

# Balanced Constrained Partitioning of Distinct Objects

Joseph Varghese Kureethara, Jayanta Biswas, Debabrata Samanta, Naived George Eapen

**Abstract**— Partition is the process of dividing  $\eta$  elements of a set into  $v$  non-empty subsets where every element is included in exactly one subset. Set partitioning problems play an important role in real-world scenarios such as scheduling, management and grouping. In this paper,  $n$  distinct objects are arranged into different groups and the problem of partition is evaluated concerning a people assignment problem of a multi-site company. New constraints such as minimum and the maximum number of people in a single-site company (respectively  $m$  and  $M$ ) that are fixed for all sites are introduced and the solutions for the partition problem for the constraints are also established.

## I. INTRODUCTION

Partitioning is the common initial step that is performed in any real-life distributed problem-solving. The occurrence of the requirement for dividing a number of distinct objects into a fixed number of groups is frequent. The sizes of each group have to be either perfectly balanced or bordering to that to treat all the groups in an unbiased manner. To produce a balanced approach in the partitioning for optimal outcome, we introduce constraints such as minimum and the maximum number of objects that a group could hold. The first and second kinds of Stirling Numbers have been highly studied with respect to their mathematical properties that were discovered independently and repeatedly. Accompanying the related applications in various domains such as combinatorial, probabilistic and statistical, several extensions and refinements were suggested during the last decade. A Stirling number of the second kind, denoted as  $S(\eta, v)$  is the number of ways a set of  $\eta$  elements can be partitioned into  $v$  non-empty sets. The applications of set partitioning problems are very vast. Many real-life problems can be modeled as set partitioning problems. A larger set of values can be divided into smaller groups in many ways and this concept has very typical applications in day-to-day life which are more or less found in grouping and scheduling activities. Algorithms can be developed for typical scenarios and randomly generated values could be used for testing efficiency. Crew Scheduling, Vehicle Routing, Faculty Location Problems, Steiner Problem, Minimum Test Set Problem, etc. are some of the stock examples for set partitioning problems [1].

**Revised Manuscript Received on September 10, 2019.**

**Joseph Varghese Kureethara**, Professor of Mathematics, Christ (Deemed to be University) Bengaluru, Karnataka, India.  
(Email: frjoseph@christuniversity.in)

**Jayanta Biswas**, Department of Computer Science, Christ (Deemed to be University) Bengaluru, Karnataka, India.  
(Email: jayantab2002@gmail.com)

**Debabrata Samanta**, Department of Computer Science, Christ (Deemed to be University) Bengaluru, Karnataka, India.  
(Email: debabrata.samanta369@gmail.com)

**Naived George Eapen**, Department of Computer Science, Christ (Deemed to be University) Bengaluru, Karnataka, India.

## II. RELATED WORK

C. Wu, E. Kamar, and E. Horvitz [2] has introduced the concept of Clustering for set partitioning through a case study in carpooling. In their work, an interconnection between clustering and set partitioning is proposed with the objective to solve any set partitioning problem through the clustering approach. The K-set partitioning problem has benefited tremendously as it is considered as centroid-based clustering problem in the paper. A case study on carpooling of 100 people with a single destination source is evaluated with the results of the study[2]. M. Chen and H. Niu [3] has considered a case of integrated vehicle and crew scheduling in a multiple-depot environment. In their study, two models along with algorithms are established based on a combination of column generation and Lagrangian relaxation. Each vehicle and crew-member is exactly grouped into a depot which is modeled as a set-partition problem. Further, the grouping information is used for optimal scheduling purposes and a simulation of an urban bus network is carried out. [3] The study of H. D. Chu, E. Gelman, and E. L. Johnson [4] were based on a large scale crew scheduling problem. A graph-based branching heuristic is applied to a restricted set partitioning problem representing the ideal pairings. The associated problem has diverged into sub linear programming problems that have integer solutions and the set partitioning problem is solved by modeling the crew pairing problem as a problem of finding a set of node-disjoint directed cycles spanning the nodes of a graph.[4] A. Mingozzi, M. Boschetti, L. Bianco, and S. Ricciardell [5] discussed another set partitioning approach to the crew scheduling problem where a branch-and-bound algorithm is used to solve the set partitioning problem after minimizing the number of variables as a result of several bounding procedures. The minimum number of ongoing duties is always a constraint to a minimum value  $m$ . A variable reduction method is used to minimize the size of the Set Crew and the optimal solution is found by replacing the original set  $R$  with  $\bar{R}$ [5]. It can be seen that in any of the above-mentioned works, there are no constraints for the minimum or maximum number of elements in a subdivided group. However, it is necessary to have a uniform distribution (balanced partitioning) of objects in the group. In this paper, we propose the approach to solve a set partitioning problem with the above-mentioned constraints applied.



III. METHODS & RESULTS

A partition of  $\eta$  objects into  $\nu$  groups is one of the possible ways of subdividing the  $\eta$  objects into  $\nu$  groups ( $\nu \leq \eta$ ). This approach has two rules,

1. The objects could be assigned in any order
2. Only one group can be assigned per object

Let  $a_1, a_2, \dots, a_\eta$  be  $\eta$  objects and  $g_1, g_2, \dots, g_\nu$  be  $\nu$  (with  $\nu \leq \eta$ ) groups to which  $\eta$  objects are assigned.  $\eta_1$  objects can be assigned to group  $g_1$ ,  $\eta_2$  objects can be assigned to group  $g_2$  and so on.

Let  $P_{\eta_1, \eta_2, \dots, \eta_\nu}$  denote the number of possible partitions into the  $\nu$  groups (where group  $i$  contains  $\eta_i$  objects). Then,

$$P_{\eta_1, \eta_2, \dots, \eta_\nu} = \frac{\eta!}{\eta_1! \eta_2! \dots \eta_\nu!}$$

Even though (1) gives the total number of partitions into  $\nu$  groups, it lacks the uniformity or balance in distribution. It is possible that one group may contain only 1 element and the other contains any  $\eta_j$  number of elements. The RHS of the (1) could be represented as a

multinomial coefficient,  $\binom{\eta}{\eta_1, \eta_2, \dots, \eta_\nu}$ . A multinomial coefficient could be expanded as :

$$\binom{\eta}{\eta_1, \eta_2, \dots, \eta_\nu} \cdot \kappa^{\eta_1} \cdot \kappa^{\eta_2} \dots \kappa^{\eta_\nu} = \sum_0 \binom{\eta}{\eta_1, \eta_2, \dots, \eta_\nu} \cdot \kappa^{\eta_1} \cdot \kappa^{\eta_2} \dots \kappa^{\eta_\nu} \tag{2}$$

where the total number of objects,  $\eta = \eta_1 + \eta_2 + \dots + \eta_\nu$ . [6]

3.1 Introducing new constraints on number of variables

When grouping a number of objects if lower limit and upper limit is not established beforehand, the partition would go unbalanced. In the above method, we could calculate the total number of possibilities of dividing  $\eta$  objects into  $\nu$  subsets. Let  $N$  denote the number of possible partitions into the  $\nu$  groups where the number of objects in a group are constrained. However, this requires us to know the exact number of elements in each subset. In most of the real life scenarios, we will not be having any control over defining number of elements in each subset. Moreover, in a scenario where there are a large number of objects, it is impossible to define the number of elements for each subset. For instance, consider a set of 1000 employees that has to be partitioned into 100 sets. In the previous approach, one has to define the number of employees in all the sets from 1 to 100, which is not practical. Therefore in order to have balanced partitioning over a set of constraints (say 100 sub groups of 8 to 12 employees), we propose the method below.

3.2 Scenario 1: The minimum number of objects in a group is fixed at 'm'

Consider  $\eta$  distinct objects and  $\nu$  groups. Each group must have at least  $m$  objects. Thus,

$$N = [\nu C_m^{\eta-m} C_m^{\eta-2m} C_m \dots \eta^{-(\nu-1)m} C_m] \cdot \omega \tag{3}$$

Each set of  $m$  distinct objects selected goes to one group, and this process repeats  $\nu$  times. Now there are  $(\eta - \nu m)$

objects remaining; each of these objects may go to any one of the  $\nu$  groups.

$\therefore$  Each object has  $\nu$  distinct choices,

$$\omega = \nu \cdot \nu \cdot \nu \dots \nu \quad (\eta - \nu m \text{ times}) \\ = \nu^{(\eta - \nu m)}$$

Now, applying the binomial formula, (3) becomes,

$$(4) \quad N = \frac{\eta!}{(m!)^\nu (\eta - (\nu - 1)m)!} \cdot \nu^{(\eta - \nu m)}$$

3.3 Scenario 2 : The maximum number of objects in a group is fixed at 'M'

Let  $\eta$  be the total number of distinct objects,  $M$  be the maximum number of objects in a group and  $\nu$  be the total number of groups.

$N =$  Coefficient of  $\kappa^\eta$  in

$$\begin{aligned} (\kappa^1 + \kappa^2 + \dots + \kappa^M)^\nu &= \kappa^\nu \cdot \left( \frac{1 - \kappa^M}{1 - \kappa} \right)^\nu \\ &= \kappa^\nu \cdot \frac{(1 - \kappa^M)^\nu}{(1 - \kappa)^\nu} \\ &= \kappa^\nu (1 - \kappa)^{-\nu} (1 - \kappa^M)^\nu \end{aligned} \tag{5}$$

IV. ILLUSTRATED EXAMPLES: MULTI-SITE COMPANY EMPLOYEE ASSIGNMENT PROBLEM

4.1 Description

Employee assignment with respect to a multi-site company accommodates the process of allocating a number of employees to several locations (sites).

4.2 Formulating it as Balanced Constrained Partitioning of Distinct Objects

The primary assumptions and requirements are:

- The employees are of the same level (the order of distribution do not matter)
- There are  $\eta$  employees and  $\nu$  locations
- The minimum and maximum number of employees that are required in each locations are  $m$  and  $M$  respectively
- No employees can work in two locations simultaneously
- The locations are of uniform size and there are no priority assigned to any of the locations in terms of the required number of employees

V. FUNDAMENTAL EXAMPLES

5.1 Scenario 1 : The minimum number of objects in a group is fixed at 'm'

Consider a case where 5 employees have to be divided into two groups where maximum number of employees in a group are constrained to 3.

That is,  $\eta = 5, \nu = 2$  and  $m = 3$ .

We know that the number of possibilities will be  $\binom{5}{3, 2}$  which is the same as  ${}^5C_3 = 10$



From equation (I) of scenario 1 (2.2.1), the total number of possibilities will be:

$$\begin{aligned}
 N &= \frac{6!}{(3!) (6 - (2 - 1) 3)!} \cdot 2^{(6 - 2 \times 3)} \\
 &= \frac{720 \times 1}{6 \times (6)} \\
 &= \frac{720}{36} \\
 &= 20 \qquad (6)
 \end{aligned}$$

## VI. CONCLUSION

Balanced constrained partition of distinct objects play an important role in many applications. The earlier approach to find the total number of groups involved addition, multiplication and division of several terms that involve factorials. The newly proposed method in our study eliminates the complexity of calculations...

## VII. REFERENCES

1. E. El-Darzi and G. Mitra, "Set covering and set partitioning: A collection of test problems," *Omega*, vol. 18, no. 2, pp. 195 – 201, 1990.4
2. C. Wu, E. Kamar, and E. Horvitz, "Clustering for set partitioning: A case study in carpooling," 2016.
3. M. Chen and H. Niu, "A model for bus crew scheduling problem with multiple duty types," *Discrete Dynamics in Nature and Society*, vol. 2012, 09 2012.
4. H. D. Chu, E. Gelman, and E. L. Johnson, "Solving large scale crew scheduling problems," *European Journal of Operational Research*, vol. 97, no. 2, pp. 260 – 268, 1997.
5. Mingozzi, M. Boschetti, L. Bianco, and S. Ricciardelli, "A set partitioning approach to the crew scheduling problem," *OPERATIONS RESEARCH*, vol. 47, pp. 883–888, 12 1999.
6. M. Taboga, *Lectures on Probability and Statistics*, 2010.
7. Jossy P. George, Joseph Varghese Kureethara, "Anefficient 2-Step DNA symmetric cryptography algorithm based on dynamic data structures," *International Journal of Engineering & Technology*, 7 (2.6) (2018) 141-146.
8. Dulal Chandra Samanta, Debabrata Samanta, "The Magic Square from Myth to Mystery", *International Journal of Engineering Research and Development (IJERD)* pp. 30-38, July 2012 issue, ISSN: 2278-067X.
9. Dulal Chandra Samanta, Debabrata Samanta, "Human Computer Calendar for ZERO to INFINITE", *IOSR Journal of Mathematics (IOSR-JM)* pp. 13-15, Volume 1, Issue 1 (May-June 2012), ISSN: 2278-5728.
10. Syed K A Khadri,, Debabrata Samanta,M Paul, "Message Encryption Using Pascal Triangle Multiplication: in Cryptology," *Asian Journal of Mathematics and Computer Research*, [ISSN: 2395-4205 (Print), 2395-4213 (Online)], Vol 13(4): 262-270, 2016.
11. Ranjini K.S., Debabrata Samanta, *Applied Remedial Mathematics*, Publishing house: Lambert Academic Publishing, Germany, ISBN-13: 978-613-9-87021-9, July, 2018.