

## Development of an Algorithm Improving Label Arrangements in Offset Printing

**Geun Soo Jang · Taehyeong Kim ·  
Ki Man Kong · Jeong Rye Park ·  
Jong-hyeon Seo · Sang-hyup Seo<sup>†</sup> ·  
Shin won Yoon**

Received: date / Accepted: date

**Abstract** One of the most classic problems in the manufacturing industry is inventory processing. There is a way to effectively reduce inventory by changing the array of the pieces on the printing plates in the offset printing. It is setting an upper limit of acceptable for each plate, and carrying out complete enumeration. This method dra-

---

Geun Soo Jang  
Finance.Fishery.Manufacture Industrial Mathematics Center on Big Data, Pusan National University, Busan, 46241, Republic of Korea  
E-mail: sand621@naver.com

Taehyeong Kim  
Finance.Fishery.Manufacture Industrial Mathematics Center on Big Data, Pusan National University, Busan, 46241, Republic of Korea  
E-mail: xogud7936@pusan.ac.kr

Ki Man Kong  
World Komax Co., Ltd., 1505, Centum Jungang-Ro 48, Haeundae-Gu, Busan, 48059, Republic of Korea  
E-mail: kennethkong@worldkomax.net

Jeong Rye Park  
Finance.Fishery.Manufacture Industrial Mathematics Center on Big Data, Pusan National University, Busan, 46241, Republic of Korea  
E-mail: parkjr@pusan.ac.kr

Jong-hyeon Seo  
Chubu University Academy of Emerging Science, Kasugai, 487-0027, Japan  
E-mail: hyeonni94@gmail.com

Sang-hyup Seo  
<sup>†</sup>Corresponding Author  
Finance.Fishery.Manufacture Industrial Mathematics Center on Big Data, Pusan National University, Busan, 46241, Republic of Korea  
E-mail: saibie1677@gmail.com

Shin won Yoon  
Finance.Fishery.Manufacture Industrial Mathematics Center on Big Data, Pusan National University, Busan, 46241, Republic of Korea  
E-mail: ysw0123@pusan.ac.kr

matically reduces the operating time of the algorithm. The advantage of this method is that there is only change the arrangement of the pieces on the plates.

**Keywords** Offset Printings · Combinations with Repetition · Label Arrangements · Inventory Managements

## 1 Introduction

A combination with repetition is the number of cases, where  $k$  elements are selected from among different  $n$  elements allowing repetition [1]. It is indicated with the symbol  ${}_nH_k$  and the following is established.

$${}_nH_k = {}_{n+k-1}C_k = \frac{(n+k-1)!}{(n-1)!k!} \quad (1)$$

For instance, the combination with repetition  ${}_2H_4$  to select four elements from among two elements A and B comprises the following five cases.

- (1) [A, A, A, A] : a list consisting of four A's
- (2) [A, A, A, B] : a list consisting of three A's and one B
- (3) [A, A, B, B] : a list consisting of two each of A's and B's
- (4) [A, B, B, B] : a list consisting of one A and three B's
- (5) [B, B, B, B] : a list consisting of four B's

Among the above cases, if we want to obtain three A's and nine B's, then we can choose  $[A, B, B, B] \times 3$ . Also, we consider the following as another case.

$$[A, A, B, B] \times 1 + [A, B, B, B] \times 1 + [B, B, B, B] \times 1 \quad (2)$$

In this case, we can get the three A's and nine B's. However, the former case seems to be a 'Better' because there are three different lists in (2). Let us examine the other case.

$$[A, A, A, A] \times 1 + [B, B, B, B] \times 3 - [A] \times 1 - [B] \times 3 \quad (3)$$

$[A, B, B, B] \times 3$  also seems to be 'Better' because there is no loss. Under the following conditions,  $[A, B, B, B] \times 3$  is the 'Best' method.

- (1) Minimize the number of list.
- (2) Minimize the loss of lists.

Offset printing, also called offset lithography, or litho-offset, in commercial printing, widely used printing technique in which the inked image on a printing plate is printed on a rubber cylinder and then transferred (i.e., offset) to paper or other material. The rubber cylinder gives great flexibility, permitting printing on wood, cloth, metal, leather, and rough paper (see Figure 1) [3].

Offset printing is one of the most common ways of creating printed materials. A few of its common applications include: newspapers, magazines, brochures, stationery, and books. Compared to other printing methods, offset printing is best suited

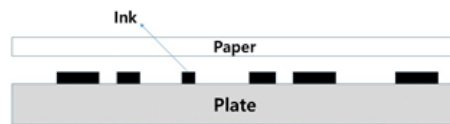


Fig. 1: Offset Printing

for economically producing large volumes of high quality prints in a manner that requires little maintenance [2]. Therefore, how to make the initial plates is an important issue. Above example means that what is the best arrangement in such print method.

In the past, production was based on ordering of products from companies and predictions of consumption. However, as the internet market has become popular, the production systems have been changed by consumers. Now, many factories produce only products ordered by consumers.

World Komax is a company that produces labels using the offset printing. The labels refer to stickers containing bar-codes attached to garments or shoes as follows (see Figure 2). Each bar-code in the label contains fixed information such as product names and colors, and variable information such as the date of manufacture.

Fig. 2: Sneaker Review : Nike Air Huarache (<http://visla.kr/?p=18146>)

Since the program of World Komax is not suitable for small quantity batch productions, the number of labels loss increased compared to the past.

Section 2 of this paper will describe the process of label printing using offsets. The modeling of the problem will be carried out in Section 3, and examples to help the understanding of the problem will be prepared in Section 4. The final results of the algorithm will be described in Section 5.

## 2 Offset label printing process

### 2.1 Label printing process

Prior to describing the label printing process, we define the following terms first.

\* **Plate** : A printing plate for the offset printing (see Figure 3)

\* **Loss** : The number of labels printed in excess of the order-quantity

The offset label printing process is as follows. At first, we receive orders from customers. The order includes many types of labels and order-quantities by type (see Figure 4). Thereafter, offset printing plates are made. Many types of labels are placed on each plate so that many labels are printed at one printing. After the plates are made, we produce more labels than order-quantities using each plate. Then, the sheets are cut to the sizes of labels. As the final process, the labels are collected by type.

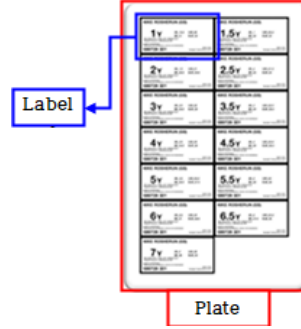


Fig. 3: Plate and Label

## 2.2 Major points for cost saving

The constraint conditions and major points that will be considered in this paper for cost saving are as follows. First, one type of label should be placed on only one plate. This is to prevent different types of labels from being mixed when collected by type after the printed sheets are cut. Meanwhile, the total number of labels placed on each plate is also constant because the sizes of individual plates are constant and the sizes of labels in one order are also constant. In addition, the number of plates should be minimized as little as possible because plates are made using molds and the costs are high. Finally, the Loss should be minimized because overprinted labels cannot be used and should be entirely discarded.

## 2.3 Sorting reports output program

Plate fabrication and the use of printing paper incur costs. To reduce the costs, order details are inputted to output appropriate methods to place labels on the plate as sorting reports. The plate makers produce plates according to the instructions in the sorting reports (see Figure 4).

Since the existing sorting report output method was not suitable for small quantity batch production systems, the algorithm should be improved. Therefore, this study was conducted to develop new algorithms suitable for small quantity batch production systems too.

## 3 Modeling and flowchart

To formulate our problem as a mathematical optimization problem, we first need to define our notation.

- Let  $I$  be a set of products.
- For each product  $i \in I$ , let  $b_i$  be the number of order.

The figure consists of two parts. On the left is an 'Order Form' for 'NB206 - 71706-SANDAL SLIDE'. It includes a barcode, a table of quantities for different sizes (e.g., 10.5, 11.5, 12.5, 13.5, 14.5, 15.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, 30.5, 31.5, 32.5, 33.5, 34.5, 35.5, 36.5, 37.5, 38.5, 39.5, 40.5, 41.5, 42.5, 43.5, 44.5, 45.5, 46.5, 47.5, 48.5, 49.5, 50.5, 51.5, 52.5, 53.5, 54.5, 55.5, 56.5, 57.5, 58.5, 59.5, 60.5, 61.5, 62.5, 63.5, 64.5, 65.5, 66.5, 67.5, 68.5, 69.5, 70.5, 71.5, 72.5, 73.5, 74.5, 75.5, 76.5, 77.5, 78.5, 79.5, 80.5, 81.5, 82.5, 83.5, 84.5, 85.5, 86.5, 87.5, 88.5, 89.5, 90.5, 91.5, 92.5, 93.5, 94.5, 95.5, 96.5, 97.5, 98.5, 99.5, 100.5), and a total quantity of 366518. On the right is a 'Sorting Report' for 'NB206 - 71706-SANDAL SLIDE'. It is a detailed table showing the arrangement of labels on plates, with columns for plate number, label number, and quantity. The table is organized into sections for different sizes (e.g., 10.5, 11.5, 12.5, 13.5, 14.5, 15.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, 30.5, 31.5, 32.5, 33.5, 34.5, 35.5, 36.5, 37.5, 38.5, 39.5, 40.5, 41.5, 42.5, 43.5, 44.5, 45.5, 46.5, 47.5, 48.5, 49.5, 50.5, 51.5, 52.5, 53.5, 54.5, 55.5, 56.5, 57.5, 58.5, 59.5, 60.5, 61.5, 62.5, 63.5, 64.5, 65.5, 66.5, 67.5, 68.5, 69.5, 70.5, 71.5, 72.5, 73.5, 74.5, 75.5, 76.5, 77.5, 78.5, 79.5, 80.5, 81.5, 82.5, 83.5, 84.5, 85.5, 86.5, 87.5, 88.5, 89.5, 90.5, 91.5, 92.5, 93.5, 94.5, 95.5, 96.5, 97.5, 98.5, 99.5, 100.5).

Fig. 4: An Order Form(left) and a Sorting Report(right)  
(Source: World Komax Co., Ltd.)

- Let  $\mathbf{b} = (b_i | i \in I)$  be a vector of order-quantities.
- $k$  is the total number of labels that can be placed in one Plate.
- $\pi$  is a partition of  $I$  such that  $1 \leq |P| \leq k$  for any  $P \in \pi$ .
- Let  $\Gamma_\pi$  be a set of matrix  $A \in \text{Mat}_{\pi \times I}(\mathbf{Z})$  satisfy the following:
  - For all  $(P, i) \in \pi \times I$ ,  $A_{P,i} \geq 0$ , and  $A_{P,i} = 0$  if and only if  $i \notin P$ .
  - For each  $P \in \pi$ ,  $\sum_{i \in P} A_{P,i} = k$

Using the above notation,

$$\left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \mid i \in P \right\} \right\rceil \quad (4)$$

is the number of printing of Plate  $P$ , where  $\lceil \cdot \rceil$  means the ceiling. Assume that  $\alpha$  is the cost to produce one Plate, and  $\beta$  is the cost of the loss of one label. Then, our goals is to obtain the following

$$\min_{\pi} \{ \alpha |\pi| + \beta E_{A,b} | A \in \Gamma_\pi \} \quad (5)$$

where

$$E_{A,b} = \sum_{P \in \pi} \sum_{i \in I} \left( \left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \mid i \in P \right\} \right\rceil \cdot A_{P,i} - b_i \right) \quad (6)$$

means the total number of losses of labels.

Naturally, the complete enumeration using combinations with repetition is the surest way. However, this method has a problem of taking too much time. For instance, when  $n = 65$  and  $k = 24$ , the combination with repetition  ${}_{65}H_{24}$  comprises  $2.36 \times 10^{21}$  cases, and the calculation of the cases takes more than 658 hours, that is, more than 27 days using a super computer that can calculate  $10^{15}$  partitions per second. Given that there are limits of the time from the date of receipt of orders to the delivery date, this is a very long computation time.

In this algorithm, this problem was solved by introducing any positive integers as tolerances. Tolerances means the allowed amount of losses occurring in each plate. Adopting a partition that does not exceed the tolerance will dramatically reduce the time taken.

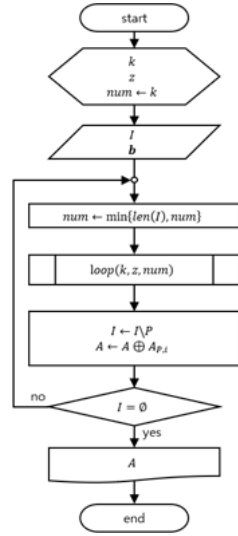


Fig. 5: The Main Flow Chart

Based on the foregoing, the flowchart of the algorithm can be set forth as follows. This algorithm outputs matrix  $A$  containing the label of each product when  $I$  and  $\mathbf{b}$  have been inputted for  $z$ ,  $k$ ,  $num$  (see Figure 5).

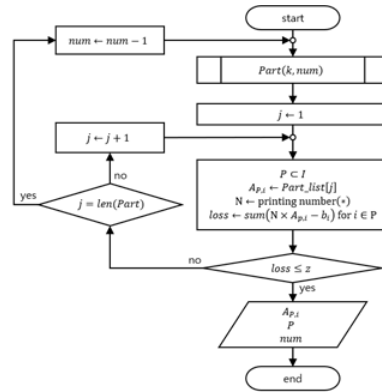


Fig. 6: The Flow Chart of Loop

Meanwhile, in the case of the  $loop(k, z, num)$  function, the following flowchart should be followed. First, the  $Part(k, num)$  function finds partitions using the combination with repetition  ${}_{num}H_k$  (method to select  $k$  pieces of products from  $num$  pieces of products allowing repetition) and indicated in the form of a list. We set the result as  $Part\_list$ . An appropriate  $P$  that has  $num$  pieces of products is selected from  $I$  and the loss is obtained using  $Part\_list$  and the printing number.

If the loss exceeds the tolerance  $z$ , another *Part\_list* will be selected and the foregoing will be repeated while adjusting *num* until the tolerance is not exceeded. As a result of this process,  $A_{P,i}$  that does not exceed the tolerance is obtained (see Figure 6).

#### 4 Examples

This example was described to help the understanding of the problem. In this section, we assume that  $k$  is equal to 4.

*Example 1* Let  $I = \{A, B, C\}$  and  $\mathbf{b} = (50, 30, 20)$ .

We consider  $\pi = \{\{A, B\}, \{C\}\}$  is a partition for the order-quantity vector  $\mathbf{b}$ . Without loss of generality, assume that  $P_1 = \{A, B\}, P_2 = \{C\}$ . Since  $k = 4$ , the matrix  $A$  can be found as follows.

$$A = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad (7)$$

In this case, the printing numbers (4) of  $P_1$  and  $P_2$  are as follows, respectively.

$$\left\lceil \max \left\{ \frac{b_i}{A_{P_1,i}} \mid i \in P_1 \right\} \right\rceil = \left\lceil \max \left\{ \frac{50}{2}, \frac{30}{2} \right\} \right\rceil = 25 \quad (8)$$

and

$$\left\lceil \max \left\{ \frac{b_i}{A_{P_2,i}} \mid i \in P_2 \right\} \right\rceil = \left\lceil \max \left\{ \frac{20}{4} \right\} \right\rceil = 5. \quad (9)$$

We can see the Figure 7 for more details.

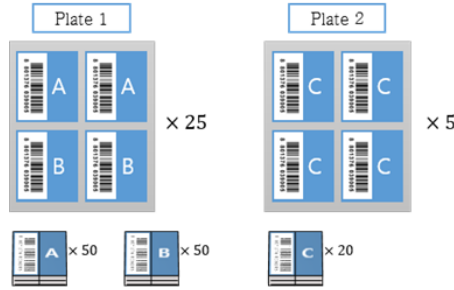


Fig. 7

It can be seen that the total number of Loss  $E_{A,\mathbf{b}}$  is 20.

Now, we consider a new partition  $\pi = \{\{A, B, C\}\}$ . In this case,  $A = (2 \ 1 \ 1)$ , and printing number (4) is

$$\left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \mid i \in P \right\} \right\rceil = \left\lceil \max \left\{ \frac{50}{2}, \frac{30}{1}, \frac{20}{1} \right\} \right\rceil = 30. \quad (10)$$

In addition, it can be easily seen that the total number of Loss  $E_{A,\mathbf{b}}$  is 20 (see Figure 8). The array shown in Figure 8 is a more efficient because its Loss is the same but its number of Plates is smaller.

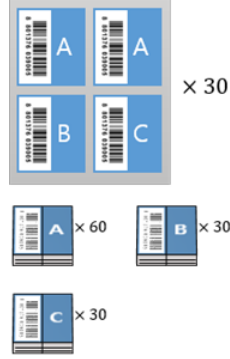


Fig. 8

*Example 2* Assume that  $I = \{A, B\}$ , and the order-quantity vector  $\mathbf{b} = (50, 20)$ .

We consider  $\pi = \{\{A, B\}\}$  is a partition for the quantity vector  $\mathbf{b}$ . Since  $k = 4$ , the matrix  $A = \begin{pmatrix} 2 & 2 \end{pmatrix}$ , and the printing number (4) is

$$\left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \mid i \in P \right\} \right\rceil = \left\lceil \max \left\{ \frac{50}{2}, \frac{20}{2} \right\} \right\rceil = 25. \quad (11)$$

The total number of Loss  $E_{A,\mathbf{b}}$  is 30 (see Figure 9).



Fig. 9

For the same partition  $\pi$ , the matrix  $A = \begin{pmatrix} 3 & 1 \end{pmatrix}$  can be considered. In this case, the printing number (4) is

$$\left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \mid i \in P \right\} \right\rceil = \left\lceil \max \left\{ \frac{50}{3}, \frac{20}{1} \right\} \right\rceil = 20. \quad (12)$$



and the total number of Loss  $E_{A,b}$  is 10 (see Figure 10). The array shown in Figure 10 is more efficient because its number of Plate is the same but fewer losses occur.

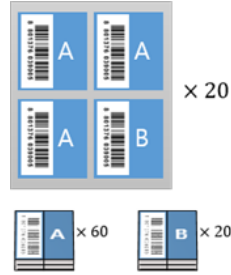


Fig. 10

## 5 Result

Each sorting report include the number of losses corresponding to one plate. Using this, the total cost can be calculated by replacing each plate with loss. We used 30 sorting report samples. The algorithm can be improved for every sample data. The total cost was reduced by from minimum 0.4%(sample no. 15) to maximum 15.96%(sample no. 8) (see Figure 11).

We used the *paired t-test* to verify the efficiency of the algorithm. The pared t-test is one of the *two sample t-test*, and it is a test that verifies whether the two groups are different. The two populations are as follows.

- population1: total cost before applying the algorithm
- population2: total cost after applying the algorithm
- sample1: sample of 30 items from population1
- sample2: sample of 30 items from population2

In order to proceed with the two-sample t-test, the two groups have to satisfy the normality and homoscedasticity. We choose 30 sample data from the two populations, which can be satisfied the normality by the *central limit theorem*. In addition, we identified the homoscedasticity of the two samples through the *var.test* of R. R is a programming language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis.

The null hypothesis ( $H_0$ ) of *var.test* is that ‘the variances of the two groups are equal’, and the alternative hypothesis ( $H_1$ ) is that ‘the variances of the two groups are different’. If the p-value is below the significance level, the null hypothesis ( $H_0$ ) will be rejected and if the p-value is not lower than the significance level, the alternative hypothesis ( $H_1$ ) will be rejected. The result of *var.test* with the significance level to 0.05 is as follows.(see Figure 12.)

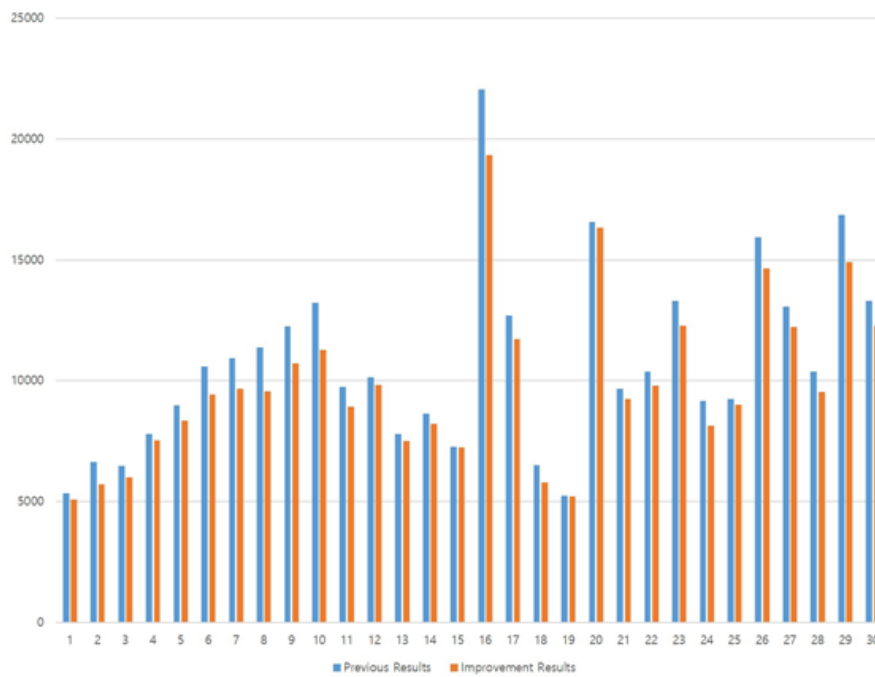


Fig. 11: Comparing the Results

```
> var.test(sample1,sample2)

      F test to compare two variances

data:  sample1 and sample2
F = 1.2647, num df = 26, denom df = 26, p-value = 0.5536
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5763597 2.7751597
sample estimates:
ratio of variances
 1.26471
```

Fig. 12: var.test of R

Since the p-value is not lower than the significance level(0.05), the alternative hypothesis ( $H_1$ ) is rejected and the null hypothesis ( $H_0$ ) is accepted. We say that the variances of the two groups can be equal.

Since the two groups satisfy normality and homoscedasticity, we verified whether the difference between the two groups is significant through paired t-test. In this test, the null hypothesis ( $H_0$ ) is 'the total cost will be the same after applying the algorithm.' and the alternative hypothesis ( $H_1$ ) is 'the total cost will be reduced after applying algorithm.' The result of the paired t-test with the significance level to 0.05 is as follows.(see Figure 13.)

```
> t.test(sample1,sample2, paired=TRUE)

Paired t-test

data: sample1 and sample2
t = 6.6852, df = 26, p-value = 4.306e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 572.3077 1080.5071
sample estimates:
mean of the differences
 826.4074
```

Fig. 13: t.test of R

Since the p-value is below the significance level(0.05), the null hypothesis ( $H_0$ ) is rejected and the alternative hypothesis ( $H_1$ ) is accepted. We say that the difference between the two groups can be significant.

**Notice.** Please note that the detailed idea of the algorithm cannot be described for confidentiality of the company.

**Acknowledgements** This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIP) (NRF-2017R1A5A1015722, NRF-2018R1D1A1B07048197). The authors extend thanks to Professor Hyun-Min Kim and Professor Sang-il Kim of Pusan National University who provided many support and advises for the writing of this paper.

## References

1. R. A. Brualdi, Introductory Combinatorics, Pearson Prentice Hall (2004)
2. H. Kipphan, Handbook of Print Media: Technologies and Production Methods, Springer (2001)
3. "Offset printing", Encyclopædia Britannica, inc., 19-06-2016.  
url: <https://www.britannica.com/technology/offset-printing>.