ASYMPTOTIC BEHAVIOR OF THE EDGE METRIC DIMENSION OF THE RANDOM GRAPH

NINA ZUBRILINA

Department of Mathematics
Stanford University

e-mail: nina57@stanford.edu

Abstract

Given a simple connected graph $G(V, E)$, the edge metric dimension, denoted $\text{edim}(G)$, is the least size of a set $S \subseteq V$ that distinguishes every pair of edges of $G$, in the sense that the edges have pairwise different tuples of distances to the vertices of $S$. In this paper we prove that the edge metric dimension of the Erdős-Rényi random graph $G(n, p)$ with constant $p$ is given by

$$\text{edim}(G(n, p)) = (1 + o(1)) \frac{4 \log n}{\log(1/q)},$$

where $q = 1 - 2p(1 - p)^2(2 - p)$.

Keywords: random graph, edge dimension, Suen's inequality.

2010 Mathematics Subject Classification: 05C12, 05C80.

1. Introduction

Let $G(V, E)$ be a finite, simple, connected graph, and define the distance $d(x, y)$ between two vertices $x, y \in V$ to be the length of the shortest path connecting $x$ and $y$. The metric dimension of $G(V, E)$, denoted $\text{dim}(G(V, E))$, is the minimal cardinality of a set $S \subseteq V$ such that for any distinct $x, y \in V$ there exists $v \in S$ which satisfies $d(v, x) \neq d(v, y)$.

The metric dimension was introduced by Slater [12] in 1975 in connection with the problem of uniquely recognizing the location of an intruder in a network, and independently by Harary and Melter in [4] a year later. Graphs with $\text{dim}(G) = 1$ and 2 were characterized in [8], and graphs with $\text{dim}(G) = |V| - 1$ and $|V| - 2$ were described in [3]. This graph invariant is useful in areas like robot navigation [8], image processing [10], and chemistry [2, 3, 6].
In [1], Bollobás, Mitsche and Pralat computed the asymptotic behavior at infinity of the metric dimension of the Erdős-Rényi random graph for a wide range of probabilities \( p(n) \) (viewed as functions of \( n \)). For instance, for constant \( p \in (0, 1) \), it was shown that

\[
\dim(G(n, p)) = (1 + o(1)) \frac{2 \log n}{\log(1/Q)} ,
\]

where \( Q = p^2 + (1 - p)^2 \). In this paper we generalize those calculations to a variation on the metric dimension called the edge metric dimension, introduced by Kelenc, Tratnik and Yero in [7] in 2016. While the metric dimension is about uniquely identifying the vertices of a graph in terms of distances to a set, the edge metric dimension is about identifying the edges of a graph in the same way.

For an edge \( e = xy \in E \) and a vertex \( v \in V \), let \( d(e, v) = \min\{d(x, v), d(y, v)\} \).

The edge metric dimension (denoted \( \text{edim} \)) of a graph \( G(V, E) \) is defined as the minimal cardinality of a set \( S \subseteq V \) such that for any distinct \( e_1, e_2 \in E \), there exists \( v \in S \) satisfying \( d(v, e_1) \neq d(v, e_2) \).

Kelenc, Tratnik and Yero computed the edge metric dimension of a range of families of graphs, showed \( \text{edim}(G) \) can be less, equal to, or more than \( \dim(G) \), and showed computing \( \text{edim}(G) \) is NP-hard in general ([7]). Zubrilina ([13]) showed that the \( \text{edim}(G) / \dim(G) \) ratio is not bounded from above and classified graphs \( G \) with \( \text{edim}(G) = |V| - 1 \). Kratica, Filipović and Kartelj studied the edge metric dimension of the generalized Petersen graph \( GP(n, k) \) in [9]. In this paper, we prove the following theorem.

**Theorem 1.1.** Let \( G(n, p) \) be the Erdős-Rényi random graph with constant \( p \). Then

\[
\text{edim}(G(n, p)) = (1 + o(1)) \frac{4 \log n}{\log(1/q)} ,
\]

where \( q = 1 - 2p(1 - p)^2(2 - p) \).

For a set \( R = \{r_1, \ldots, r_{|R|}\} \subseteq V \), we define the distance tuple \( d_R : V \cup E \to \mathbb{N}^{|R|} \) via \( (d_R(x))_i = d(x, r_i) \). We say \( R \) distinguishes \( v_1, v_2 \in V \) if \( d_R(v_1) \neq d_R(v_2) \), and similarly that \( R \) distinguishes \( e_1, e_2 \in E \) if \( d_R(e_1) \neq d_R(e_2) \). \( R \) is a generating set of \( G \) if it distinguishes any two distinct vertices, and an edge generating set if it distinguishes any two distinct edges of \( G \).

We say \( f(n) = \Theta(g(n)) \) if there exists a constant \( C > 0 \) such that \( \lvert f(n) \rvert \leq C \lvert g(n) \rvert \), and \( f(n) = o(g(n)) \) if \( f = g \cdot o(1) \), where \( o(1) \xrightarrow{n \to \infty} 0 \).

We say a property holds asymptotically almost surely (denoted a.a.s.) for the random graph if the probability that it holds for \( G(n, p) \) goes to 1 as \( n \) goes to infinity. We denote probability with \( P \) and expected value with \( E \). All the graphs are assumed to be finite, simple, connected and undirected.
2. The Upper Bound

In this section we prove the following theorem.

**Theorem 2.1.** For the random graph $G(n, p)$ with $p$ constant, we have

$$\text{edim}(G(n, p)) \leq (1 + o(1))\frac{4 \log n}{\log(1/q)},$$

where $q = 1 - 2p(1 - p)^2(2 - p)$.

In order to prove Theorem 2.1, we will need some lemmas.

**Lemma 2.2.** Let $G = G(n, p)$ be the random graph, and let $V, E$ denote its vertex and edge sets. Let $\omega \in \{1, \ldots, n\}$ be such that for any two distinct edges $e_1, e_2 \in E$, a uniformly random subset $W \subseteq V$ of size $|W| = \omega$ satisfies

$$P(W \text{ does not distinguish } e_1, e_2) \leq \frac{1}{n^4 p^2}.$$

Then

$$\text{edim}(G) \leq \omega.$$

**Proof.** We use the probabilistic method. Note that

$$\mathbb{E}[|E|] = p\binom{n}{2} < pn^2/2,$$

so the expected number of distinct pairs of edges is no more than $\binom{pn^2/2}{2} \leq p^2n^4/8$. Then by our hypothesis the expected number of pairs not distinguished by some $W \subseteq V$ with $|W| = \omega$ is less than $p^2n^4/8p^2n^4 = 1/8$. Since this is strictly less than 1, there must be at least one such set $W$ that distinguishes all the pairs. $lacksquare$

**Lemma 2.3.** In $G(n, p)$, the probability that a vertex $v$ doesn’t distinguish two uniformly random edges $e_1, e_2$ is $(1 + o(1))q$, where $q = 1 - 2p(1 - p)^2(2 - p)$.

**Proof.** There are two types of distinct edge pairs.

1. $ab, bc$ for some $a, b, c \in V$.
2. $ab, cd$ for $a, b, c, d \in V$ and $\{a, b\} \cap \{c, d\} = \emptyset$.

Note that

the expected number of type 2 pairs $= 3\binom{n}{4}p^2 = \frac{n^4p^2}{8}(1 + o(1)),$

and

the expected number of type 1 pairs $\leq n^3 = o\left(\frac{n^4p^2}{8}\right).$
Thus, we can neglect the type 1 pairs. Let $xy, zt$ be a type 2 pair and $v$ a uniformly random vertex. Clearly, $\mathbb{P}(v \in \{x, y, z, t\}) = o\left(\frac{n^4p^2}{S}\right)$, so we can assume $v$ is not a vertex of $xy$ or $zt$. Since the random graph has diameter 2 a.a.s. (see [11]), $v$ has distance 1 or 2 to $x, y, z, t$ a.a.s.; moreover, $\mathbb{P}(d(v, x) = 1) = p$, so a.a.s. $\mathbb{P}(d(v, x) = 2) = 1 - p$. It is easy to see that $v$ has distance 1 to $xy$ and 2 to $zt$ if and only if one of the following cases holds.

1. $(d(v, x), d(v, y), d(v, z), d(v, t)) = (1, 1, 2, 2)$ (with probability $p^2(1 - p)^2$).
2. $(d(v, x), d(v, y), d(v, z), d(v, t)) = (1, 2, 2, 2)$ (with probability $p(1 - p)^3$).
3. $(d(v, x), d(v, y), d(v, z), d(v, t)) = (2, 1, 2, 2)$ (with probability $p(1 - p)^3$).

The same probabilities hold for $xy$ and $zt$ switched. Thus, a.a.s. $\mathbb{P}(v$ distinguishes $xy, zt) = (1 + o(1)) \cdot 2[p^2(1 - p)^2 + 2p(1 - p)^3] = (1 + o(1)) \cdot 2p(1 - p)^2(2 - p) = (1 + o(1))(1 - q)$.

This gives us the desired result.

Lemma 2.4. Let $V, E$ be the vertex and edge sets of $G(n, p)$. Consider a uniformly random subset $W \subseteq V$ with $|W| = (1 + o(1)) \frac{4\log n}{\log(1/q)}$.

Then for uniformly random $e_1$ and $e_2 \in E$,

$$
\mathbb{P}(W \text{ does not distinguish } e_1, e_2) \leq (1 + o(1))/n^4p^2.
$$

Proof. Using Lemma 2.3, we see that

$$
\mathbb{P}(W \text{doesn’t distinguish } e_1, e_2)
\leq (1 + o(1))\mathbb{P}(\text{uniformly random vertex } v \text{ doesn’t distinguish } e_1, e_2)^{|W|}
\leq (1 + o(1))q^{(1+o(1))\frac{4\log n}{\log(1/q)}} = (1 + o(1))q^{-\log_q(n^4)}
= (1 + o(1))\frac{1}{n^4} \leq (1 + o(1))\frac{1}{p^2n^4}.
$$

Proof of Theorem 2.1. Combining Lemmas 2.4 and 2.2, we see that $\text{edim}(G(n, p))$ is at most

$$
(1 + o(1)) \frac{4\log n}{\log(1/q)},
$$

which concludes the proof of Theorem 2.1.
3. The Lower Bound

The goal of this section is to prove the following theorem.

**Theorem 3.1.** For the random graph $G(n, p)$ with $p$ constant, we have

$$\text{edim}(G(n, p)) \geq (1 + o(1)) \frac{4 \log n}{\log(1/q)},$$

where $q = 1 - 2p(1-p)^2(2-p)$.

Let

$$\varepsilon := 3 \log \frac{\log n}{\log n} = o(1).$$

We will show that a.a.s. there is no edge generating set $R$ of cardinality less than

$$r := (4 - \varepsilon) \frac{\log n}{\log(1/q)}.$$

To do that we will use a theorem which is a version of Suen’s inequality demonstrated by Janson in [5]. First we introduce some notation

- $\{I_i\}_{i \in \mathcal{I}}$ — a finite family of indicator random variables;
- $\Gamma$ — the associated dependency graph ($\mathcal{I}$ is the set of vertices of $\Gamma$);
- For $i, j \in \mathcal{I}$, write $i \sim j$ if $i, j$ are adjacent in $\Gamma$;
- $\mu := \sum_i P(I_i = 1)$;
- $\Delta := \sum_{i \sim j} E[I_i I_j]$;
- $\delta := \max_i \sum_{i \sim j} P(I_j)$;
- $S := \sum_i I_i$.

**Theorem 3.2** (Suen’s inequality, Theorem 2 of [5]).

$$\mathbb{P}(S = 0) \leq \exp \left( -\mu + \Delta \varepsilon^{2\delta} \right).$$

We now apply this theorem to our problem.

Let $V, E$ be the vertex and edge sets of $G(n, p)$. Let $R \subseteq V$ with $|R| = r$. Let

$$\mathcal{I} := \{(xy, zt) \mid xy, zt \in E, xy \neq zt\}$$

be the set of pairs of distinct edges, and for any $(xy, zt) \in \mathcal{I}$ let $A_{xy,zt}$ be the event $d_R(xy) = d_R(zt)$ (with $I_{xy,zt}$ being the corresponding indicator function). Let $S = \sum_{(xy, zt) \in \mathcal{I}} I_{xy,zt}$. Then

$$\mathbb{P}(R \text{ is an edge generating set}) = \mathbb{P}(S = 0).$$
The associated dependency graph has $I$ as vertices and $(x_1y_1, z_1t_1) \sim (x_2y_2, z_2t_2)$ if and only if \{x_1, y_1, z_1, t_1\} \cap \{x_2, y_2, z_2, t_2\} \neq \emptyset (here, again, $\sim$ denotes adjacency). Then by Theorem 3.2,

\[ P(S = 0) \leq \exp(-\mu + \Delta \varepsilon^2), \]

where

\[ \mu = \sum_{(e,f) \in I} P(A_{e,f}), \]

\[ \Delta = \sum_{(e_1,f_1) \sim (e_2,f_2)} \mathbb{E}[I_{e_1f_1}I_{e_2f_2}], \]

\[ \delta = \max_{(e_1,f_1) \in I} \sum_{(e_2,f_2) \sim (e_1,f_1)} P(A_{e_2f_2}). \]

We now show the following estimate for $\mu$.

**Lemma 3.3** (Evaluation of $\mu$).

\[ \mu = (1 + o(1))p^2n^{\varepsilon}/8. \]

**Proof.** Using Lemma 2.3, we can derive that that

\[ P(A_{e,f}) = (1 + o(1))q^r, \]

so, since the expected number of pairs is $(1 + o(1))(n^4p^2/8)$, we indeed get

\[ \mu = (1 + o(1))n^4p^2q^r/8. \]

Since $r = \frac{(4-\varepsilon)\log n}{\log(1/q)}$,

\[ q^r = q^{-(4-\varepsilon)\log_q(n)} = n^{(4-\varepsilon)/4}. \]

Thus,

\[ (1 + o(1))n^4p^2q^r/8 = (1 + o(1))n^4p^2n^{(4-\varepsilon)/4}/8 = (1 + o(1))p^2n^{\varepsilon}/8. \]

This means that, indeed,

\[ \mu = (1 + o(1))p^2n^{\varepsilon}/8. \]

Now we estimate $\Delta$ and show the following.

**Lemma 3.4** (Evaluation of $\Delta$).

\[ \Delta = o(\mu). \]
Proof.

Claim 3.5. In calculating $\Delta$, we may only consider the adjacent pairs

$$(x_1 y_1, z_1 t_1), (x_2 y_2, z_2 t_2) \in I$$

for which

$$|\{x_1, y_1, z_1, t_1\} \cap \{x_2, y_2, z_2, t_2\}| = 1.$$

Proof. Consider two adjacent elements of $I : (x_1 y_1, z_1 t_1) \sim (x_2 y_2, z_2 t_2)$. Suppose $|\{x_1, y_1, z_1, t_1, x_2, y_2, z_2, t_2\}| = 7$. The expected number of such pairs is

$$p^4 \cdot \frac{n!}{4 \cdot (n-7)!} = (1 + o(1))p^4 n^4 / 4.$$

Now consider two adjacent elements of $I$ with $|\{x_1, y_1, z_1, t_1, x_2, y_2, z_2, t_2\}| \leq 6$. There are no more than

$$n^6 = o(p^4 n^7)$$

such pairs of pairs. 

Thus we can and will only consider pairs of elements of $I$ with only one vertex in common.

We will now compute the probability that $I_{(x_1 y_1, z_1 t_1)} I_{(x_1 y_2, z_2 t_2)} = 1$. Consider a uniformly random vertex $v$. We can neglect the case when $v \in \{x_1, y_1, z_1, t_1, y_2, z_2, t_2\}$ because it happens with probability $o(1)$. Since the random graph has diameter 2 a.a.s., $I_{(x_1 y_1, z_1 t_1)} I_{(x_1 y_2, z_2 t_2)} = 1$ in the following cases.

Case 1. $d_v(x_1) = 1$. Then $v$ has to have distance 1 to all four edges. $v$ has distance 1 to $z_1 t_1$ (or $z_2 t_2$) with probability $p^3 + 2p(1-p) = p(2-p)$, and the distances from $v$ to $y_1, y_2$ don’t affect anything, so

$$\mathbb{P}(I_{(x_1 y_1, z_1 t_1)} I_{(x_1 y_2, z_2 t_2)} = 1 \mid \text{Case 1 holds}) = p^3 (2-p)^2.$$

Case 2. $d_v(x_1) = 2$. Then $v$ has distance 2 to both $x_1 y_1$ and $z_1 t_1$ with probability $(1-p)^3$ and distance 1 to both $x_1 y_1$ and $z_1 t_1$ with probability $p^2 (2-p)$. So $v$ is equidistant from the two edges with probability $(1-p)^3 + p^2 (2-p)$. Thus,

$$\mathbb{P}(I_{(x_1 y_1, z_1 t_1)} I_{(x_1 y_2, z_2 t_2)} = 1 \mid \text{Case 2 holds}) = (1-p)((1-p)^3 + p^2 (2-p))^2.$$

Hence the total probability

$$\mathbb{P}(I_{(x_1 y_1, z_1 t_1)} I_{(x_1 y_2, z_2 t_2)} = 1) = (1-p)((1-p)^3 + p^2 (2-p))^2 + p^3 (2-p)^2.$$

We will henceforth refer to this constant as $s_p$.

$$s_p := (1-p)((1-p)^3 + p^2 (2-p))^2 + p^3 (2-p)^2.$$
It follows that
\[ \Delta = (1 + o(1))p^4n^7s_p^r/4. \]
Using (2), we get
\[ \Delta = (1 + o(1))p^4n^7s_p^r/4 = (1 + o(1))p^4n^3n^r4^{-\varepsilon}s_p^r/4 \]
\[ = (1 + o(1))2p^2n^3 \left( \frac{s_p}{q} \right)^r \frac{p^2n^r}{8} = (1 + o(1))2p^2n^3 \left( \frac{s_p}{q} \right)^r \mu. \]
Notice that
\[ \left( \frac{s_p}{q} \right)^r = \left( \frac{s_p}{q} \right)^{(4-\varepsilon)\log n/\log(1/q)} = n^{(4-\varepsilon)\log\left( \frac{s_p}{q} \right)/\log(1/q)} \]
\[ = n^{(4-\varepsilon)\left( -\log(s_p) +1 \right) \log(q)} \leq n^{(4-\varepsilon)(-\log_q s_p +1)} \leq n^{\varepsilon-4} \]
(since \( q, s_p \leq 1 \)). Thus,
\[ (1 + o(1))2p^2n^3 \left( \frac{s_p}{q} \right)^r \mu \leq (1 + o(1))2p^2n^3n^{\varepsilon-4}\mu = o(\mu). \]
This concludes the proof that
\[ \Delta = o(\mu). \]

Finally, we estimate \( \delta \) and show the following.

**Lemma 3.6** (Evaluation of \( \delta \)).

\[ \delta = o(1). \]

**Proof.** Note that for fixed \( f_1, e_1 \),
\[ \mathbb{P}(A_{e_2,f_2} \mid (e_2, f_2) \text{ uniformly random}, (e_2, f_2) \sim (e_1, f_1)) \]
\[ = \mathbb{P}(A_{e,f} \mid e, f \text{ uniformly random}). \]
Thus, the maximum for \( \delta \) is achieved for \((e_1, f_1)\) with the largest possible number of adjacent edge pairs \((e_2, f_2)\). Clearly, this number is the greatest when \( e_1 \) and \( f_1 \) don’t share vertices. The expected number of adjacent edge pairs in this case is \((1 + o(1))2n^3p^2\). Since \( q^r = \mathbb{P}(A_{e,f}) \) for uniformly random edges \( e, f \) we have
\[ 2\delta = (1 + o(1))2n^3p^2q^r. \]
Using (2), we get
\[ \delta = (1 + o(1))2p^2n^{\varepsilon-1} = o(1). \]
We are now ready to prove Theorem 3.1.

**Proof of Theorem 3.1.** Substituting the results of Lemmas 3.3, 3.4, 3.6 into inequality (1), we obtain

\[
\log(\Pr(S = 0)) \leq (1 + o(1)) \left(-\mu + o(\mu) e^{o(1)}\right) \leq (1 + o(1))(-\mu + o(\mu))
\]

\[
\leq -(1 + o(1))\mu \leq -(1 + o(1))p^2 n^\varepsilon / 8 \leq -p^2 n^\varepsilon / 16
\]

for sufficiently large \( n \). Then the expected number of edge generating sets of cardinality \( r \) is no greater than

\[
\binom{n}{r} \exp(-p^2 n^\varepsilon / 16) \leq n^r \exp(-p^2 n^\varepsilon / 16)
\]

\[
= O\left(\exp[(4 - \varepsilon) \log^2(n)/\log(1/q) - p^2 n^\varepsilon / 16]\right)
\]

\[
\leq O\left(\exp[\log^2(n) - \log^3(n)p^2 / 16]\right) = o(1).
\]

This concludes the proof of Theorem 3.1, and together with Theorem 2.1, this proves the main result, Theorem 1.1.

\[\blacksquare\]

### 4. Concluding Remarks

We have shown that

\[
edim(G(n, p)) = (1 + o(1))\frac{4 \log n}{\log(1/q)},
\]

where

\[
q = 1 - 2p(1-p)^2(2-p).
\]

As demonstrated by Bollobas et al. in [1],

\[
dim(G(n, p)) = (1 + o(1))\frac{2 \log n}{\log(1/Q)},
\]

where \( Q = p^2 + (1-p)^2 \). Since \( 2/\log(1/Q) < 4/\log(1/q) \), this means that

\[
dim(G(n, p)) < \dim(G(n, p))
\]

a.a.s. for all \( p \in (0, 1) \).

While random graphs with constant edge probability don’t help in resolving the problem of finding more examples of graphs \( G \) for which \( \dim(G) < \dim(G) \) posed in [7], perhaps this problem could be addressed with random graphs of
non-constant probability \( p(n) \). Because of this it would be interesting to calculate \( \text{edim}(G(n, p(n))) \) for non-constant \( p(n) \). The analogous results for \( \text{dim}(G(n, p(n))) \) can be found in [1].

Acknowledgments

The research was conducted during the Undergraduate Mathematics Research Program at University of Minnesota Duluth, and supported by grants NSF-1358659 and NSA H98230-16-1-0026. I would like to thank Joe Gallian very much for creating a marvelous working environment and all his invaluable support throughout the program. I’m very thankful to thank Matthew Brennan for suggesting the problem; I would also like to thank Levent Alpoge for pointing out the paper [1], encouraging me to work on this problem and helpful commentary and remarks.

References

doi:10.1016/0734-189X(84)90051-3


doi:10.1016/j.disc.2018.04.010

Received 10 October 2018
Revised 9 February 2019
Accepted 9 February 2019