International Journal of Scientific & Engineering Research, Volume 7, Issue 4, April-2016 1165
ISSN 2229-5518

Comparative study of Pareto, Knapsack and Greedy Algorithm in the field of industrial Maintenance

Amal Boukili[1], Mohamed Fri[2], Mohammed El Hammoumi[3], Fouad Belmajdoub[4]

Abstract— In this paper, we have compared these three methods (Pareto, Knapsack problem (KP) and Greedy Algorithm) in order to deduce which among them is the most effective. The purpose of this study is to remedy halts and unforeseen breakdowns on chains of production as well as to increase the efficiency of production, which is the major concern of the industry leaders. The choice of method of maintenance management directly affects the stability of the productivity and the cost of maintenance actions. We will compare the results obtained from these three methods: Pareto, which is widely used in the maintenance; Knapsack Problem, which is widely used in Freight management; and Greedy Algorithm, which is used in scheduling.

We will apply these three methods to an actual case study in order to improve industrial maintenance structure in the production of soft drinks. The synthesis will be based on two variables: the number of failure times for each machine for two months and the budget to repair them.

We have shown in this document that the Greedy algorithm for scheduling gave beneficial results of 3.50% more than those obtained using the Knapsack problem (KP) for the management of transportation of the goods. In the same perceptual using, the latter gave results, which advance 8% vis-a-vis the use of Pareto method that is specific to industrial maintenance. Our future goal is to design a new method that incorporates the most effective elements from each of these three existing methods.

Index Terms— Pareto, Knapsack Problem, Greedy Algorithm, science of decision, Industrial Maintenance.

—————————— ◆ ——————————

1. INTRODUCTION

RECENT studies on the effectiveness of maintenance management have shown that more than a third of maintenance costs come from unnecessary or poorly executed transactions. Poor maintenance policy has disastrous consequences on the quality of products. The main reason for this inefficiency is the absence of real information that would identify the immediate needs for repair or maintenance.

Maintenance costs often represent the bulk of operating costs in a number of production units. These costs can be significantly reduced by making more suitable decisions. The choice of method of maintenance management directly influences the rate of profitability and efficiency, it should therefore be very important to prepare the methods to ensure that management.

Among these methods, we find AMDEC, 5S, ….. These methods are based on the opinions of experts and therefore are not based on real scientific calculations. However, there are other methods that do rely on precise calculations which we focus on instead. First, we find "Pareto analysis": this method is widely used in the classification of equipment in order to determine which machines have the most critical issues. But we strive to find a classification system that gives better results than those of the Pareto analysis [3][4][5]. To do so, we thought to use two other methods: Knapsack Problem [1],[2],[6],[7],[8] which is widely used in Freight management and Greedy Algorithm [9],[10],[11],[12] which is used in scheduling. Since neither of the latter two methods have ever been used in research on maintenance, we cannot base our research on any previous findings, so we aim to create a new method that incorporates elements from all three.

The comparison of these three methods, contributes to making optimal decision on the choice of equipment. Then one can determine the most critical equipment, and therefore, to achieve optimum operation of the budget for improving the production chain in order to minimize the failure rate.

In this context, a study was made in one of the companies that makes the production of soft drinks, which allowed us to apply these three methods based on two variables : numbers of hours of downtime and the cost for each equipment in the hope of improving industrial maintenance, productivity and management of the company budget, actual measurements were analyzed by these three methods leading to satisfying results.

After identifying various critical equipment, it will be easy to identify the different categories of anomalies, focus on the most disastrous, and assign the appropriate response in the context of preventive maintenance.

This article describes the approach to meet our goal. It is structured as follows:

The first part will be devoted to the general presentation of

————————————————

- *1, 2, 3, 4: University of Sidi Mohamed Ben Abdellah, Faculty of Science and Technology, Laboratoire Techniques Industrielles, Road Imouzzer B.P. 2202 Fez, Morocco*
- *1: Amal-boukili@usmba.ac.ma*
- *2: Mohamed.fri@usmba.ac.ma*
- *3: m_elhammoumi@yahoo.fr*
- *4: fbelmajdoub@yahoo.fr*

Pareto analysis, while the second part will focus on the general format of the bag of bones problem. In part, three we will give an overview on the greedy algorithm. In the fourth will show a case study and the end we will give the general rule.

2. The Pareto (ABC method)

Without hierarchisation, all actions of organization can be long and tedious. By using the Pareto law [3][4][5] we can highlight the most important elements of a problem to guide our action. Because of this, the elements having little influence on the criterion studied will be eliminated. The ABC method is a tool for decision support, which defines priorities for actions. This means that the Pareto chart shows the most important causes that are causing most of effects.

The elements will be ranked by order of importance indicating the percentages for a given criterion. This study requires a three steps approach:

• Defininge the nature of the elements to be classified: the classification of these elements depends on the criteria studied.

These elements can be physical, causes of failures, types of failures, work orders, items in stock, etc…

• Choosing the classification criterion: The most common criteria are costs and time, according to the character studied, other criteria can be used, including: The number of accidents, the number of incidents, the number of rejects, the number of operating hours, the number of kilometers covered, annually consumed value which is often necessary for the management of stocks, etc..

• Defininge the limits of the study and classify the elements.

The Pareto chart is a column chart that presents information in descending order and thus brings out the most important elements, which explain a phenomenon or situation. Generally, 20% of the number of elements represents 80% of the criteria: it is the class A; 30% on the number of elements represents 15% of the criteria studied: it is the class B; and the remaining 50% of elements represents only 5% of the criterion studied: it is the class C.

By cumulating the decreasing values of the criterion studied, the ABC curve shows three zones, reason why it "ABC curve", see “Fig.1”.

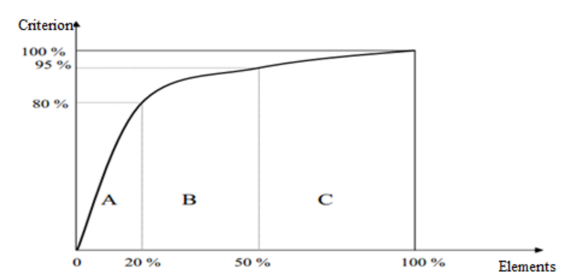

Fig. 1. Pareto ABC

3. Knapsack Problem

The knapsack Problem (KP) or rucksack problem is a problem of combinatorial optimization: Given a set of elements, with a mass and a value each, it determines the element to include in a collection so that the total weight is less than or equal to a given limit and the total value is as large as possible. It derives its name from the problem faced by someone who is constrained by a fixed-size knapsack and must fill it with the most valuable elements [1],[2],[6],[7],[8] then the parameters may be the volume of the bag or container and the value or price.

The data of the problem can be expressed in mathematical terms. Objects are numbered by index i varying from 1 to n. Numbers Wi and Pi are respectively the weight and the value of the object numbered i. The capacity of the bag will be noted W.

There are many different ways to complete the Knapsack. To describe one of them must be indicated for every element whether it is taken or not. We can use a binary coding: the state of the element i will have the value xi = 1 if the element is in the bag, or xi = 0 if it is left out. A way of filling the bag is completely described by a vector called vector content, or simply content: X = (x1, x2... xn), and the associated weight and value this filling can then be expressed as a function of the vector content.

For a given content X, the total value in the bag is naturally:

$$Z(X) = \sum_{\{i, x_i = 1\}} P_i = \sum_{i=1}^{n} x_i P_i$$

Similarly, the sum of the weights of selected objects is:

$$W(X) = \sum_{\{i, x_i = 1\}} W_i = \sum_{i=1}^{n} x_i W_i$$

The problem can then be reformulated as the search for a content vector X = (x1, x2, ...,xn) (which components have the value 0 or 1), achieving the maximum total value function Z(X) under duress (1) :

$$W(X) = \sum_{i=1}^{n} x_i W_i \le W \qquad (1)$$

This is to say that the sum of the weights of objects selected does not exceed the capacity of the Knapsack.

In general, the following constraints are added to avoid singular cases:

$\sum W_i > W$: We can not put all the objects ;

$W_i \le W, \forall i \in \{1, \ldots, n\}$: no object is heavier than the bag can carry ;

$P_i > 0, \forall i \in \{1, \ldots, n\}$: any object has a value and brings a gain ;

$W_i > 0, \forall i \in \{1, \ldots, n\}$: all objects have a certain weight and consumes resources ;

Terminology:

Z(X) : is called objective function;

Every vector X satisfying the constraint (1) is said to be feasible;

If the value of Z(X) is maximum, then X is said optimal.

4. 4 Greedy Algorithm

As for most decision problems, it may be enough to find workable solutions even if they are not optimal. Preferably, however, the approximation comes with a guarantee on the difference between the value of the solution found and the value of optimal solution.

The terminology adopted is "Efficiency of an object" which is the ratio of its value over its weight. If the value of the object is large compared to what they consume, then the object is more efficient.

The idea of greedy algorithm as illustrated in “Fig.2” is to add in priority the most effective objects until the saturation of the bag [9],[10],[11],[12] :

1- Sort the objects in decreasing order of effectiveness
2- w_conso: = 0
3- for i = 1 to n
4 - if w [i] + w_conso ≤ W then
5 - x [i]: = 1
6 - w_conso: = w_conso + w [i]
7 - else
8 - x [i]: = 0
9 - end if

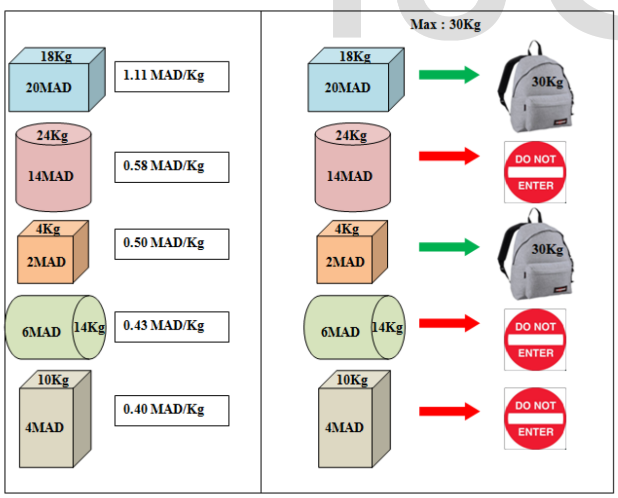

Fig. 2. Illustration of the problem of the Knapsack Problem solved using the Greedy Algorithm

The two phases of the greedy algorithm are presented. Left: sorting of the boxes in order of interest (here in MAD per kilogram). Right: insertion in the order boxes, if possible. We get in this example a solution of 22 MAD for 22 kg.

5. Case study.

5.1. Introduction

All enterprises have a lucrative purpose that is to say "producing more" and consequently downtime must be minimized; in order to make this, enterprises reserve the budgets allocated to improve their productivities. In our article, we look at a case study of a packaging of soft drinks. At first, we begin with a case study where we will use simple problem analysis tool “Pareto”, then we will integrate the knapsack Problem to the problem of Pareto and finally the Greedy algorithm.

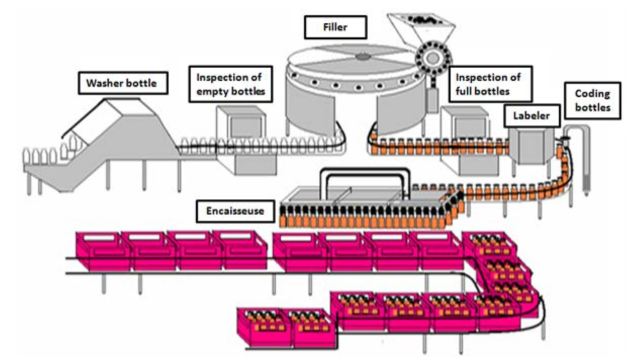

Fig. 3. Schematic illustrative of the different steps of the preparation of soft drinks

To improve the efficiency of a production line (called line 2) and maintain the majority of its equipments in good condition during production “Fig.3”, a budget of 300,000.00 MAD is proposed. To do this we will study the downtime and the maintenance costs during 2 months for each machine on line 2, given in the following table (Table I):

The elements of the line 2	Downtime (h)	The cost of maintenance (MAD)
FILLER O + H L2	7,24	72400
VISSEUSE L2	6,73	67300
CONVEYOR BOTTLE L2	6,02	60400
ENCAISSEUSE KETTNER L2	5,59	56400
CAPPING L2	4,47	44300
CONVEYOR CASIERS L2	4,14	41900
PALLETIZER L2	3,73	18000
WASHER BOTTLES O + H L2	3,3	33100
LABEL KRONES L2	2,41	24200
DECRATER KETTNER L2	2,4	24300
DEPALETISOR L2	1,51	15500
INSPECTOR L2	1,12	11900
DATEUSE L2	1,03	10900
MIXER L2	0,26	3300
Total	**49,95**	**483900**

Fig. 4. The cost of maintenance and downtime for each machine of line 2

The method Pareto consists in classifying machines in order of severity which is calculated by (downtime of the machine / Total downtimes)*100

The table below (see Table II) represents the percentage of downtime for each machine of the line 2 during two months:

ISSN 2229-5518

Machine	Downtime (h)	Cumulative (h)	Downtime (%)	Cumulative (%)
FILLER O + H L2	7,24	7,24	14,49	14,49
VISSEUSE L2	6,73	13,97	13,48	27,97
CONVEYOR BOTTLE L2	6,02	19,99	12,05	40,02
ENCAISSEUSE KETTNER L2	5,59	25,58	11,19	51,21
CAPPING L2	4,47	30,05	8,96	60,17
CONVEYOR CASIERS L2	4,14	34,19	8,29	68,46
PALLETIZER L2	3,73	37,92	7,47	75,93
WASHER BOTTLES O + H L2	3,3	41,22	6,61	82,54
LABEL KRONES L2	2,41	43,63	4,82	87,36
DECRATER KETTNER L2	2,4	46,03	4,80	92,16
DEPALETISOR L2	1,51	47,54	3,02	95,18
INSPECTOR L2	1,12	48,66	2,24	97,42
DATEUSE L2	1,03	49,69	2,06	99,48
MIXER L2	0,26	49,95	0,52	100,00

Fig. 5. Percentage of the breakdown of each machine of the line 2

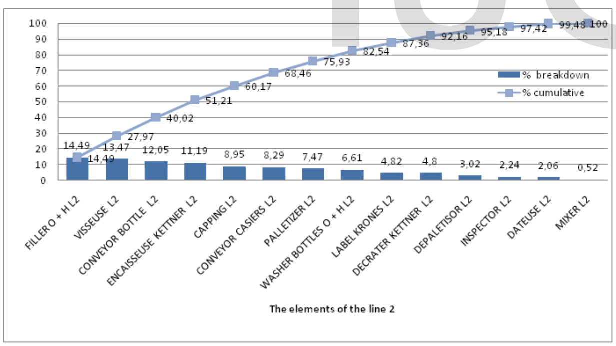

Fig. 6. Machines which cause a high percentage of discontinuation of production in line 2

According to the Pareto diagram, we find that 82.52% of the problems, which cause the stopping of line 2, are due to stopping of the FILLER, the NUTRUNNER, the CONVEYOR BOTTLE, the ENCAISSEUSE, the CAPPER, the CONVEYOR CASIERS, the PALLETIZER and the WASHER BOTTLES, provoking downtime taking a sizeable proportion of working time and consequently stopping the production.

For the budget (300 000.00MAD), we note that with the Pareto method we can address the problems of the following machines:

- FILLER O+H L2
- NUTRUNNER L2
- CONVOYOR BOTTLE L2
- ENCAISSEUSE KETTNER LV2

This solution allows us to minimize to 51.2% of downtimes with an amount of 258 500.00 MAD, but the arising questions are:

- Is it the most optimal solution?
- Can we exploit the rest of the budget to get a better solution than this?

To answer these questions, we will use the knapsack problem.

5.3. Resolution by the method of Knapsack Problem

The method of Knapsack Problem consists on putting objects in a bag without exceeding its capacity, until the saturation of the knapsack, if the object i is in the bag we have $x_i = 1$, else $x_i = 0$.

The application of the method of Knapsack Problem is illustrated in the following table (Table III):

Machine	The cost of maintenance (MAD)	x_i
FILLER O + H L2	72400	1
NUTRUNNER L2	67300	1
CONVEYOR BOTTLE L2	60400	1
ENCAISSEUSE KETTNER L2	56400	1
CAPPER L2	44300	0
CONVEYOR CASIERS L2	41900	1
PALLETIZER L2	18000	0
WASHER BOTTLES O + H L2	33100	0
LABEL KRONES L2	24200	0
DECRATER KETTNER L2	24300	0
DEPALETISOR L2	15500	0
INSPECTOR L2	11900	0
DATEUSE L2	10900	0
MIXER L2	3300	0

Fig. 7. Results of the application of the Knapsack Problem

From the results of Knapsack Problem, we see that we can solve the problems of the following machines:

- FILLER O+H L2
- NUTRUNNER L2
- CONVOYOR BOTTLE L2
- ENCAISSEUSE KETTNER LV2
- CONVEYOR CASIERS L2

This solution allows us to minimize up to 59.49% of downtime with an amount of 298,400 .00MAD, but, does it exist any other more optimal solution?

- To answer this question, we will use the Greedy algorithm in order to compare the results to better exploit the budget by minimizing downtime in line 2.

5.4. Resolution by the Greedy Algorithm

Greedy algorithm makes the classification of objects by the efficiency; calculated by dividing the cost of maintenance by the downtime. The choice of machines to be corrected is made through the method of filling the knapsack; the use of this algorithm in our case study provides results that are illustrated in the following table (Table IV):

Machine	Downtime (h)	The cost of maintenance (KMAD)	Efficiency (KMAD/h)	x_i
MIXER L2	0,26	3,3	12,69	1
INSPECTOR L2	1,12	11,9	10,63	1
DATEUSE LV2	1,03	10,9	10,58	1
DEPALETISOR L2	1,51	15,5	10,26	1
DECRATER KETTNER L2	2,4	24,3	10,13	1
CONVOYOR CASIERS L2	4,14	41,9	10,12	1
ENCAISSEUSE KETTNER L2	5,59	56,4	10,09	1
LABEL KRONES L2	2,41	24,2	10,04	1
CONVEYOR BOTTLE L2	6,02	60,4	10,03	1
WASHER BOTTLES O + H L2	3,3	33,1	10,03	1
FILLER O + H L2	7,24	72,4	10,00	0
NUTRUNNER L2	6,73	67,3	10,00	0
CAPPER L2	4,47	44,3	9,91	0
PALLETIZER L2	3,73	18	4,83	1

Fig. 8. Application results of the Greedy algorithm

After the application of the Greedy algorithm, we find that we can solve the problems of the following machines:

- MIXER L2
- INSPECTOR L2
- DATEUSE L2
- DEPALETISOR L2
- DECRATER KETTER L2
- CONVOYOR CASIERS L2
- ENCAISSEUSE KETTER L2
- LABEL KRONES L2
- CONVOYOR BOTTLE L2
- WASHER BOTTLES O+H L2
- PALLETIZER L2

This solution allows us to minimize over 63.07% of downtimes with an amount of 299 900 .00 MAD.

The results of three tools: Pareto, Knapsack Problem and Greedy algorithm are shown in the following table (Table V):

	downtime corrected (%)	The cost of intervention (MAD)	Value remedied in 2 months (MAD)
Pareto	51.20	258500	92160
Knapsack Problem	59.49	298400	107082
Greedy algorithm	63.07	299900	113526

Fig. 9. The percentage of downtime corrected and the cost of intervention for the four tools

The comparison of downtime percentage corrected by Knapsack Problem compared to Pareto; Greedy Algorithm compared to Pareto and Knapsack Problem is shown in the following table (Table VI):

	Pareto	Knapsack Problem	Greedy algorithm
Knapsack Problem	8.29%	-	-
Greedy algorithm	11.87%	3.58%	-

Fig. 10. The percentage of downtime corrected for each tool

The application of Greedy Algorithm provides the most optimal result compared to the other tools, and therefore the gains in MAD from Greedy Algorithm compared to other tools are shown in the following graph “Fig.5”:

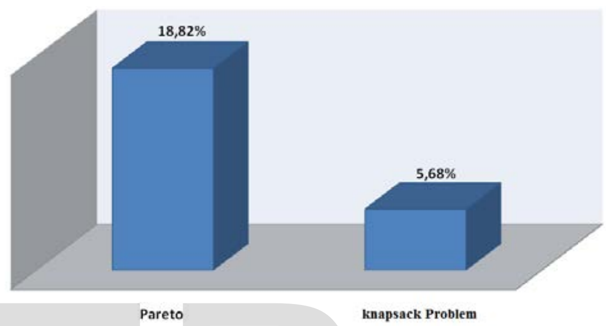

Fig. 11.Quantification of monetary gains of our methodology compared to the other three tools

6. GENERAL RULE

In our article, we showed that the greedy algorithm (GA) gives a more optimal result than knapsack problem (KP) and the latter gives a more optimal result than the method of Pareto (P).

Result of our case study: GA>KP>P

However, after several tests on different cases we found the following results:

Case 1 : GA>KP>P
Case 2 : GA=KP>P
Case 3 : GA>KP=P
Case 4 : GA=KP=P

Then these results can be generalized to the following equation:

$$GA \geq KP \geq P$$

This conclusion can be applied only if we have two variables (in our case study number of hours of breakdown and cost of intervention). For the case of one can use the methods of the back and Pareto bag problem, but in the case where the variable number is greater than 2, it must find another method that take into consideration three variables or use the three tools with all cases and compare all results to find the most optimal solution. This solution is not practical, that is why we are planning to develop a new tool that will consider three or more variables.

7. Conclusion

In our article, we showed that the Greedy Algorithm is intended to give the best results in industrial maintenance with regards to failures of machines and repairs to these machines. The application of these findings will improve productivity and conditions for maintenance in the soft-drink industry. Specifically, this method can help to optimize budgets by minimizing the amount of time that machines remain unfixed, minimizing the amount of money spent on the maintenance of these machines, and by maximizing the rate of production.

Application of the Greedy algorithm is based on two variables and gives more optimal results than traditional methods. However, if we want to consider additional conditions (eg agents, equipment ...), we will have to develop a method that is capable of analyzing three or more variables.

References

[1] H. Kellerer, U. Pferschy, D. Pisinger, "Knapsack problems," Spinger, 2004.

[2] M. Hifi, H. Mhallah, S. Sadfi, "Adaptive algorithms for the Knapsack problem," European Journal of Industrial Engineering, 2 (2), 2008, pp. 134-152.

[3] C. Brooks," What Is a Pareto Analysis? ", Business News Daily Senior, March 29, 2014

[4] R. Cirillo, "The Economics of Vilfredo Pareto", Routledge, 12 nov. 2012 - 148 pages

[5] M. Doostparast, Narayanaswamy Balakrishnan, "Pareto analysis based on records", A Journal of Theoretical and Applied Statistics,Volume 47, Issue 5, 2013, pages 1075-1089

[6] X. Song, R. Lewis, J. Thompson, and Y. Wu, "An incomplete m-exchange algorithm for solving the large-scale multi-scenario knapsack problem," Computers & Operations Research, vol. 39, no. 9, pp. 1988–2000, 2012.

[7] P.Sharafi, L. H. The, M. N. S. Hadi, "Conceptual design optimization of rectilinear building frames: A knapsack problem approach", Engineering Optimisation, Taylor& Francis, DOI: 10.1080/0305215X.2014.963068, 2014.

[8] F. Furini, M. Iori, S. Martello, M. Yagiura, "Heuristic and exact algorithms for the interval min-max regret knapsack problem", INFORMS Journal on Computing, 2015. INFORMS. DOI:10.1287/ijoc.2014.0632

[9] B. Diop, D. Diongue, O. Thiaré, "Greedy Algorithms for Target Coverage Lifetime Management Problem in Wireless Sensor Networks", International Journal of Control and Automation Vol. 8, No. 2, pp. 232-250, 2015

[10] Z. Zou, Z. Li, S. Shen, R. Wang, "Energy-Efficient Data Recovery via Greedy Algorithm for Wireless Sensor Networks", International Journal of Distributed Sensor Networks, Article ID 856764, 2015.

[11] L. Turner, M. Ehrgott, H. W. Hamacher; "On the generality of the greedy algorithm for solving matroid base problems", Discrete Applied Mathematics, 2014

[12] F. Glover, "Advanced greedy algorithms and surrogate constraint methods for linear and quadratic knapsack and covering problems", European Journal of Operational Research, Volume 230, Issue 2, 16 October 2013, Pages 212-225.