Journal of Graph Algorithms and Applications
http://jgaa.info/

Corrections to the JGAA Paper "A Binomial Distribution Model for the Traveling Salesman Problem Based on Frequency Quadrilaterals" by Y. Wang and J. B. Remmel

Y. Wang[1] *J. B. Remmel*[2]

[1]School of Renewable Energy,
North China Electric Power University, Chang Ping, Beijing 102206. China
[2]Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112. USA

Abstract

In a recently published paper "Y. Wang and J. B. Remmel, A Binomial Distribution Model for the Traveling Salesman Problem Based on Frequency Quadrilaterals, Journal of Graph Algorithms and Applications, vol. 20, no. 2, pp. 411-434 (2016)", we assume the probability $p_{\{3,5\}} = \frac{2}{3}$ when we compute the standard deviation $\sigma(\epsilon_e)$ according to the three frequency sets $\{5, 3\}, \{5, 1\}$ and $\{3, 1\}$. This assumption should change because the probability $p_{\{3,5\}}$ is computed according to the three frequency sets $\{5, 3\}, \{5, 1\}$ and $\{3, 1\}$ rather than the six frequency quadrilaterals. In addition, we derive the 32 frequency pentalaterals according to the regular pentagon. The number of frequency pentalaterals is at least 60 for general graphs K_5. These two errors are corrected in this short notice.

This erratum has been published on August 2017.

Research supported by Grant NSFC-51205129, Fundamental Research Funds for the Central Universities-2015ZD10

E-mail addresses: yongwang@ncepu.edu.cn (Y. Wang) jremmel@ucsd.edu (J. B. Remmel)

This short notice is to report corrections and comments on our published paper "Y. Wang and J. B. Remmel, A Binomial Distribution Model for the Traveling Salesman Problem Based on Frequency Quadrilaterals, Journal of Graph Algorithms and Applications, vol. 20, no. 2, pp. 411-434 (2016)".

The following are corrected versions of some results. The changed contents are stated.

1. Line 7 in the last paragraph of page 423 of the original paper.

The sentence " The ϵ_e corresponding to an edge e is equal to $2(1+\delta_e)p_{\{3,5\}}$, where $p_{\{3,5\}} = \frac{2}{3}$ according to our assumption about the distribution of the frequency of e in the six frequency quadrilaterals in our binomial distribution model." is changed into " The ϵ_e corresponding to an edge e is equal to $2(1+\delta_e)p_{\{3,5\}}$ where $p_{\{3,5\}}$ is the probability that e has frequency 3 or 5 according to some of the the six frequency quadrilaterals in our binomial distribution model."

2. The results associates with the $p_{\{3,5\}}$ in page 424 of the original paper are revised.

For every edge, its frequency is 5, 3 or 1 in a frequency quadrilateral. Therefore, we draw the pairwise frequency from $\{5,\ 3\ ,1\}$ to form three frequency sets $\{5,\ 3\}$, $\{5,\ 1\}$ and $\{3,\ 1\}$. In the three frequency sets, the corresponding δ_e is either 0.5, 1.0, 0 and each occurs with probability $\frac{1}{3}$. Meanwhile, the corresponding $p_{\{3,5\}}$ is equal to $p_{\{3,5\}} = 1$, $p_{\{3,5\}} = p_5 = \frac{1}{2}$ and $p_{\{3,5\}} = p_3 = \frac{1}{2}$, respectively, considering each of the two frequency quadrilaterals. Of course, the expected value $\mu(\delta_e) = 0.5$. For every edge e which corresponds to $\delta_e = 0.5\ (1.0,\ 0)$ and $p_{\{3,5\}} = 1.0,\ 0.5,\ 0.5$, the corresponding ϵ_e is 3 (2, 1) and the expected value of ϵ_e, $\mu(\epsilon_e) = 2$. It follows that $\sigma^2(\epsilon_e) = \frac{2}{3} \approx 0.6667$ and $\sigma(\epsilon_e) \approx 0.8165$. One can compute that in the normal distribution $\epsilon_e \sim \mathcal{N}(2, \frac{2}{3})$. $P(\epsilon_e \geq 4) \leq 0.007153$. However, we know that $P(\epsilon_e > 4) = 0$ for every edge e.

Note that there are in total $6\binom{n}{4}$ ϵ_es because a K_n has $\binom{n}{4}$ quadrilaterals and each quadrilateral contains 6 edges. Let ϵ_e denote the ϵ associated with edge e. If we draw N ϵ_es, i.e., $\{\epsilon_{e^1},\ \epsilon_{e^2},\ \cdots,\ \epsilon_{e^N}\}$ where ϵ_{e^k} means the k^{th} ϵ_e, at random, then we let $\epsilon = \frac{1}{N}\sum_{k=1}^{N}(\epsilon_{e^k})$ denote the associated mean value and $\sigma^2(\epsilon) = \sigma^2(\frac{1}{N}\sum_{k=1}^{N}(\epsilon_{e^k}))$ denote the associated variance. Obviously, $\sqrt{N}(\epsilon - \mu(\epsilon))$ conforms to a normal distribution based on the central limit theorem. Here $\sqrt{N}(\epsilon-\mu(\epsilon)) \sim \mathcal{N}(0, \frac{2}{3})$ or $\sqrt{N}\epsilon \sim \mathcal{N}(2\sqrt{N}, \frac{2}{3})$. The maximum and minimum ϵ are 4 and 0, respectively. As N becomes big, the $\sqrt{N}\epsilon$ also increases. However, the variance of all these ϵs remains unchanged. This means that the probability that ϵ is close to 4 becomes smaller as N becomes large.

The ϵ computed according to formula (5) for every edge e is just the mean value of the $\binom{n-2}{2}$ ϵ_es. The ϵ of every edge will conform to the normal distribution $\sqrt{N}(\epsilon-\mu(\epsilon)) \sim \mathcal{N}(0, \frac{2}{3})$ where $N = \binom{n-2}{2}$. This means that the probability that the ϵs deviate from their expected value $\mu(\epsilon) = 2$ and approach 4 tends to zero as $n \to \infty$. Thus, the number of ϵs close to 4 is very small. In the next section, we will see the ϵs of the OHC edges increase with the scale of TSP n until they approach the maximum value 4.

A linear transformation does not change the probability properties of random variables. Therefore, we can use the ϵs computed according to formula (5) to analyze their distribution for TSP. For the $\binom{n}{2}$ ϵs, the expected value $\mu(\epsilon)$ and variance $\sigma^2(\epsilon)$ are computed as follows. We assume $M = \binom{n}{2}$, $N = \binom{n-2}{2}$ and $\{\epsilon_1, \ \epsilon_2, \ \cdots, \ \epsilon_M\}$ are the ϵs of the M edges. For the j^{th} edge, $\epsilon_j = \frac{1}{N}\sum_{i=1}^{N}(\epsilon_{ij})$, where every $\epsilon_{ij} = \epsilon_e \in \{1,2,3\}$. In addition, we suppose all ϵ_{ij}s are independently and uniformly distributed. The expected value of ϵ_j is $\mu(\epsilon_j) = \frac{1}{N}\sum_{i=1}^{N}\mu(\epsilon_{ij}) = 2$. The variance of ϵ_j is $\sigma^2(\epsilon_j) = \frac{1}{N^2}\sum_{i=1}^{N}\sigma^2(\epsilon_{ij}) = \frac{2}{3N}$. This holds in our binomial distribution model because we are assuming that the frequency of edge e_j being 1, 3, or 5 has the equal probability $\frac{1}{3}$. In real graphs, it is often the case that short edges have a high probability of having frequency 5 and 3 in their frequency quadrilaterals. On the other hand, it is often the case that for long edges, there is a small probability that the edge will have frequency 5 or 3 in their frequency quadrilaterals. Based on the $N \times M$ matrix of ϵ_{ij}s, we can derive the expected value and variance of them. The expected value of the ϵ_{ij}s is $\mu(\epsilon_{ij}) = \frac{1}{MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\mu(\epsilon_{ij}) = 2$. Meanwhile, the variance of the ϵ_{ij}s is computed as $\sigma^2(\epsilon_{ij}) = \frac{1}{MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\sigma^2(\epsilon_{ij}) = \frac{2}{3}$.

3. Line 8 below the formula (11) in page 425 of the original paper.

The sentence "If we take a threshold at 0.0025 as a small probability (which is reasonable considering the 3σ rule for the normal distribution), then $t_{\max} = 2.819$ and $\sigma(\epsilon) \approx 0.7094$ which is bigger than the theoretical value 0.5443 (or $(\frac{2}{3})^{\frac{3}{2}}$) of the ideal case." is changed into " If we take a threshold at 0.0025 as a small probability (which is reasonable considering the 3σ rule for the normal distribution), then $t_{\max} = 2.819$ and $\sigma(\epsilon) \approx 0.7094$ which is approximately equal to the theoretical value 0.8165 (or $(\frac{2}{3})^{\frac{1}{2}}$) based on the frequency quadrilaterals."

4. Line 3 in the second paragraph of the conclusion in page 431 of the original paper.

The number 32 is changed into 60.