Triangles in random graphs

Martin Loebl, Jiří Matoušek, Ondřej Pangrác

Department of Applied Mathematics and Institute for Theoretical Computer Science, Charles University, Malostranské nám. 25, 118 00 Prague, Czech Republic. [1]

Abstract

We show the number of triangles of $G_{n,\frac{1}{2}}$ is almost uniformly distributed among residue classes modulo q, where q is a prime number bounded by $\Theta(\log n)$. This implies a consequence of a conjecture of Bollobás, Pebody and Riordan (that almost every random graph $G_{n,\frac{1}{2}}$ is uniquely determined by its Tutte polynomial): almost every pair of independently chosen random graphs $G_{n,\frac{1}{2}}$ has different Tutte polynomials.

Key words: Number of triangles, Random graph, Tutte polynomial

1 Introduction

The Tutte polynomial of a graph $G = (V, E)$ is the bivariate polynomial

$$T(G; x, y) = \sum_{A \subseteq E} (x-1)^{r(E)-r(A)} (y-1)^{|A|-r(A)},$$

where $r(A) = |V| - k(A)$ and $k(A)$ denotes the number of components of the graph (V, A).

It is clear that two isomorphic graphs have the same Tutte polynomial. On the other hand, there are examples of non-isomorphic graphs with the same Tutte

Email addresses: loebl@kam.mff.cuni.cz (Martin Loebl), matousek@kam.mff.cuni.cz (Jiří Matoušek), pangrac@kam.mff.cuni.cz (Ondřej Pangrác).

[1] Institute for Theoretical Computer Science is supported by Ministry of Education of Czech Republic as project LN00A056.

polynomial (see [3,1]). In this context, the question of how much information about the graph its Tutte polynomial contains is natural.

A random graph $G_{n,\frac{1}{2}}$, for n natural number, is a graph on an n element vertex set such that any pair of vertices forms an edge with probability $\frac{1}{2}$ independently of other pairs. Let $\mathcal{G}_{n,\frac{1}{2}}$ denote the corresponding probability space. Note that such a random graph is dense; the expectation of the number of edges is $\frac{1}{2}\binom{n}{2}$ and the expectation of the number of triangles is $\frac{1}{8}\binom{n}{3}$. The following conjecture was presented in [1].

Conjecture 1.1 *Almost every graph $G \in \mathcal{G}_{n,\frac{1}{2}}$ is such that $T(G;x,y) = T(G';x,y)$ implies $G \cong G'$.*

The main result of this paper stated below implies a weaker result.

Theorem 1.2 *There are numbers q_0 and n_0 and a function $f \in \Theta(\log n)$ such that for any $n > n_0$, any prime number q, $q_0 < q < f(n)$, and any choice of $k \in \{0, 1, \ldots, q-1\}$ the probability that a random graph $G_{n,\frac{1}{2}}$ contains k triangles modulo q is $q^{-1}(1 + o(1))$.*

Since the number of triangles of any simple graph can be easily derived from its Tutte polynomial, we can state this easy consequence of Theorem 1.2.

Corollary 1.3 *The probability that two independently chosen random graphs $G_{n,\frac{1}{2}}$ have the same Tutte polynomial is of order $O(1/\log n)$.*

Another open problem is to find or to prove existence of a family containing almost every graph such that the Tutte polynomial distinguishes any pair of graphs inside this family. Solving this problem affirmatively may give us better understanding and a hope for positive answer to Conjecture 1.1.

2 Counting 2-paths

Let G be a random graph from $\mathcal{G}_{n,\frac{1}{2}}$ with the vertex set V. Let $V_1 = \{u_1, v_1, u_2, v_2, \ldots, u_{\lfloor \frac{n}{4} \rfloor}, v_{\lfloor \frac{n}{4} \rfloor}\}$ be a set of $2\left\lfloor \frac{n}{4} \right\rfloor$ vertices partitioned into pairs (u_i, v_i), $i = 1, 2 \ldots, \left\lfloor \frac{n}{4} \right\rfloor$. Let V_2 be a subset of $V - V_1$ and denote the size of V_2 by l. For $w \in V_2$, let X_w^i be an indicator of the existence of both edges $\{u_i, w\}$ and $\{v_i, w\}$. Then $X^i = \sum_{w \in V_2} X_w^i$ counts the number of 2-paths connecting vertices u_i, v_i and having the middle vertex in V_2. Denote by $p_j^i(l)$ the probability that $X^i = j$. Since edges from distinct vertices of V_1 going to V_2 are independent, the number of such 2-paths is independent for distinct pairs of vertices of V_1 and hence we let $p_j^i(l) = p_j(l)$. Note also that $p_j(l)$ depends only on $l = |V_2|$,

not on the particular choice of V_2.

Assume q to be an arbitrary fixed integer $1 \leq q \leq l$. Denote by $s_i(l) = \sum_{j \equiv i (\mathrm{mod}\, q)} p_j(l)$ the probability that the number of connecting 2-path with the middle vertex in V_2 is equal to i modulo q.

Lemma 2.1 *Let $1 \leq q \leq |V - V_1|$. There exists a number $c = c(q) > 0$ such that $s_i(l) \geq c$ for $i = 0, 1, \ldots, q-1$ and $l \geq q-1$.*

Proof: Note that the probability of a given 2-path is $\frac{1}{4}$. Using the recursion $p_i(l) = \frac{1}{4} p_{i-1}(l-1) + \frac{3}{4} p_i(l-1)$ one can derive $s_i(l) = \frac{1}{4} s_{i-1}(l-1) + \frac{3}{4} s_i(l-1)$ (where $s_{-1}(l)$ means $s_{q-1}(l)$).

It is easy to observe that if $s_i(l_0) \geq c$ for all possible i's and some l_0 and c, then also $s_i(l) \geq c$ for all $l \geq l_0$. Note that $s_i(q-1) = p_i(q-1)$, and clearly, each $p_i(q-1) \geq \left(\frac{1}{4}\right)^{q-1}$. Hence for any $l \geq q-1$ and any $i = 1, 2, \ldots, q-1$ we have $s_i(l) \geq \left(\frac{1}{4}\right)^{q-1}$. By setting $c(q) = \left(\frac{1}{4}\right)^{q-1}$ we complete the proof. □

In the rest of the paper, we assume the pairs $(u_1, v_1), (u_2, v_2), \ldots, (u_{\lfloor \frac{n}{4} \rfloor}, v_{\lfloor \frac{n}{4} \rfloor})$ forming set V_1 are fixed and $V_2 = V - V_1$.

Lemma 2.2 *Let n be sufficiently large natural number. Then with probability at least $1 - \frac{5}{3} \cdot \frac{\log_2 n}{n^{1/3}}$ a graph $G \in \mathcal{G}_{n,\frac{1}{2}}$ satisfies, for any $q \leq q_0 = \lfloor \log_2 n^{1/3} \rfloor$, the following property: There exists a $(q-1)$-tuple $(i_1, i_2, \ldots, i_{q-1})$ such that the number of 2-paths connecting u_{i_k} and v_{i_k} is equal to k modulo q for any $k = 1, 2, \ldots, q-1$.*

Proof: By Lemma 2.1 the probability that the number of 2-paths with the middle vertex in V_2 connecting u_i and v_i is equal to k modulo q is at least $c(q)$ for any $k = 0, 1, \ldots, q-1$ (and hence at most $1-(q-1)c(q)$). Moreover, these probabilities are independent for distinct pairs u_i, v_i and u_j, v_j.

Denote the event that there is a pair u_i, v_i (among the pairs from V_1) connected by exactly k 2-paths (counted modulo q) by Y_k ($Y_k = 1$ if such a pair exists, $Y_k = 0$ otherwise). Then the probability $P[Y_1, Y_2, \ldots, Y_{q-1}]$ is equal to $1 - P\left[\exists k : \overline{Y_k}\right]$. Assume the edges inside V_1 are fixed. The probability that vertices u_i, v_i are connected by k 2-paths is $s_{i'}(|V_2|)$, where i' is such that the number of 2-paths with the middle vertex inside V_1 plus i' is equal to k modulo q. Hence $P\left[\overline{Y_k}\right] = \prod_{i=1}^{\lfloor \frac{n}{4} \rfloor} (1 - s_{i'}(|V_2|)) \leq (1 - c(q))^{\lfloor \frac{n}{4} \rfloor}$.

$$P\left[\exists k : \overline{Y_k}\right] \leq \sum_{k=1}^{q-1} P\left[\overline{Y_k}\right] \leq \sum_{k=1}^{q-1} (1 - c(q))^{\lfloor \frac{n}{4} \rfloor} = (q-1)(1 - c(q))^{\lfloor \frac{n}{4} \rfloor}.$$

Let $q_0 = \lfloor \log_2 n^{1/3} \rfloor$. Since $q \leq q_0$ and $c(q) \geq c(q_0)$ we can bound

$$P\left[\exists k : \overline{Y_k}\right] \leq q_0 \left(1 - c(q_0)\right)^{\lfloor \frac{n}{4} \rfloor} = q_0 \frac{\left(1 - c(q_0)^2\right)^{\lfloor \frac{n}{4} \rfloor}}{\left(1 + c(q_0)\right)^{\lfloor \frac{n}{4} \rfloor}} \leq$$
$$\leq \frac{q_0}{\left(1 + c(q_0)\right)^{\lfloor \frac{n}{4} \rfloor}} \leq \frac{q_0}{1 + \left\lfloor \frac{n}{4} \right\rfloor c(q_0)} .$$

From Lemma 2.1 we have $c(q_0) = \left(\frac{1}{4}\right)^{q_0 - 1} \geq \left(\frac{1}{4}\right)^{\log n^{1/3}} = 2^{-2 \log n^{1/3}} = n^{-2/3}$. Hence

$$P\left[\exists k : \overline{Y_k}\right] \leq \frac{\log n^{1/3}}{1 + \left\lfloor \frac{n}{4} \right\rfloor n^{-2/3}} \leq \ frac \log n^{1/3} \frac{n}{5} n^{-2/3} = \frac{5}{3} \frac{\log n}{n^{1/3}} .$$

□

3 Proof of the Theorem

Using the results of [2] we can state the following useful proposition.

Proposition 3.1 *Let q be a sufficiently large prime number; then the number of subsets $S \subseteq \{1, 2, \ldots, q-1\}$ such that the sum $\sum_{i \in S} i \equiv k \mod q$ is $2^{q-1} q^{-1} (1 + o(1))$ for any $k = 0, 1, \ldots, q-1$.*

Combining the result of Lemma 2.2 with the previous proposition we can prove Theorem 1.2.

Proof of Theorem 1.2: Let us consider a graph $G \in \mathcal{G}_{n,\frac{1}{2}}$ with a fixed set V_1 partitioned into $\left\lfloor \frac{n}{4} \right\rfloor$ disjoint pairs of vertices as above. By Lemma 2.2 we can find with probability at least $1 - \frac{5}{3} \frac{\log n}{n^{1/3}}$ a $(q-1)$-tuple $(i_1, i_2, \ldots, i_{q-1})$ such that the number of connecting 2-paths between vertices u_{i_k} and v_{i_k} is equal to k modulo q for all $k = 1, 2, \ldots, q-1$. Choose such a $(q-1)$-tuple lexicographically minimal. Let us define graphs G_S for any $S \subseteq \{1, 2, \ldots, q-1\}$ obtained from G by adding all edges $\{u_{i_k}, v_{i_k}\}$ with $k \in S$ (if they are not in G) and by deleting all edges $\{u_{i_k}, v_{i_k}\}$ with $k \notin S$ (if they are in G). The probability of all G_S in the probability space $\mathcal{G}_{n,\frac{1}{2}}$ is the same and the $(q-1)$-tuple $(i_1, i_2, \ldots, i_{q-1})$ is lexicographically minimal with required property for any graph G_S.

Moreover, graphs G_1 and G_2 of $\mathcal{G}_{n,\frac{1}{2}}$ have assigned the same (lexicographically minimal) $(q-1)$-tuples $(i_1, i_2, \ldots, i_{q-1})$ and their edge sets differ only on the edges $\{u_{i_k}, v_{i_k}\}$ $(k = 1, 2, \ldots, q-1)$ if and only if they are G_{S_1} and G_{S_2} for some G and $S_1, S_2 \subseteq \{1, 2, \ldots, q-1\}$.

From now on, all computation will be done modulo q. Let t_0 be the number of triangles of $G_\emptyset$. Since the pairs u_{i_k}, v_{i_k} are disjoint, edges $\{u_{i_k}, v_{i_k}\}$ do not affect the number of 2-paths between these pairs. Then the number of triangles of G_S is $t_0 + \sum_{k \in S} k$. By Proposition 3.1, if S is chosen uniformly at random then the probability that the number of triangles of G_S is equal to k modulo q is $q^{-1}(1 + o(1))$.

Up to a $\left(\frac{5}{3}\frac{\log n}{n^{1/3}}\right)$-fraction, we have partitioned the graphs of $\mathcal{G}_{n,\frac{1}{2}}$ into classes of equal size such that in each class the number of triangles computed modulo q is almost uniformly distributed. Since $\frac{\log n}{n^{1/3}}$ tends to 0 for n going to infinity, we can conclude that the probability that $G_{n,\frac{1}{2}}$ contains exactly t triangles modulo q is $q^{-1}(1 + o(1))$ for any $t = 0, 1, \ldots, q-1$.

□

4 Acknowledgements

The authors would like to thank to D. Welsh for helpful discussions and to M. Klazar for pointing out the paper of Erdős and Heilbronn. The authors would also like to thank to referees for suggesting several improvements.

References

[1] B. Bollobás, L. Pebody, O. Riordan. Contraction-deletion invariants for graphs. *J. Combin. Theory Ser. B* **80** (2000), 320–345

[2] P. Erdős and H. Heilbronn. On the addition of residue classes mod p. *Acta Arithmetica* **9** (1964), 149–159

[3] H. Whitney. 2-isomorphic graphs. *Amer. J. Math.* **55** (1933), 245–254