq 1$, we have*

$$\Pr[d(u, v) \leq t] < \frac{2\log_a n}{n}\left(\frac{[(1 + \epsilon)d]^{t+1}}{(1 + \epsilon)d - 1} + \frac{3et^2[(1 + \epsilon)d]^t}{n}\right).$$

Theorem 2.1 implies the following result by taking ϵ small enough.

Corollary 2.1. *Let $n \in \mathbb{N}$ and ϵ be a positive constant, and let u and v be two different nodes of a random graph in $\mathcal{G}(n, d/n)$ with $d \in (1 + \epsilon, n)$. For $t \leq (1 - \epsilon)\log_d n$, we have*

$$\Pr[d(u, v) \leq t] = o\left(n^{-\epsilon/3}\right).$$

Let $t \in \mathbb{N}$, and we write $N_t(u)$ as the set of nodes at distance at most t to u in a random graph $G \in \mathcal{G}(n, p)$. By symmetry, $\Pr[d(u, v) \leq t] = \Pr[d(u, w) \leq t]$ for every $w \neq u$, we can therefore take a node w uniformly at random from $V \setminus \{u\}$ where $V = V(G)$ so that

$$\Pr[d(u, w) \leq t,\ w \in V \setminus \{u\}] = \Pr[d(u, v) \leq t]$$