DEVELOPMENT OF AN ALGORITHM IMPROVING LABEL ARRANGEMENTS IN OFFSET PRINTING

GEUN SOO JANG, TAEHYEONG KIM, HYUN-MIN KIM, KI MAN KONG, JEONG RYE PARK, JONG-HYEON SEO, SANG-HYUP SEO†, AND SHIN WON YOON

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 pieces on the printing plates in the offset printing. It is by setting an acceptable upper limit for each plate, and carrying out complete enumeration. This method drastically reduces the operating time of the algorithm. The advantage of this method is that it focuses on changing the arrangement of the pieces on the plates.

1. INTRODUCTION

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

$$(1.1) \qquad _nH_k =_{n+k-1} C_k = \frac{(n+k-1)!}{(n-1)!k!}$$

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, we can choose [A, B, B, B] × 3. Also, we consider the following as another case.

$$(1.2) \qquad [A, A, B, B] \times 1 + [A, B, B, B] \times 1 + [B, B, B, B] \times 1$$

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 (1.2). Let us examine the other case.

$$(1.3) \qquad [A, A, A, A] \times 1 + [B, B, B, B] \times 3 - [A] \times 1 - [B] \times 3$$

[A, B, B, B] × 3 also seems to be a 'Better' because there is no loss. Under the following conditions, [A, B, B, B] × 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, is a 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 1.1) [5].
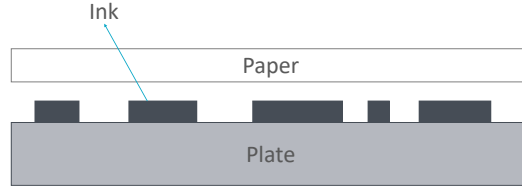

FIGURE 1.1. Offset Printing

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 for economically producing large volumes of high quality prints in a manner that requires little maintenance [6]. To improve printing process, there were many researches in [2, 3, 7]. As another improvement, how to make the initial plates is an important issue. The above example denotes what is the best arrangement in such print method.

In the past, production was based on the ordering of products from companies and predictions of consumption. However, as the internet market has become popular, the production systems have been altered 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 1.2). Each bar-code in the label contains fixed information such as product names and colors, and variable information such as the manufactured date.

FIGURE 1.2. Sneaker Review : Nike Air Huarache

Since the program of World Komax is not suitable for small quantity batch productions, the number of label 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.** Before discussing the process of label printing, we define the following terms.

* **Plate** : A printing plate for the offset printing (see Figure 2.1)
* **Loss** : The number of labels printed in excess of the order-quantity

The offset label printing process is as follows. First, we receive orders from customers. The order includes many types of labels and order-quantities by type (see Figure 2.2). Thereafter, offset printing plates are made. Many types of labels are placed on each plate so that many labels are printed at once. After the plates are made, we produce more labels than order-quantities using each plate. Then, the sheets are cut into the label size. For the final process, the labels are collected by type.

2.2. **Major points for cost saving.** The constraints 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 as minimized 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.
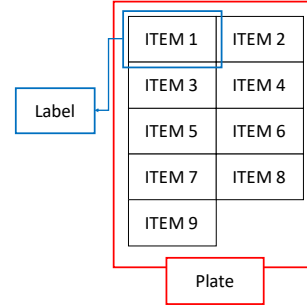

FIGURE 2.1. Plate and Label

2.3. **Sorting reports output program.** The production of plates 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 2.2).

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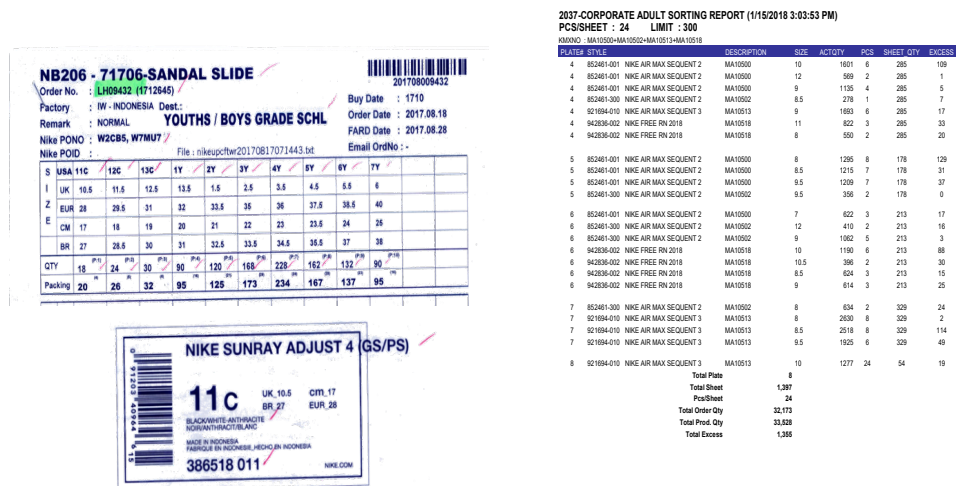

FIGURE 2.2. An Order Form(left) and a Sorting Report(right)
(Source: World Komax Co., Ltd.)

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.
- 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.
- π is a partition of I such that $1 \leq |P| \leq k$ for any $P \in \pi$.
- Let Γ_π be a set of matrix $A \in \mathrm{Mat}_{\pi \times I}(\mathbf{Z})$ 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,

$$(3.1) \qquad \left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \,\middle|\, i \in P \right\} \right\rceil$$

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

$$(3.2) \qquad \min_\pi \{ \alpha |\pi| + \beta E_{A,b} | A \in \Gamma_\pi \}$$

where

$$(3.3) \qquad E_{A,b} = \sum_{P \in \pi} \sum_{i \in I} \left(\left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \,\middle|\, i \in P \right\} \right\rceil \cdot A_{P,i} - b_i \right)$$

means the total number of losses of labels.

Naturally, the complete enumeration using combinations with repetition is the surest way. However, this method takes too much time. For instance, when $n = 65$ and $k = 24$, the combination with repetition $_{65}H_{24}$ comprises about 2.36×10^{21} cases. Then, the calculation 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 time constraints from the date of order receipts to the delivery date, this is a very long computation time.

In this algorithm, this problem was solved by introducing any positive integers as thresholds. Thresholds mean the allowed amount of losses occurring in each plate. Adopting a partition that does not exceed the threshold will dramatically reduce the time taken. Based on the foregoing, the flowchart of the algorithm can be set forth as follows (see Figure 3.1).

In Figure 3.1, z is the threshold and the others are explained in the notation above. This algorithm outputs matrix $A = \bigoplus_P A_{P,i}$ containing the label of each product when I and $\mathbf{b}$ have been inputted for z, k, num. One of the results of $loop(k, z, num)$, P is the set of products that contains the Plate (see Figure 3.2). By removing P from I, each plate does not have the same product. The algorithm repeats until I is empty.

Meanwhile, in the case of the $loop(k, z, num)$ function, the flowchart in Figure 3.2 should be followed. First, the $Part(k, num)$ function finds partitions using the combination with repetition $_{num}H_k$ which is indicated in the form of a list. We set the result as $Part\_list$. For example, let $k = 6$ and $num = 3$, then $Part\_list = Part(6, 3) = \{[4, 1, 1], [3, 2, 1], [2, 2, 2]\}$. N is a printing number that expressed in (3.1). 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.
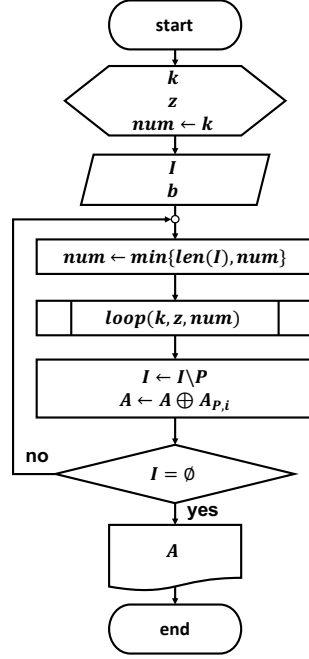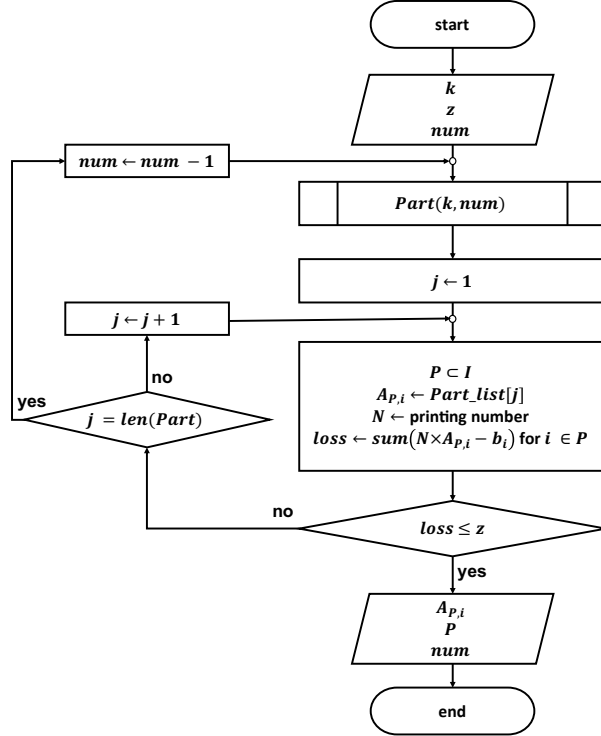
FIGURE 3.1. The Main Flowchart

FIGURE 3.2. The Flowchart of $loop(k, z, num)$

If the loss exceeds the threshold z, another *Part_list* will be selected and the foregoing will be repeated while adjusting num until the threshold z is not exceeded. As a result of this process, $A_{P,i}$ whose loss does not exceed the threshold is obtained.

4. EXAMPLES

This example was described to help the understanding of the problem. For the next two examples, we assume that k is equal to 4.

Example 4.1. Assume that $I = \{1, 2, 3\}$, and $\mathbf{b} = (50, 30, 20)$.

We consider $\pi = \{\{1, 2\}, \{3\}\}$, a partition that matches $\mathbf{b}$. Without loss of generality, assume that $P_1 = \{1, 2\}, P_2 = \{3\}$. Since $k = 4$, the matrix A can be found as follows.

$$(4.1) \qquad A = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

In this case, the printing numbers (3.1) of P_1 and P_2 are as follows, respectively.

$$(4.2) \qquad \left\lceil \max \left\{ \frac{b_i}{A_{P_1,i}} \,\middle|\, i \in P_1 \right\} \right\rceil = \left\lceil \max \left\{ \frac{50}{2}, \frac{30}{2} \right\} \right\rceil = 25$$

and

$$(4.3) \qquad \left\lceil \max \left\{ \frac{b_i}{A_{P_2,i}} \,\middle|\, i \in P_2 \right\} \right\rceil = \left\lceil \max \left\{ \frac{20}{4} \right\} \right\rceil = 5.$$
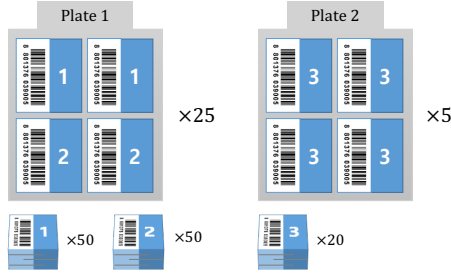
We can see the Figure 4.1 for more details.

FIGURE 4.1

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

Now, we consider a new partition $\pi = \{\{1, 2, 3\}\}$. In this case, $A = (\, 2 \quad 1 \quad 1 \,)$, and printing number (3.1) is

$$(4.4) \qquad \left\lceil \max \left\{ \frac{b_i}{A_{P,i}} \,\middle|\, i \in P \right\} \right\rceil = \left\lceil \max \left\{ \frac{50}{2}, \frac{30}{1}, \frac{20}{1} \right\} \right\rceil = 30.$$

In addition, it can be easily seen that the total number of Loss $E_{A,\mathbf{b}}$ is 20 (see Figure 4.2). The array shown in Figure 4.2 is more efficient since its Loss is the same but its number of plates is smaller.
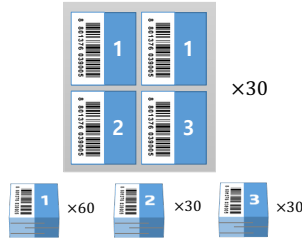

FIGURE 4.2

Example 4.2. Assume that $I = \{1, 2\}$, and $\mathbf{b} = (50, 20)$.

We consider $\pi = \{\{1, 2\}\}$, a partition that matches $\mathbf{b}$. Since $k = 4$, the matrix $A = (\ 2 \quad 2\ )$, and the printing number (3.1) is

$$(4.5) \qquad \left\lceil \max\left\{ \frac{b_i}{A_{P,i}} \,\middle|\, i \in P \right\} \right\rceil = \left\lceil \max\left\{ \frac{50}{2}, \frac{20}{2} \right\} \right\rceil = 25.$$

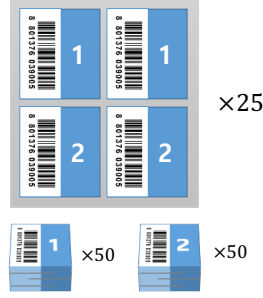The total number of Loss $E_{A,\mathbf{b}}$ is 30 (see Figure 4.3).

FIGURE 4.3

For the same partition π, matrix $A = (\ 3 \quad 1\ )$ can be considered. In this case, the printing number (3.1) is

$$(4.6) \qquad \left\lceil \max\left\{ \frac{b_i}{A_{P,i}} \,\middle|\, i \in P \right\} \right\rceil = \left\lceil \max\left\{ \frac{50}{3}, \frac{20}{1} \right\} \right\rceil = 20.$$

and the total number of Loss $E_{A,\mathbf{b}}$ is 10 (see Figure 4.4). The array shown in Figure 4.4 is more efficient because its number of plate is the same but fewer losses occur.
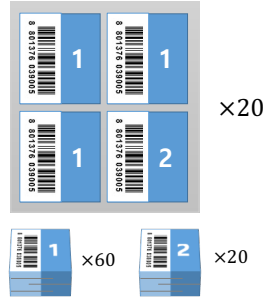

FIGURE 4.4

If there are too many products and large order-quantities, it requires a lot of iteration. Now, we consider real problem of World Komax. In this case, $k = 18$, $\alpha = 300$, $\beta = 1$. Table 4.1 is an actual sorting report of an order from World Komax.

As we can see from Table 4.1, there are 14 products and the largest order-quantity is 1059. It uses 3 plates and total Loss is 611. Total cost is 1511.

Example 4.3. We obtain another partition of the sorting report using our algorithms. Table 4.2 shows this result.

TABLE 4.1. the sorting report of World Komax

Plate	product	order-quantity	psc	printing-number	production	Loss
1	1	59	2	32	64	5
	2	156	5		160	4
	3	9	1		32	23
	4	162	6		192	30
	5	102	4		128	26
2	6	1059	5	228	1140	81
	7	886	4		912	26
	8	228	1		228	0
	9	832	4		912	80
	10	862	4		912	50
3	11	532	4	133	532	0
	12	532	5		665	133
	13	482	4		532	50
	14	582	5		665	103

TABLE 4.2. the result of our algorithm

Plate	product	order-quantity	psc	printing-number	production	Loss
1	7	886	5	178	890	4
	10	862	5		890	28
	11	532	3		534	2
	12	532	3		534	2
	4	162	1		178	16
	2	156	1		178	22
2	6	1059	9	118	1062	3
	14	562	5		590	28
	8	228	2		236	8
	5	102	1		118	16
	1	59	1		118	59
3	9	832	11	76	836	4
	13	482	7		532	50
4	3	9	18	1	18	9

The matrix A can be found as followed.

$$
\begin{array}{c}
\text{column index}\\
\begin{array}{cccccccccccccc}
7 & 10 & 11 & 12 & 4 & 2 & 6 & 14 & 8 & 5 & 1 & 9 & 13 & 3
\end{array}
\end{array}
$$

$$
A = \begin{pmatrix}
5 & 5 & 3 & 3 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 9 & 5 & 2 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 & 7 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 18
\end{pmatrix}
$$

The total Loss $E_{A,b}$ is 251. So, total cost is $\alpha|\pi| + \beta E_{A,b} = 300 \times 5 + 1 \times 251 = 1451$. We use one more Plate, but the total cost is reduced by 60.

5. RESULT

Each sorting report includes the number of losses corresponding to one plate. Using this, the total cost can be calculated by replacing each plate with loss. We

used 82 sorting report samples. The total cost was reduced by a minimum of -6.85%(sample no. 15) to a maximum of 27.5%(sample no. 74) (see Figure 5.1).
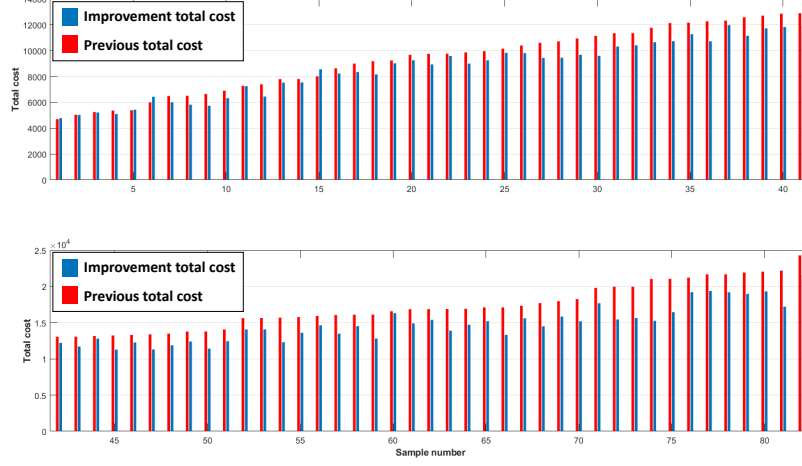

FIGURE 5.1. Comparing the Results

We used the *paired t-test* [8] to verify the efficiency of the algorithm. The paired t-test is one of the two sample t-tests, and it is a test that verifies whether the two groups are different. The two populations are as such.

- population1: total cost before applying the algorithm
- population2: total cost after applying the algorithm
- sample1: sample of 82 items from population1
- sample2: sample of 82 items from population2

In order to proceed with the two-sample t-test, the two groups have to satisfy normality and homoscedasticity. We choose 82 sample data from the two populations, which can satisfy normality by the *central limit theorem* [4]. 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 graphic.

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 5.2).

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

        F test to compare two variances

data:  sample1 and sample2
F = 1.5552, num df = 81, denom df = 81, p-value = 0.04847
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 1.002998 2.411501
sample estimates:
ratio of variances
         1.555227
```

FIGURE 5.2. 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 5.3).

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

        Paired t-test

data:  sample1 and sample2
t = 10.766, df = 81, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1240.884 1803.531
sample estimates:
mean of the differences
             1522.207
```

FIGURE 5.3. 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.

Finally, we check the efficiency over the number of products which is the size of I(cf. Section 3). We see that the efficiency increases as the number of products increases in Figure 5.4. For each 82 samples, we calculated the efficiency as the following formula.

$$(5.1) \qquad \text{Efficiency} = \left[1 - \frac{\text{Improvement Total Cost}}{\text{Previous Total Cost}} \right] \times 100$$

The result of the linear fitting shows that the efficiency and the number of products have a positive correlation. Thus, we see that the improved algorithms have better results in many products. This suggests that the algorithm fits well into small quantity batch productions.
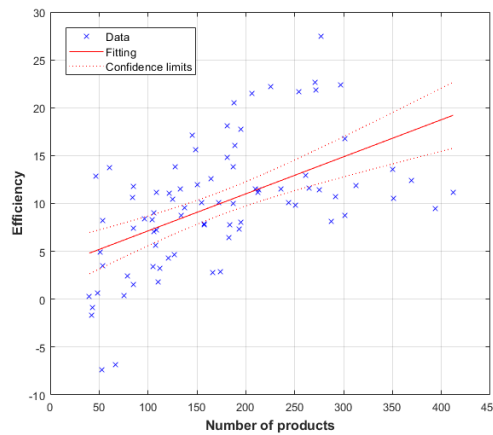

FIGURE 5.4. The linear fitting for the efficiency over the number of products

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

References

[1] R. A. Brualdi. *Introductory Combinatorics*. Pearson Prentice Hall, 2004.

[2] Albert W. Chan and P. W. Udo Graefe. An interactive sequencing aid for printing presses. *Computers and Industrial Engineering*, 3(4):321–325, 1979.

[3] Andriele Busatto do Carmo, Mateus Raeder, Thiago Nunes, Mariana Kolberg, and Luiz Gustavo Fernandes. A job profile oriented scheduling architecture for improving the throughput of industrial printing environments. *Computers and Industrial Engineering*, 88:191–205, 2015.

[4] Richard Durrett. *Probability: theory and examples*. Cambridge University Press, 4 edition, 2004.

[5] Encyclopædia Britannica, inc. "offset printing", June 2016. https://www.britannica.com/technology/offset-printing.

[6] H. Kipphan. *Handbook of Print Media: Technologies and Production Methods*. Springer, 2001.

[7] K. J. Musselman. Complex scheduling of a printing process. *Computers and Industrial Engineering*, 39(3–4):273–291, 2001.

[8] John A. Rice. *Mathematical Statistics and Data Analysis*. Duxbury Advanced, 3rd edition, 2006.

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

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

Hyun-Min Kim, Finance.Fishery.Manufacture Industrial Mathematics Center on Big Data, Pusan National University, Busan, 46241, Republic of Korea
Email address: hyunmin@pusan.ac.kr

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

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

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

Sang-hyup Seo, Finance.Fishery.Manufacture Industrial Mathematics Center on Big Data, Pusan National University, Busan, 46241, Republic of Korea
Email address: saibie1677@gmail.com

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