æ

Edge Disjoint Polyp Packing

GYULA Y. KATONA*

Mathematical Institute of the
Hungarian Academy of Sciences

Abstract. A graph is called *p–polyp* if it consists of p simple paths of the same length and one endvertex of all these paths is a common vertex. The Polyp Packing problem is a generalization of the well known Bin Packing problem: How to pack a set of paths with different lengths to a set of polyps edge disjointly? It is proved that the Polyp Packing problem is NP-complete and that a modification of the First Fit algorithm gives a reasonable approximation.

Keywords: Bin Packing problem, First Fit algorithm, packing of paths
Subject classification: 68Q25, 68R10

1. Introduction

Both of the following problems may be considered as a generalization of the Bin Packing problem [3]: Given a simple path P, a set of simple paths possibly with many different lengths and a positive integer K. Is it possible to embed all the paths edge/vertex disjointly into K copies of P? (In the edge disjoint case the embeded path may contain a vertex more than once, but each edge only once.)

G. O. H. Katona generalized this question: Given a graph G, a set of simple paths possibly with many different lengths and a positive integer K. Is it possible to embed all the paths edge/vertex disjointly into K copies of G?

The vertex disjoint problem is equivalent to the Bin Packing problem if G has a Hamiltonian path. The edge disjoint problem is equivalent to the Bin packing problem if G has an Eulerian path.

In this paper this question is investigated in a special case when the given graph is a polyp. A graph is called *p–polyp* if it consists of p simple paths of the same length and one endvertex of all these paths is a common vertex — the center of the polyp. The paths attached to the center are called *arms*. These problems are called *edge/vertex disjoint Polyp Packing problems*. Obviously if the given graph is a p–polyp then the problem is called p–Polyp Packing. The main difference between the edge and vertex disjoint case is that several packed paths may contain the center of the polyp in the first case but only one in the second case. Note that there is no difference if $p = 3$. (see FIG 1.)

First it is proved that both Polyp Packing problems are also NP–complete like the original Bin Packing problem. Then we consider a variant of the First Fit algorithm and determine how many more polyps are needed to pack the paths with the First Fit algorithm than with the optimal packing in the edge disjoint case. Some of the proofs are similar to the proofs in Johnson at al [3] which proves similar theorems about the original Bin Packing problem. In a forthcoming paper similar results will be presented in the vertex disjoint case which looks harder to prove.

2. NP–completeness

THEOREM 1. *Both the edge and vertex disjoint p–Polyp Packing problems are NP–complete for any given $p \geq 3$.*

PROOF: First of all it is trivial that both problems are in NP, since one can check the validity of a given packing in polynomial time.

*This work was partially supported by the Hungarian National Foundation for Scientific Research, OTKA Grant Numbers T 4271 and T 014302

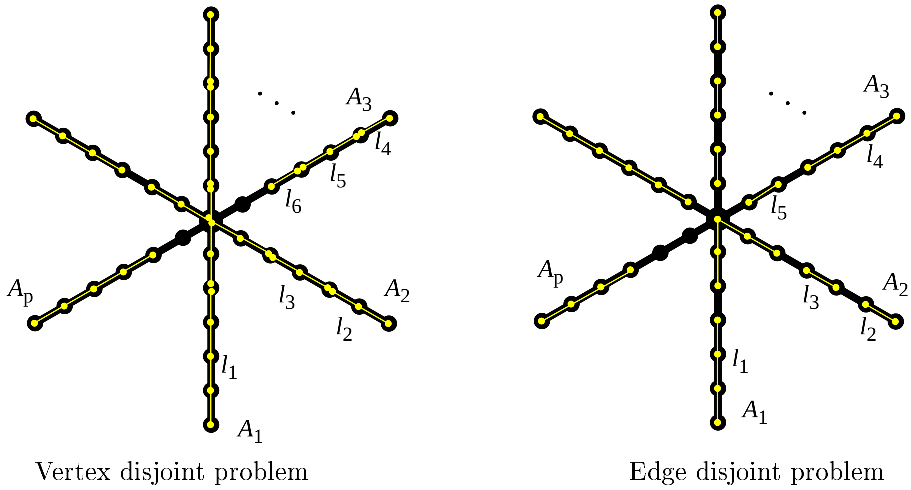

FIG. 1.

Edge disjoint case. We transform Bin Packing to p–Polyp Packing. If in the Bin Packing problem we consider only integer size items and a "large" integer size bin instead of rational size items, the problem remains NP–complete, trivially. There is no real difference if odd integer size bins are not allowed.

So we really transform this restricted problem to the Polyp Packing problem. There are two cases depending on the parity of p. In the first case p is odd. Let the given set of integer sizes be $a_1, a_2, \dots, a_t$, the size of the bin is $2s$ and the integer in the question is K. We create a set of paths with lengths $a_1, a_2, \dots, a_t$ then add $\frac{p-1}{2}K$ copies of a $4s$ long path. So there are $t + \frac{p-1}{2}K$ paths. It is possible to pack the t items into K bins of size $2s$ if and only if one can pack these paths into K copies of a p–polyp with $2s$ long arms.

The second case is the even p. A set of paths is created as before, paths with length $a_1, a_2, \dots, a_t$ and additionally $\frac{p-2}{2}K$ paths of length $2s$ but the arms of the polyp now have length s.

If it is possible to pack the t items into K bins of size $2s$ then it is easy to see that one can pack the above set of paths into K p–polyps with armlength s. $\frac{p-2}{2}$ paths of size $2s$ are put into the first polyp and the paths corresponding the content of the first bin are put into the remaining space in the first polyp, etc.

On the other hand let us show that if one can pack the items into K p–polyps then there is a packing into K bins of size $2s$. Delete the arms containing one of the paths of the additional paths with size $2s$, clearly $2K$ arms remains. These arms contain all the paths which corresponds to the sizes $a_1, a_2, \dots, a_t$. Now a matching is made between these arms. If there is path which has a part in two different arms, these arms are matched. The remaining arms of a polyp can be matched arbitrarily. It is clear that exactly K pairs are obtained in this way. So put the items corresponding to the contents of one pair to one bin.

Vertex disjoint case. Let the given set of integer sizes be $a_1, a_2, \dots, a_t$, the size of the bin s and the integer in question K. A set of paths with lengths $a_1 - 1, a_2 - 1, \dots, a_t - 1$ is created, K copies of a path with length s and $(p-2)K$ copies of a path with length $s-1$ is added to this set.

First it is proved that if the items can be packed into K bins of size s then there is a packing of the above paths into K p–polyps with arm size s. Put one path of size s to the first polyp so that it covers one arm, thus the endvertex of it will cover the center of the polyp. Then we put $p-2$ paths of size $s-1$ to the other arms, the only way for this is when each of them covers all vertices of an other arm. Finally the paths corresponding to the contents of the first bin will fit into the remaining arm of the first polyp vertex disjointly. Then this procedure is repeated with each polyp.

To complete the proof it must be shown that if there is a packing of the above paths into K p–polyps then we can pack the items to K bins of size s. It is clear that if there is a packing of the paths to the polyps then there is exactly one path of size s in every polyp. This path must cover the center. It is also possible that it covers vertices in two different arms, but the the paths in these two arms contain at most s vertices. Therefore we may assume that it covers one arm completely and the center, and the other paths are in a separate arm. A path of size $s-1$ can be only in an arm alone. Thus the paths corresponding to the items are packed into K different arms. So the content of one arm goes simply to one bin and the proof is complete. ■

3. Worst case of First Fit

Let us define the algorithm which will be investigated.

Let the p–polyps be indexed as $P_1, P_2, \ldots$, initially each empty. The arms of P_i are also indexed $A_{i,1}, \ldots, A_{i,p}$. The paths $l_1, l_2, \ldots, l_t$ will be placed in that order. To place l_i, we try to put it into $A_{1,1}$ so that it completely fits into it. If it is not possible we try to put it into $A_{1,2}$ and so on. If we find an arm which is not "full" we put l_i to this arm as far as possible from the center. If all the arms of P_1 are "full" we try to put l_i to $A_{1,1}$ and $A_{1,2}$ so that l_i will contain the center of P_1 and $A_{1,1}$ is filled completely. Then we try to put it into $A_{1,1}$ and $A_{1,3}$ and so on, we continue using the lexicographic order of the pair of arms. If l_i does not fit to P_1, we move on to P_2, etc.

This is called the First Fit algorithm. The number of polyps needed to pack a set L of paths is denoted by $FF_p(L)$, while the number of polyps in an optimal packing is $OPT_p(L)$. The set of paths may be viewed in the First Fit algorithm like a list since it gives us the order of the paths in the First Fit algorithm. So L really denotes this list while $|L|$ denotes the number of paths in the list. One more notation:

$$R_p = \lim_{|L|\to\infty} \frac{FF_P(L)}{OPT_P(L)}.$$

To make computations easier, from now on, the arm length of the polyp will be the "unity". This means that every length will be divided by s, the length of the polyps arm. Thus the arm of the polyp will have size one and the sizes of the paths will be rational numbers between 0 and 2. For example a path of size $\frac{1}{2} + \varepsilon$ originally had length $s/2 + 1$, supposing that s is large.

4. Even number of arms

Theorem 2. *If $p \geq 4$ is even then for every list L*

$$FF(L) \leq OPT(L)\frac{17}{10} + 4.$$

Proof: The proof was inspired by the work of Johnson at al [3]. They defined a weight function on the items. We denote their function now by $W'(x)$. We use a similar weight function $W(x)$ on the paths. In fact

$$W(x) = \frac{2}{p} W'\left(\frac{x}{2}\right).$$

Explicitly,

$$W(x) = \begin{cases} \frac{2}{p}\frac{6}{10}x, & \text{if } 0 < x \leq \frac{1}{3} \\ \frac{2}{p}\frac{9}{10}x - \frac{2}{p}\frac{1}{10}, & \text{if } \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{2}{p}\frac{6}{10}x + \frac{2}{p}\frac{1}{10}, & \text{if } \frac{2}{3} \leq x \leq 1 \\ \frac{2}{p}, & \text{if } 1 < x \leq 2. \end{cases}$$

This means that

$$W\left(\frac{1}{3}\right) = \frac{2}{p}\frac{1}{2}, \quad W\left(\frac{2}{3}\right) = \frac{2}{p}\frac{1}{5}, \quad W\left(1\right) = \frac{2}{p}\frac{7}{10}.$$

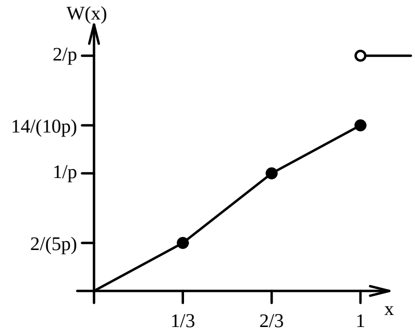

Fig. 2.

First we are going to prove a few lemmas. These are also similar to the ones in [3] but sometimes more complex.

LEMMA 1. *Let a polyp be filled with paths $l_1, l_2, \ldots, l_t$. Then*

$$\sum_{i=1}^{t} W(l_i) \leq \frac{17}{10}.$$

PROOF: If a polyp is filled with the above paths then some of the paths may cover some parts of two arms of the polyp. If this is the case then these two arms are called *paired.* We define a complete matching on the arms. If there are some paired arms they will be pairs in the matching too. The number of unpaired arms must be even since the number of arms is even. An arbitrary matching can be made between the leftover arms. In this way, a pair of arms filled with some paths is exactly the same as a bin filled with some items. But the size of the two arms is 2 here so if we want the bin to be size 1 then we must have items of size half of the corresponding path.

It is proved in [3] that if a bin of size 1 is filled with items $b_1, b_2, \ldots, b_m$ then

$$\sum_{i=1}^{m} W'(b_i) \leq \frac{17}{10}.$$

Let $h_1, \ldots, h_m$ denote the lengths of paths put in a specific pair of arms. Here $b_i = \frac{h_i}{2}$, so in one pair of arms

$$\sum_{i=1}^{m} W(h_i) = \sum_{i=1}^{m} \frac{2}{p} W'(b_i) \leq \frac{2}{p}\frac{17}{10}.$$

There are exactly $\frac{p}{2}$ pairs of arms so if we sum the weights in all pairs we obtain

$$\sum_{i=1}^{t} W(l_i) \leq \frac{17}{10}. \qquad \blacksquare$$

Let us define the *coarseness* of a polyp. The coarseness of the polyp P_d is the largest α_d such that after completing the First Fit algorithm an additional path of size α_d would fit to some polyp with smaller index than d. The coarseness of P_1 is 0. Therefore if the First Fit algorithm puts a path of size l to polyp P_d then $l > \alpha_d$.

LEMMA 2. *Let a polyp P_d of coarseness $\alpha_d \leq 1$ be filled with paths $l_1 \geq l_2 \geq \ldots \geq l_t$ with the completed First Fit algorithm. If $\sum_{i=1}^{t} W(l_i) < 1$, then $\alpha_{d+1} > \alpha_d$.*

PROOF: Unfortunately it is not so easy to use the lemmas of [3] to prove the present lemma like in the previous case. Here we have to go through all the cases, again. Start with an observation which is used in each case.

There must be a pair of arms where the sum of the weights is less then $\frac{2}{p}$ otherwise the overall weight sum should be over 1 since there are $\frac{p}{2}$ pairs. Let A denote this pair of arms and $b_1, b_2, \ldots, b_m$ the paths in it. If there is a path of size over 1 in A then the weight sum in this pair would be larger than $\frac{2}{p}$. So all paths in A are at most 1. We will show in each case that a path of size larger than α_d fits into this pair of arms.

Case 1: $\alpha_d \leq \frac{1}{3}$.

For the proof it is sufficient to use the following lower estimate of $W(x)$ in $(0, 1]$, which may be seen in FIG. 3:

$$W(x) \geq \frac{2}{p}\frac{6}{10}x.$$

In this way

$$\sum_{A} b_i \frac{2}{p}\frac{6}{10} \leq \sum_{A} W(b_i) < \frac{2}{p}$$

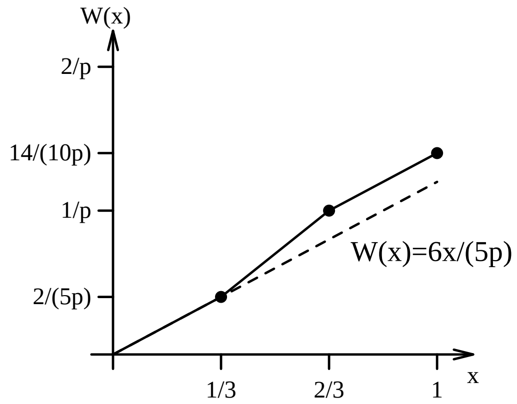

FIG. 3.

is obtained which implies

$$\sum_A b_i < \frac{10}{6} = \frac{5}{3}.$$

This means that the empty space in this pair of arms is larger than $\frac{1}{3}$. If the paths are packed with the First Fit algorithm then a new path is always placed such that one of its endvertex is the endvertex of some earlier placed path except if it is placed into an empty arm. Anyway it is clear that there may not be two separated empty parts in a pair of arms. Thus the previous inequality shows that a path of size $> \frac{1}{3}$ will fit to this polyp. Since $\alpha_d \leq \frac{1}{3}$ this gives that $\alpha_{d+1} > \alpha_d$.

Case 2: $\frac{1}{3} < \alpha_d \leq \frac{2}{3}$.

By the definition of coarseness it is known that $b_i > \alpha_d > \frac{1}{3}$ holds for all b_i. Suppose there are q paths of size in $(\frac{1}{3}, \frac{2}{3}]$ and r paths of size in $(\frac{2}{3}, 1]$. The sets of these paths is denoted by Q and R, respectively. The following inequalities are clearly true:

$$\sum_A W(b_i) = \frac{2}{p}\frac{9}{10}\left(\sum_Q b_i\right) - q\frac{2}{p}\frac{1}{10} + \frac{2}{p}\frac{6}{10}\left(\sum_R b_i\right) + r\frac{2}{p}\frac{1}{10} < \frac{2}{p},$$

$$\frac{9}{10}\left(\sum_Q b_i\right) + \frac{6}{10}\left(\sum_R b_i\right) + \frac{r-q}{10} < 1,$$

$$\frac{3}{10}\left(\sum_Q b_i\right) + \frac{6}{10}\left(\sum_A b_i\right) + \frac{r-q}{10} < 1,$$

$$\sum_A b_i < \frac{10}{6} + \frac{q-r}{6} - \frac{1}{2}\sum_Q b_i.$$

So the empty space in A is more than

$$S = \frac{1}{3} + \frac{r-q}{6} + \frac{1}{2}\sum_Q b_i.$$

It will be proved that $S \geq \alpha_d$.

Our first observation is that $b_i > \alpha_d$ implies

$$S > \frac{1}{3} + \frac{r-q}{6} + \frac{1}{2}q\alpha_d.$$

In the case $q \geq 2$ one can verify that the trivial inequality

$$q \geq 2 \geq \frac{\alpha_d - \frac{1}{3} - \frac{r}{6}}{\frac{1}{2}\left(\alpha_d - \frac{1}{3}\right)}$$

implies $S > \alpha_d$.

It is easy to see that $0 \leq r \leq 1$ otherwise the weight sum in A would be at least $\frac{2}{p}$ since $W\left(\frac{2}{3}\right) = \frac{2}{p}\frac{1}{2}$. So there are only a few cases left.

If $q = 0$ and $r = 0$ or $r = 1$ then a paths of size 1 fits clearly into A. If $q = 1$ and $r = 0$ then, again, a path of size 1 fits into A. If $q = 1$ and $r = 1$ then

$$S \geq \frac{1}{3} + \frac{1}{2}\alpha_d \geq \alpha_d$$

since $\frac{1}{3} \geq \frac{1}{2}\alpha_d$.

Case 3: $\frac{2}{3} < \alpha_d$.

This is the easiest part. If there is at most one path in A it must be at most 1, therefore a path of size 1 fits into A. On the other hand if there are two paths then the weight sum is at least $\frac{2}{p}$ since both paths must be over $\frac{2}{3}$ by the definition of α_d and $W\left(\frac{2}{3}\right) = \frac{2}{p}\frac{1}{2}$. ■

LEMMA 3. *Let a polyp P_d of coarseness $\alpha_d \leq 1$ be filled with paths $b_1 \geq b_2 \geq \ldots \geq b_t$ in the completed First Fit algorithm. If $\sum_{i=1}^{t} W(b_i) = 1 - \beta_d \quad (0 < \beta_d < 1)$ then $\alpha_{d+1} \geq \min(\alpha_d + \frac{5}{9}\beta_d, 1)$.*

PROOF: Suppose that $\alpha_{d+1} < \alpha_d + \frac{5}{9}\beta_d$. Then no path of length $\alpha_d + \frac{5}{9}\beta_d$ can fit to P_d when the list of paths L was packed to the polyps with the First Fit algorithm. A modified list L' will be created in the following way. Some of the paths in P_d will be enlarged.

It is possible that some of the paths cover vertices from two different arms. These two arms will be called paired, again. In a pair of arms the empty part is "together" because the First Fit algorithm was used to pack the paths. Let e denote the length of the empty part. This empty part does not contain the center, so it is contained in one arm completely. This arm contains a part of the path which covers the center and it must contain some paths which are as far from the center as possible. Let b_e be the closest path to the center among the latter paths. This path will be enlarged, let $b'_e = b_e + \min(\frac{5}{9}\beta_d, e)$. In this way the size of this empty part decreases by $\frac{5}{9}\beta_d$ or it becomes 0.

There may be some unpaired arms also. The empty part in these arms are attached to the center and all paths in the arms are as far from the center as possible. (There must be at least one path in each arm.) Let $e_1 \geq e_2 \geq \ldots \geq e_k$ be the sizes of the empty parts in the unpaired arms, b_i the closest path to the center in arm A_i. By our assumption $\alpha_d + \frac{5}{9}\beta_d > e_1$. If $e_1 < \alpha_d + \frac{5}{18}\beta_d$ then let $b'_i = b_i + \min(\frac{5}{18}\beta_d, e_i)$ for all i. If $e_1 \geq \alpha_d + \frac{5}{18}\beta_d$ then $\frac{5}{18}\beta_d > e_2 \geq e_3 \geq \ldots$ holds since $e_1 + e_2 < \alpha_d + \frac{5}{9}\beta_d$. Let $b'_1 = b_1 + \min(\frac{5}{9}\beta_d - e_2, e_1)$ and $b'_i = b_i + e_i$ for $i = 2, 3, \ldots, k$.

The modified list $L' = b'_1, b'_2, \ldots$ is obtained by replacing the above paths by the enlarged paths. If L' is packed with the First Fit algorithm then every path will get the same place as the corresponding path in L, because no path fits to an earlier place (therefore α_d does not change), but every path fits to its "old" place. On the other hand it is easy to see that after packing L' there will be not enough empty space to insert an additional path of size $> \alpha_d$.

The slope of W in the interval $(0, 1]$ is at most $\frac{2}{p}\frac{9}{10}$ and the sum of the enlargements is $\leq \frac{p}{2}\frac{5}{9}\beta_d$. Hence

$$\sum W(b'_i) \leq \sum W(b_i) + \left(\frac{2}{p}\frac{9}{10}\right)\frac{p}{2}\frac{5}{9}\beta_d = 1 - \frac{\beta_d}{2} < 1$$

is obtained. Therefore L' would be a counterexample to Lemma 2 which is a contradiction.

The proof is similar if $\alpha_d + \frac{5}{9}\beta_d > 1$. ■

Now we are prepared to complete the proof of Theorem 2.

PROOF OF THEOREM 2: Let $Q_1, \ldots, Q_f$ denote the polyps for which $\sum_{i=1}^{t} W(b_j) = 1 - \beta_i$ holds with $\beta_i > 0$ in the First Fit algorithm in the same order as they appear in the list of polyps. The other polyps have weight sum at least 1. Let γ_i be the coarseness of Q_i. Suppose $1 \leq \gamma_{f-1} \leq \gamma_f$. Consequently all paths in Q_{f-1} are larger than 1, therefore all paths have weight $\frac{2}{p}$. Hence the number of paths in Q_{f-1} is $\leq \frac{p}{2} - 1$ since the weight sum is less than 1. This means that there must be a completely empty pair of arms. Therefore Q_{f-1} must be the last nonempty polyp, since any further path would fit to Q_{f-1}. This contradicts our assumption. Hence $\gamma_1 \leq \ldots \leq \gamma_{f-1} < 1$.

Lemma 3 and the definition of coarseness imply

$$\gamma_{i+1} \geq \gamma_i + \frac{5}{9}\beta_i \quad \text{for} \quad 1 \leq i \leq f-2.$$

As $\beta_{f-1}, \beta_f \leq 1$ thus

$$\sum_{i=1}^{f} \beta_i \leq \frac{9}{5}\sum_{i=1}^{f-2}(\gamma_{i+1}-\gamma_i) + 2 = \frac{9}{5}(\gamma_{f-1}-\gamma_1) + 2 < 4$$

holds.

Summing the weights for all paths $l_1, \ldots, l_m$ in L

$$FF(L) - 4 \leq \sum_{i=1}^{FF(L)-f} 1 + \sum_{i=1}^{f} 1 - \beta_i \leq \sum_{j=1}^{m} W(l_j)$$

is obtained. Finally, applying Lemma 1 we conclude

$$FF(L) \leq \sum_{i=1}^{m} W(l_i) + 4 \leq OPT(L)\frac{17}{10} + 4,$$

completing the proof. ■

THEOREM 3. *If the arm size of the polyp is suitably large there exists a list L for any large enough M such that $OPT(L) \geq M$ and*

$$FF(L) \geq 1.6904 \cdot OPT(L).$$

PROOF: Let N be a positive integer divisible by 42 and let ε a be a suitably small rational number. The list of paths denoted by L consists of the following regions of paths in the order of appearance

a) $N\frac{p}{2}$ paths of size $\frac{1}{21} - 3\varepsilon$,
b) $N\frac{p}{2}$ paths of size $\frac{2}{7} + \varepsilon$,
c) $N\frac{p}{2}$ paths of size $\frac{2}{3} + \varepsilon$,
d) $N\frac{p}{2}$ paths of size $1 + \varepsilon$.

If these paths are packed with the First Fit algorithm then $\frac{p}{2}$ pairs of arms of the first $\frac{N}{42}$ polyps will contain 42 paths of region a). The coarseness of the next polyp will be $42 \cdot 3\varepsilon$ which is chosen small enough to exclude that one of the later coming paths could fit in. Namely $42 \cdot 3\varepsilon < \frac{2}{7} + \varepsilon$ must be true. In this way all the paths in region b) will go to later polyps. Exactly 3 will fit to the first arm, then 3 to the second etc. and after all arms are filled there is not enough place in the center for another one. This makes $3p$ in one polyp. Thus region b) will fill up $\frac{N}{6}$ polyps. The coarseness of the next polyp will be $\frac{2}{7} - 2\varepsilon$ so no later paths will fit into these polyps. In the same way, region c) will fill up $\frac{N}{2}$ polyps. Finally, only one path of size $1 + \varepsilon$ may fit into every pair of arms in each polyp. This occupies N more polyps. In this way

$$N\left[\frac{1}{42} + \frac{1}{6} + \frac{1}{2} + 1\right] \geq N \cdot 1.6904$$

polyps were needed to pack these paths.

The optimal packing of the list L packs one path of each region to one pair of arms. This packing is optimal since the sum of the sizes of the paths is equal to the sum of the sizes of the polyps. ■

It is easy to improve this construction with the paths $1+\varepsilon, 2/3+\varepsilon, 2/7+\varepsilon, 2/43+\varepsilon, 2/1807+\varepsilon, \ldots$ but the improvement is only about 0.001 in the ratio. On the other hand the following conjecture can be made.

CONJECTURE. . *If the arm size of the polyp is suitably large there exists a list L for any large enough M such that $OPT(L) \geq M$ and*

$$FF(L) \geq 1.7 \cdot OPT(L).$$

5. Odd number of arms

The method is almost the same as in the case of even arms but the arm without a pair makes some differences here. The more arms the less difference. For example if $p = 3$ the difference is quite big. But for $p = 1001$ the results are very close to the even case.

THEOREM 4. *If $p \geq 7$ is odd then for every list L*

$$FF(L) \leq OPT(L)\left[1 + \frac{7p^2 + 6p - 1}{10p^2 - 2p}\right] + 4.$$

For some values of p this gives

$$\begin{aligned} R_7 &< 1.807 \\ R_9 &< 1.783 \\ R_{11} &< 1.768 \\ R_{15} &< 1.750 \\ R_\infty &< 1.7 \end{aligned}$$

PROOF: The weight function used in this case is the following:

$$W(x) = \begin{cases} \frac{6}{5p-1}x, & \text{if } 0 < x \leq \frac{1}{3} \\ \frac{9p-3}{p(5p-1)}x - \frac{p-1}{p(5p-1)}, & \text{if } \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{6}{5p-1}x + \frac{p-1}{p(5p-1)}, & \text{if } \frac{2}{3} \leq x \leq 1 \\ \frac{2}{p-1}, & \text{if } 1 < x \leq 2. \end{cases}$$

This implies

$$W\left(\frac{1}{3}\right) = \frac{2}{5p-1}, \quad W\left(\frac{2}{3}\right) = \frac{1}{p}, \quad W(1) = \frac{7p-1}{p(5p-1)}.$$

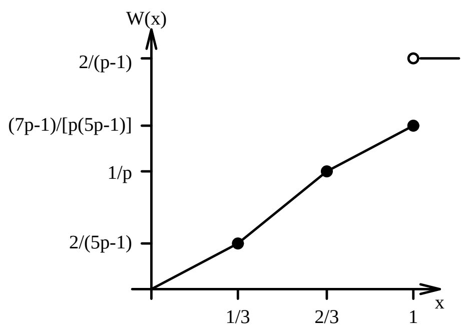

FIG. 4.

The lemmas needed to the proof are very similar to the even case.

LEMMA 4. *Let some polyp be filled with paths $l_1, l_2, \ldots, l_t$. Then*

$$\sum_{i=1}^{t} W(l_i) \leq 1 + \frac{7p^2 + 6p - 1}{2p(5p-1)}.$$

PROOF: If a polyp is filled by the above paths then some of the paths may cover some parts of two arms of the polyp. If this is the case then these two arms are called *paired arms*. A maximal matching may be defined on the arms. If there are some paired arms they will be pairs in the matching, too. The number of unpaired arms must be odd since the number of arms is odd. An arbitrary matching is made between the leftover arms missing only one arm. This arm is called the *left arm.*

Let us first consider a pair of arms. If there is a path of size b such that $\frac{2}{3} < b \leq 1$ then replace it with two paths one with size $\frac{2}{3}$ and one with size $b - \frac{2}{3}$. The weight sum in the polyp was not changed by this modification because the slope of the weight function is the same in $(0, \frac{1}{3}]$ and in $[\frac{2}{3}, 1]$. Hence it may be supposed that there are no path of size in range $(\frac{2}{3}, 1]$ in the polyp. If there are two paths of size $b_1 \leq b_2 \leq \frac{1}{3}$ then they are replaced with a single path of size $b_1 + b_2$. In this way the weight sum of the polyp may increase but cannot decrease. Therefore it can be assumed that there are at most one path smaller than $\frac{1}{3}$ in every pair of arms. These replacements can be made in the left arm also, so the assumptions hold there too.

First it will be proved that the following inequality holds in a pair of arms A:

$$\sum_A W(b_i) \leq \frac{2}{p-1} + \frac{7p-1}{p(5p-1)} = U.$$

Let m be the number of paths in A. It is easy to see that $m \leq 6$.
$\mathbf{m} = 1$. The maximum weight of a path is $\frac{2}{p-1} < U$.
$\mathbf{m} = 2$. If $b_1 > 1$ then $b_2 \leq \frac{2}{3}$ thus

$$W(b_1) + W(b_2) \leq \frac{2}{p-1} + \frac{1}{p} < U.$$

If $\frac{2}{3} \geq b_1 \geq b_2$ then

$$W(b_1) + W(b_2) \leq 2\frac{1}{p} < U.$$

$\mathbf{m} = 3$. If $b_1 > 1$ and $\frac{2}{3} \geq b_2 \geq \frac{1}{3} \geq b_3$ then

$$\sum_A W(b_i) \leq \frac{2}{p-1} + W(\frac{2}{3}) + W(\frac{1}{3}) = U.$$

If $b_1 > 1$ and $\frac{2}{3} \geq b_2 \geq b_3 \geq \frac{1}{3}$ then $b_2 + b_3 < 1$ so

$$\sum_A W(b_i) = \frac{2}{p-1} + (b_2 + b_3)\frac{9p-3}{p\,(5p-1)} - 2\frac{p-1}{p\,(5p-1)} \leq \frac{2}{p-1} + \frac{7p-1}{p\,(5p-1)} = U.$$

If $\frac{2}{3} \geq b_1 \geq b_2 \geq b_3$ then

$$\sum_A W(b_i) \leq 3W\left(\frac{2}{3}\right) < U.$$

$\mathbf{m} = 4$. If $b_1 > 1$ then $\frac{2}{3} \geq b_2 \geq b_3 \geq \frac{1}{3}$ and $\frac{1}{3} > b_4$. Thus

$$\begin{aligned}\sum_A W(b_i) &= \frac{2}{p-1} + (b_2 + b_3)\frac{9p-3}{p\,(5p-1)} - 2\frac{p-1}{p\,(5p-1)} + b_4\frac{6}{5p-1} = \\ &= \frac{2}{p-1} + (b_2 + b_3 + b_4)\frac{9p-3}{p\,(5p-1)} - 2\frac{p-1}{p\,(5p-1)} - b_4\frac{3p-3}{p(5p-1)} < U,\end{aligned}$$

since $b_2 + b_3 + b_4 < 1$ and $b_4 \geq 0$.
If $\frac{2}{3} \geq b_1 \geq b_2 \geq b_3 \geq b_4 \geq \frac{1}{3}$ then

$$\begin{aligned}&\sum_A W(b_i) = (b_1 + b_2 + b_3 + b_4)\frac{9p-3}{p\,(5p-1)} - 4\frac{p-1}{p\,(5p-1)} \leq \\ &\leq 2\frac{9p-3}{p\,(5p-1)} - 4\frac{p-1}{p\,(5p-1)} < U,\end{aligned}$$

since this inequality is equivalent to the trivial $1 < 3p^2 + 6p$.
If $\frac{2}{3} \geq b_1 \geq b_2 \geq b_3 \geq \frac{1}{3} > b_4$ then

$$\sum_A W(b_i) = (b_1 + b_2 + b_3)\frac{9p-3}{p\,(5p-1)} - 3\frac{p-1}{p\,(5p-1)} + b_4\frac{6}{5p-1} =$$
$$= (b_1 + b_2 + b_3 + b_4)\frac{9p-3}{p\,(5p-1)} - 3\frac{p-1}{p\,(5p-1)} - b_4\frac{3p-3}{p(5p-1)} < U,$$

since $b_1 + b_2 + b_3 + b_4 \leq 2$, $b_4 \geq 0$ and $1 < p^4 + 4p$.
$\mathbf{m} = 5$. If $b_1 > 1$ then three of the four other paths must be at least $\frac{1}{3}$ because there is only one path smaller than $\frac{1}{3}$. This is a contradiction. The rest of the cases can be proved with exactly the same argument as the corresponding cases in the previous paragraph.
$\mathbf{m} = 6$. Only one case is possible: $\frac{2}{3} \geq b_1 \geq \ldots \geq b_5 \geq \frac{1}{3} \geq b_6$. This can be proved in the same way as above.

It remained to prove that in the left arm A_l

$$\sum_{A_l} W(b_i) \leq \frac{7p-1}{p(5p-1)} = U'$$

holds. Let $1 - \varepsilon$ be the sum of the sizes in A_l. If $\varepsilon > 0$ then suppose that the inequality does not hold and that there is pair of arms A that contain the same paths as in A_l and an additional path b_0 of length > 1. Then

$$\sum_A W(b_i) > \frac{2}{p-1} + \frac{7p-1}{p(5p-1)} = U,$$

a contradiction to the previous case. Therefore the inequality is true for any $\varepsilon > 0$. The weight function is continuous in $(0, 1]$ therefore it must be also true for $\varepsilon = 0$.

To finish the proof of the lemma notice that there are $\frac{p-1}{2}$ pairs of arms and one left arm, thus

$$\sum_{i=1}^{t} W(l_i) \leq \frac{p-1}{2}U + U' =$$
$$= \frac{p-1}{2}\left[\frac{2}{p-1} + \frac{7p-1}{p(5p-1)}\right] + \frac{7p-1}{p(5p-1)} = 1 + \frac{7p^2+6p-1}{2p(5p-1)}. \quad \blacksquare$$

Lemma 5. *Let a polyp P_d of coarseness $\alpha_d \leq 1$ be filled with paths $l_1 \geq l_2 \geq \ldots \geq l_t$ with the completed First Fit algorithm. If $p \geq 7$ and $\sum_{i=1}^{t} W(l_i) < 1$, then $\alpha_{d+1} > \alpha_d$.*

Proof: Let A be the pair of arms where the weight sum is minimal among the pairs. Suppose that no path of size $> \alpha_d$ fits to the polyp P_d after completing the First Fit algorithm, therefore such a path does not fit neither to A nor to the left arm A_l. It will be proved in each case that this implies

$$\sum_{i=1}^{t} W(l_i) \geq 1,$$

a contradiction.

If there is a path of size > 1 in A then

$$\sum_{i=1}^{t} W(l_i) \geq \frac{p-1}{2}\frac{2}{p-1} = 1$$

holds. Since a path of size > 1 cannot fit to A_l it may be assumed that there is no path > 1 in the polyp.

Case 1: $\alpha_d \le \frac{1}{3}$.
It is easy to see in FIG. 5 that $W(x) \ge \frac{6}{5p-1}x$ holds in $(0,1]$

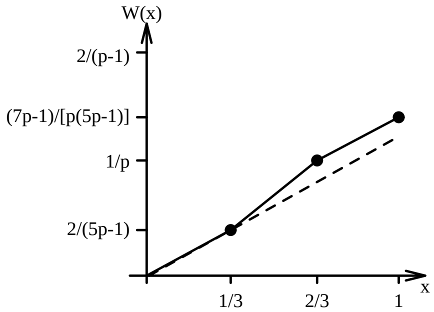

FIG. 5.

Therefore

$$\sum_A W(b_i) > (2-\alpha_d)\frac{6}{5p-1} \quad \text{and} \quad \sum_{A_l} W(b_i) > (1-\alpha_d)\frac{6}{5p-1}$$

holds, since the sum of the lengths must be $> (2-\alpha_d)$ in A and $> (1-\alpha_d)$ in A_l.

All the other pairs have weight sums at least as much as A, so

$$\sum_{i=1}^{t} W(l_i) > \frac{p-1}{2}(2-\alpha_d)\frac{6}{5p-1} + (1-\alpha_d)\frac{6}{5p-1}$$

holds. It is easy to see that this is at least 1 if $\alpha_d \le \frac{1}{3}$, a contradiction.

Case 2: $\frac{1}{3} < \alpha_d \le \frac{2}{3}$.

Let m be the number of paths in A and k the number of paths in A_l. Clearly $1 \le m \le 5$ and $1 \le k \le 2$. Let $b_1 \ge \ldots \ge b_m$ denote the paths in A and $a_1 \ge \ldots \ge a_k$ $(k \ge 1)$ the paths in A_l.

$\mathbf{m} = 1$. In this case $b_1 > 1$, since otherwise a path of size $> \alpha_d$ would fit into A, and this case is proved above.

$\mathbf{m} = 2$. If $b_1 \ge b_2 \ge \frac{2}{3}$ then

$$\begin{aligned}\sum_{i=1}^{t} W(l_i) &\ge \left[\frac{6}{5p-1}(b_1+b_2) + 2\frac{p-1}{p(5p-1)}\right]\frac{p-1}{2} + W(a_1) > \\ &> \left[\frac{6}{5p-1}(2-\alpha_d) + 2\frac{p-1}{p(5p-1)}\right]\frac{p-1}{2} + \frac{9p-3}{p\,(5p-1)}\alpha_d - \frac{p-1}{p\,(5p-1)} = \\ &= \frac{-3p^2+12p-3}{p\,(5p-1)}\alpha_d + \frac{7p^2-9p+2}{p\,(5p-1)} \ge 1,\end{aligned}$$

since $b_1+b_2 > 2-\alpha_d$, $\;a_1 > \alpha_d$, the coefficient of α_d is negative if $p \ge 3$ and $\alpha_d \le \frac{2}{3}$.

If $b_1 \ge \frac{2}{3} \ge b_2 > \alpha_d > \frac{1}{3}$ then

$$\begin{aligned}\sum_{i=1}^{t} W(l_i) &\ge \left[\frac{6}{5p-1}b_1 + \frac{9p-3}{p(5p-1)}b_2\right]\frac{p-1}{2} + W(\alpha_d) > \\ &> \left[\frac{6}{5p-1}(2-\alpha_d) + \frac{3p-3}{p(5p-1)}\alpha_d\right]\frac{p-1}{2} + W(\alpha_d) \ge 1\end{aligned}$$

using the same observations as before.

If $\frac{2}{3} > b_1 \ge b_2$ then a path of size $> \alpha_d$ fits to A which is a contradiction.

$\mathbf{m} = 3$. If $b_1 = b_2 = b_3$ then

$$\sum_{i=1}^{t} W(l_i) \geq \frac{3}{p}\frac{p-1}{2} \geq 1$$

if $p \geq 3$.

The case $b_1 \geq b_2 \geq \frac{2}{3} > b_3 > \alpha_d$ can be proved exactly like the case $m = 2$, $b_1 \geq \frac{2}{3} \geq b_2 > \alpha_d$.

If $b_1 \geq \frac{2}{3} > b_2 \geq b_3 > \alpha_d$ then

$$\begin{aligned}
&\sum_{i=1}^{t} W(l_i) \geq \\
&\geq \left[\frac{6}{5p-1}b_1 + \frac{9p-3}{p(5p-1)}(b_2+b_3) - \frac{p-1}{p(5p-1)}\right]\frac{p-1}{2} + W(\alpha_d) = \\
&= \left[\frac{6}{5p-1}(b_1+b_2+b_3) + \frac{3p-3}{p(5p-1)}(b_2+b_3) - \frac{p-1}{p(5p-1)}\right]\frac{p-1}{2} + W(\alpha_d) > \\
&> \left[\frac{6}{5p-1}(2-\alpha_d) + \frac{3p-3}{p(5p-1)}(2\alpha_d) - \frac{p-1}{p(5p-1)}\right]\frac{p-1}{2} + \frac{6}{5p-1}(1-\alpha_d)
\end{aligned}$$

holds, since $\sum_{i=1}^{k} a_i > 1 - \alpha_d$ and $W(x) \geq \frac{6}{5p-1}x$ holds in $(0,1]$ (see figure 5.). The coefficient of α_d is positive in this expression, so it takes its minimum if $\alpha_d = \frac{1}{3}$. It is easy to check that the expression is ≥ 1 in this case if $p \geq 3$.

If $\frac{2}{3} > b_1 \geq b_2 \geq b_3 > \alpha_d \geq \frac{1}{2}$ then

$$\begin{aligned}
\sum_{i=1}^{t} W(l_i) &\geq \left[\frac{9p-3}{p(5p-1)}(b_1+b_2+b_3) - 3\frac{p-1}{p\,(5p-1)}\right]\frac{p-1}{2} + W\left(\alpha_d\right) \geq \\
&\geq \left[\frac{9p-3}{p(5p-1)}(3\alpha_d) - 3\frac{p-1}{p\,(5p-1)}\right]\frac{p-1}{2} + \frac{9p-3}{p\,(5p-1)}\alpha_d - \frac{p-1}{p\,(5p-1)} \geq 1
\end{aligned}$$

holds if $p \geq 7$, since the coefficient of α_d is positive and $\alpha_d \geq \frac{1}{2}$.

If $\frac{2}{3} > b_1 \geq b_2 \geq b_3 > \alpha_d$ and $\alpha_d < \frac{1}{2}$ then we first prove that the sum of the weights in A_l is $\geq \frac{7p-1}{2p(5p-1)}$. The sum of the sizes is $> 1 - \alpha_d$ hence there is a path of size $\geq \frac{1}{2}$ or there are two paths both $> \frac{1}{3}$. It is easy to check that our claim is true in both cases. Therefore we obtain

$$\begin{aligned}
\sum_{i=1}^{t} W(l_i) &\geq \left[\frac{9p-3}{p(5p-1)}(b_1+b_2+b_3) - 3\frac{p-1}{p\,(5p-1)}\right]\frac{p-1}{2} + \frac{7p-1}{2p\,(5p-1)} \geq \\
&\geq \left[\frac{9p-3}{p(5p-1)}(2-\alpha_d) - 3\frac{p-1}{p\,(5p-1)}\right]\frac{p-1}{2} + \frac{7p-1}{2p\,(5p-1)} \geq 1,
\end{aligned}$$

if $p \geq 7$ since $b_1 + b_2 + b_3 > 2 - \alpha_d$, the coefficient of α_d is negative and $\alpha_d < \frac{1}{2}$. (Note that these are the only cases which are not true for $p = 3, 5$.)

$\mathbf{m} = 4$. The case $b_1 \geq b_2 \geq \frac{2}{3} > b_3 \geq b_4 > \alpha_d$ can be proved exactly like the case $m = 2$, $b_1 \geq \frac{2}{3} > b_2 > \alpha_d$ and the case $b_1 \geq \frac{2}{3} > b_2 \geq b_3 \geq b_4 > \alpha_d$ exactly like the case $m = 3$, $b_1 \geq \frac{2}{3} > b_2 \geq b_3 > \alpha_d$.

If $\frac{2}{3} > b_1 \geq b_2 \geq b_3 \geq b_4 > \alpha_d$ then the proof is similar to the corresponding proof in the previous case. In this case

$$\sum_{i=1}^{t} W(l_i) \geq \left[\frac{9p-3}{p(5p-1)}(b_1+b_2+b_3+b_4)\right]\frac{p-1}{2} + W\left(\frac{1}{3}\right)$$

holds. This expression takes its minimum if $b_1 + b_2 + b_3 + b_4$ achieves its minimum. $b_1 + b_2 + b_3 + b_4 > 4\alpha_d$ and $b_1 + b_2 + b_3 + b_4 > 2 - \alpha_d$ therefore the minimum, $\frac{8}{5}$, is taken when $\alpha_d = \frac{2}{5}$. Then it it easy to see that the above expression is > 1.

$\mathbf{m} = 5$. The only possible case is $\frac{2}{3} > b_1 \geq \ldots \geq b_5 > \alpha_d$. Then the lower bound $W(x) \geq \frac{6}{5p-1}$ is used again (see FIG. 5). All the paths in the paired arms have length $> \alpha_d$ and the sum of the sizes in A_l is $> (1-\alpha_d)$, therefore

$$\sum_{i=1}^{t} W(l_i) > \frac{p-1}{2}\frac{6}{5p-1}5\alpha_d + \frac{6}{5p-1}(1-\alpha_d) = \alpha_d\frac{15p-21}{5p-1} + \frac{6}{5p-1} \geq 1$$

is obtained since $\alpha_d > \frac{1}{3}$.

Case 3: $\frac{2}{3} < \alpha_d \leq 1$.

By the definition of the coarseness every path in the polyp must be larger than $\frac{2}{3}$. The sum of the sizes in a pair of arms must be $> \frac{4}{3}$, therefore there must be at least two paths in each pair. Similarly there must be at least one paths in A_l. Thus $\sum_{i=1}^{t} W(l_i) > pW\left(\frac{2}{3}\right) = 1$ is obtained completing the proof of the lemma. ■

LEMMA 6. *Let a polyp P_d of coarseness $\alpha_d \leq 1$ be filled with paths $l_1 \geq l_2 \geq \ldots \geq l_t$ in the completed First Fit algorithm. If $\sum_{i=1}^{t} W(l_i) = 1 - \beta \quad (0 < \beta < 1)$ then $\alpha_{d+1} \geq \min(\alpha_d + \frac{5}{9}\beta, 1)$.*

The proof of this lemma is very similar to the proof of Lemma 3 so it is left to the reader. ■

The rest of the proof of Theorem 4 is almost the same as the proof of Theorem 2 so it is also left to the reader. ■

By slight modification in the weight function we can prove that $R_3 < 2$ and $R_5 < 1.858$ which does not differ very much from the value of the formula in Theorem 4 for $p = 3$ and $p = 5$.

THEOREM 5. *If the arm size of the polyp is suitably large then there exists a list L for any large enough M such that $OPT(L) \geq M$ and*

$$FF(L) \geq OPT(L)\left[\frac{p+1}{42p} + \frac{p+1}{6p} + \frac{p+1}{2p} + 1\right]$$

For some values of p this gives the following lower bounds for the ratio (the upper bound is from Theorem 4).

$$\begin{aligned} 1.920 &< R_3(< 2) \\ 1.828 &< R_5(< 1.858) \\ 1.789 &< R_7(< 1.807) \\ 1.736 &< R_{15}(< 1.750) \\ 1.69 &< R_\infty(< 1.7) \end{aligned}$$

The lower and upper bounds are quite close.

PROOF: This proof is also very similar to the proof of Theorem 3 but it is worth making the differences clear.

Let N be a positive integer divisible by $42p$ and let ε a be a suitably small rational number. The list of paths denoted by L consists of the following regions of paths in the order of appearance

a) $N\frac{p-1}{2}$ paths of size $\frac{1}{21} - 3\varepsilon$,
a') N paths of size $\frac{1}{21} - 2\varepsilon$,
b) $N\frac{p+1}{2}$ paths of size $\frac{2}{7} + \varepsilon$,
c) $N\frac{p+1}{2}$ paths of size $\frac{2}{3} + \varepsilon$,
d) $N\frac{p-1}{2}$ paths of size $1 + \varepsilon$.

If these paths are packed with the First Fit algorithm then 42 paths from a) and a') fits to a pair of arms and 21 to the left arm. So

$$\frac{N\frac{p-1}{2} + N}{42\frac{p-1}{2} + 21} = N\frac{p+1}{42p}$$

polyps are needed to pack a) and a'). Similarly one can see that

$$N\frac{p+1}{6p} + N\frac{p+1}{2p}$$

additional polyps are needed to pack paths from b) and c). Note that as in the even case the coarseness of the polyps are increasing. Finally the main difference comes from the fact that only $\frac{p-1}{2}$ paths of size $1+\varepsilon$ fits to one polyp. So we need N more polyps to finish the packing.

On the other hand it is clear that these paths fit to N polyps and this is an optimal packing. ∎

There is one more construction worth mentioning. This is an other lower bound which gives a somewhat better result for $p = 3$ than Theorem 5, but for larger number of arms Theorem 5 is better.

THEOREM 6. *If the arm size of the polyp is suitably large then there exists a list L for any large enough M such that $OPT(L) \geq M$ and*

$$FF(L) \geq OPT(L)\left[\frac{p+1}{24p} + \frac{p+1}{7p-1} + \frac{p+1}{2p} + 1\right].$$

This gives $1.922 < R_3 (< 2)$.

PROOF: Let N be a positive integer divisible by $24p$ and $7p-1$. The list L of paths consists of the following regions of paths in the order of appearance

a) $N\frac{p-1}{2}$ paths of size $\frac{1}{12} - 3\varepsilon$

a') N paths of size $\frac{1}{12} - 2\varepsilon$

b) $N\frac{p+1}{2}$ paths of size $\frac{1}{4} + \varepsilon$

c) $N\frac{p+1}{2}$ paths of size $\frac{2}{3} + \varepsilon$

d) $N\frac{p-1}{2}$ paths of size $1 + \varepsilon$

The proof otherwise is the same as above so it is left to the reader. ∎

6. Conclusions

We have seen that although it looks very hard to find an optimal packing the very easily applicable First Fit algorithm works quite well. So if we can model some practical problem by the Polyp–packing one might use the First Fit algorithm.

The edge disjoint p–polyp-packing may be considered as a variant of the Bin Packing problem. Now we have groups of bins each group containing p bins. It is possible to break items into two piece and place the two parts into two different bins of the same group, but only one broken part is allowed in each bin. In [3] the reader may find some problems which are modeled by the Bin Packing problem. With this modification we may model some modified problems. For example file allocation. It is desired to place files of varying sizes on as few disks as possible, where files may be broken into two parts and placed to different tracks of the same disk.

There are some open problems left. First of all, our results are not sharp so they may be improved.

There are algorithms for the Bin Packing problem which give packing closer to the optimal. Probably such an algorithm would give a better result here also. On the other hand it is proved in [2] that for every positive ε there exists an $O(n)$-time algorithm for Bin Packing that uses at most $(1+\varepsilon)$ times more bins than the optimal packing. This may be true for Polyp Packing too.

Further investigation of this problem may include other graphs not just polyps. However there should be some restrictions in the vertex disjoint case. If we want to pack arbitrary paths into arbitrary graphs vertex disjointly then, with any algorithm, when we want to pack a particular path, it must be decided if this path fits into the graph or not. But this problem is NP-complete in general, since it contains the hamiltonian path problem as a special case. One possible restriction is that the sizes of the paths have a constant upper bound. For example if all paths are at most 100 long, this problem does not arise.

However this problem does not arise in the edge disjoint case, because there is a polynomial algorithm to decide if it is possible to pack a given path to a graph such that it may cross a vertex more than once but it may cover each edge at most once. This problem is equivalent to the Chinese Postman problem which is solved in polynomial time in [1].

References

1. J. Edmonds, E. Johnson: *Euler Tours and the Chinese Postman problem*, Math. Programming **5**, (1973) pp. 88-124.
2. W. Fernandez de la Vega and G. S. Lueker: *Bin packing can be solved within* $1 + \varepsilon$ *in linear time*, Combinatorica **1** (1981) pp. 349-355.
3. D. S. Johnson, A. Demers, J. D. Ullman, M. R. Garey and R. L. Graham: *Worst case bounds for simple one-dimensional packing algorithms*, SIAM J. Comptg. **3** (1974), pp. 299-325.

Gyula Y. Katona
Mathematical Institute of the
Hungarian Academy of Sciences
1364 BUDAPEST Pf. 127
E-mail: kiskat@ cs.elte.hu