SCIENTIFIC RESEARCH 2016

Proceedings of articles the international scientific conference
Czech Republic, Karlovy Vary - Russia, Moscow, 29-30 September 2016

SCIENTIFIC RESEARCH - 2016

Proceedings of articles the international scientific conference

Czech Republic, Karlovy Vary - Russia, Moscow, 29-30 September 2016

Czech Republic, Karlovy Vary - Russia, Kirov, 2016

UDC 001
BBK 72
N 76

Scientific editors:

Barmashov Konstantin Sergeevich, PHD, Head of the Department "Innovation management and social entrepreneurship" Russian Economic University named G.V. Plehanov

Dalgat Fatima Magomedovna, Ph.D., Associate Professor of "History of Russia of XX - XXI centuries." Dagestan State University

Ivanova Galina Viktorovna, Ph.D., assistant professor of "Elementary and preschool education" State Socio-Humanitarian University

N 76 SCIENTIFIC RESEARCH – 2016: Proceedings of articles the international scientific conference. Czech Republic, Karlovy Vary - Russia, Moscow, 29-30 September 2016 [Electronic resource] / Editors K.S.Barmashov, F.M.Dalgat, G.V.Ivanova. – Electron. txt. d. (1 файл 4.7 MB). – Czech Republic, Karlovy Vary: Skleněný Můstek – Russia, Kirov: MCNIP, 2016. - ISBN 978-80-7534-083-2 + ISBN 978-5-00090-107-6.

Proceedings includes materials of the international scientific conference « SCIENTIFIC RESEARCH – 2016», held in Czech Republic, Karlovy Vary-Russia, Moscow, 29-30 September 2016. The main objective of the conference - the development community of scholars and practitioners in various fields of science. Conference was attended by scientists and experts from Kazakhstan, Russia. International scientific conference was supported by the publishing house of the International Centre of research projects.

ISBN 978-80-7534-083-2 (Skleněný Můstek, Karlovy Vary, Czech Republic)
ISBN 978-5-00090-107-6 (MCNIP LLC, Kirov, Russian Federation)

Table of Contents

SECTION 1. PHYSICS AND MATHEMATICS

A CERTAIN PROBLEM OF CLUSTERIZATION OF DISTRIBUTED SYSTEMS

BOLOTIN E.I., ZENKEVICH S.L.

RUSSIA, MOSCOW STATE TECHNICAL UNIVERSITY NAMED N.JE.BAUMAN

Abstract. A problem of multi-level clusterization of distributed systems is considered in this paper. The problem is to find the multi-level, hierarchically organized partition of the system, which minimizes the total number of connections between objects in the system. We describe two partitioning strategies, and deduce the analytical recurrent equations allowing to find the number of the objects in each subsystem given the level of partitioning. We also conduct a number of computational experiments.

Keywords: clustering, distributed system, communication, connection.

1. Introduction

In our opinion, clusterization problems are not discussed in theory of distributed systems. However, there are some applications where certain grouping of objects constituting the system becomes quite important.

For example, [1] describes a problem of cooperation among the groups of quadro-copters, who solve the task of catching the ball. Part of the net is attached to each of the robot in the group. By solving that problem by a group of quadro-copters, they can use larger net, and increase the efficiency of a solution.

Paper [2] describes a problem of locating the moving objects inside the room, by teams of mobile robots equipped with laser rangefinders. By cooperating, the robots synchronize their idea about the environment. This allows to decrease the complexity and volume of data gathered, and more efficiently detect the moving objects.

The problem of calculating the optimal topological structure also arises in the field of constructing local and global computer networks [3]. The problems of decreasing the latency between network segments is considered, and decreasing the number of intermediate nodes.

One of the applications where clusterization issues have a great impact is the following: consider *N* objects which solve the common task. They need to communicate to each other using some communication channel [4]. Efficient and reliable operation of such channels, and decreasing the number of such channels, is quite an important task. This problem is especially challenging for the "swarm"-type systems which contain hundreds of objects. By partitioning this distributed system to subgroups we could approach such problem.

A problem of minimization of a number of links also arises when a group of objects constitutes a multi-agent system, that is it includes a number of agents which could make some decisions independently. In this situation, however, there is a an additional problem of choosing the structurization strategy.

Similar problem arises in social groups: it is complicated to ensure the normal functioning of large unstructured groups of people. We always use some kind of hierarchy, which provide some structurization of such groups. This is particularly true for scientific, administrative and industrial collaborations, where the first priority is concerned with an effective management structure.

2. Preliminary discussion

Before we define the problem formally, let us consider some illuminating examples.

2.1 Example 1

Given the group of 12 objects, for an unstructured case ("peer-to-peer") we have the number of links $L = 12 \times (12-1)/2 = 66$. Let us divide that group into two halves containing 6 objects each. Communication inside the subgroups is peer-to-peer, while communication between the groups is via chosen nodes ("switches"). Then the number of links is $L = 2 \times 6 \times (6-1)/2 + 1 = 31$. Let us now divide the group into 3 subgroups with 4 objects each. We have

$L = 3 \times 4 \times (4-1)/2 + 3 \times (3-1)/2 = 21$. Finally, if we partition the group into 12 subgroups (that is, each of the nodes is a switch), then we have the same result as for no partitioning: $L = 12 \times (12-1)/2 = 66$. It is clear that there is an optimal partitioning, where the number of links is minimum.

Notes:

1) We divide the group for the subgroups of the same size. Let us show that this is the best case. Let x_i be the size of *i*th subgroup. Then given the number *k* of subgroups the number of links for partitioning the *N* objects is $L = \sum_{i=1}^{k} x_i(x_i - 1)/2 + k(k-1)/2$, given that $\sum_{i=1}^{k} x_i = N$. Solving this simple equation for the conditional extremum of *L* function, we get $x_1 = x_2 = \cdots = x_k = N/k$. (All the expressions are symmetrical in x_i.)
2) It is obvious that for the general case where x_i is an integer number, this requirement could become invalid: for example, we could try to divide 10 objects into 3 groups. In that case we would have to search some rounded numbers as a solution.
3) We do not take into consideration the bandwidth for the communication links. It is clear that for the faulty partitioning inter-group channels could be used too much. This problem is not discussed in this paper, however it is rather important.

2.2 Example 2

Let us increase the partitioning depth, that is divide each of the subgroups into a number of sub-subgroups. Let us look at a simple example.

Consider a group of 12 objects divided into 3 subgroups of 4 objects each. Number of links equals 21. (See Example 1). Let us divide each of the subgroups into two sub-subgroups of 2 objects each. It is easy to see that the number of links becomes $L = 6 + 3 + 3 = 12$. So, the number of links decreased.

This observation allows to deduce that iterative partitioning of system of objects into smaller clusters can lead to decrease of the number of links. Note that we do not consider the possibility of increasing bandwidth requirement.

Let us look into that problem in more detail.

3. Problem statement

Let $C^{(n)}$ be the degree of the system clusterization. That is, $C^{(0)}$ means no further structure, $C^{(1)}$ means partitioning the group into subgroups, $C^{(2)}$ means further partitioning of subgroups into sub-subgroups, etc. Let x_i, x_{i+1} be a number of objects in each cluster for the *i*-th level of clusterization. Let $L^{(n)}$ and $N^{(n)} = N$ be the total number of links between the objects and full number of objects in a system using the rank-*n* clusterization (here *N* is the total number of objects in a system). Then we need to solve the problem of finding $x_i, i = 1, 2, \ldots, n+1$, which minimizes the number of links, where *n* is the given level of clusterization.

We will use the usual method of induction and Lagrange method of constrained optimization.

Let us consider two strategies of partitioning.

4. Strategy 1

This strategy consists of partitioning the group of objects into subgroups and appointing one of the objects for the role of "switch". This object does the same as the other objects, but it also provides the communication between objects of its own subgroup with objects from other subgroups.

The problem is to find such x_i that the number of links between the objects is minimum for the given level of clusterization.

Let us look at several initial induction steps.

4.1 Step $C^{(0)}$ (no structurization)

This step is needed for the following induction steps and of course it does not include the minimization procedure. The following obvious equations hold:

$$x_1 = N^{(0)}, L^{(0)} = \frac{x_i(x_i - 1)}{2}$$

4.2 Step $C^{(1)}$ (partitioning into subgroups)

Structurization discipline for this case is shown in Figure **Ошибка! Источник ссылки не найден.**.

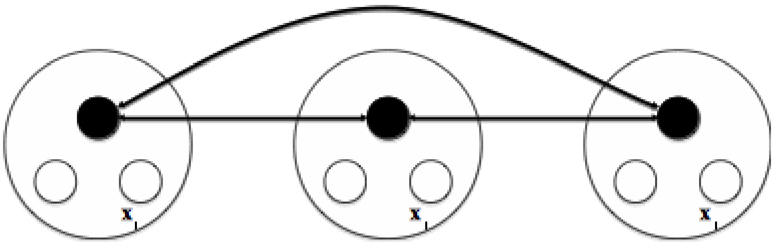

Figure 1. Structurization discipline for the case $C^{(1)}$

Let x_1 be the number of objects in subgroup, and x_2 the number of subgroups. It is easy to see then, that the following equations hold:

$$L^{(1)} = x_2 x_1 \frac{x_1 - 1}{2} + \frac{x_2(x_2 - 1)}{2} = x_2 L^{(0)} + \frac{x_2(x_2 - 1)}{2}$$

$$x_1 x_2 = N^{(1)} = x_2 N^{(0)}$$

The function to be minimized looks like this:

$$J^{(1)} = L^{(1)} + \lambda(N^{(1)} - N)$$

Taking into account the necessary extremum conditions, we get the following equations:

$$\lambda = \frac{1}{2} - x_1$$

$$\begin{cases} x_2 = \frac{1}{2}x_1^2 + \frac{1}{2} \\ x_1 x_2 = N \end{cases} \tag{1}$$

Let us introduce the additional parameter which we'll need in the following text:

$$I^{(1)} = L^{(1)} + \lambda N^{(1)} = L^{(0)}x_2 + \frac{x_2(x_2-1)}{2} + \lambda x_2 N^{(0)} = x_2 I^{(0)} + \frac{x_2(x_2-1)}{2} = -\frac{1}{2}x_2^2$$

It could be seen that searching for x_1, x_2 with given *N* requires, according to (1), to solve cubic equation for x_1 or x_2. We, however, will not discuss this simple problem, because the subsequent increase in level of clusterization will lead the equations of higher order, which could not be solved analytically. However, found relations allow to find x_2 and *N* given x_1, that is to solve the inverse problem.

Note:

Let us find the integer solutions of (1). From first equation it follows that x_1^2 is odd, and so x_1 is odd. Then integer solutions of 1 are as follows:

$$\begin{cases} x_1 = 2k+1 \\ x_2 = \frac{1}{2}x_1^2 + \frac{1}{2} \\ N = x_1 x_2 \end{cases}$$

where $k = 0, 1, 2, \ldots$

4.3 Example

Let us find first several integer solutions:

Table 1

k	x_1	x_2	N
0	1	1	1
1	3	5	15
2	5	13	65
3	7	25	175

4	9	41	369

For example, for the *N*=65 objects the optimal partitioning with regard of the number of links is 13 groups by 5 objects each.

4.4 Step $C^{(2)}$ (partitioning subgroups into sub-subgroups)

In this case we need to minimize the function

$$L^{(2)} = x_3 L^{(1)} + \frac{x_3(x_3-1)}{2}$$

given that

$$N^{(2)} = x_1 x_2 x_3 = x_3 N^{(1)} = N$$

which requires minimizing function

$$J^{(2)} = L^{(2)} + \lambda(N^{(2)} - N)$$

Calculations lead to the following result:

$$\lambda = \frac{1}{2} - x_1$$

$$\begin{cases} x_2 = \frac{1}{2}x_1^2 + \frac{1}{2} \\ x_3 = \frac{1}{2}x_2^2 + \frac{1}{2} \\ x_1 x_2 x_3 = N \end{cases} \qquad (2)$$

Let us calculate the additional parameter:

$$I^{(2)} = L^{(2)} + \lambda N^{(2)} = x_3 I^{(1)} + \frac{x_3(x_3-1)}{2} + \lambda x_3 N^{(1)} = -\frac{1}{2}x_3^2$$

Note: It is easy to see that the integer solutions of system (2) look like this:

$$\begin{cases} x_1 = 2k+1 \\ x_2 = \frac{1}{2}x_1^2 + \frac{1}{2} \\ x_3 = \frac{1}{2}x_2^2 + \frac{1}{2} \\ N = x_1 x_2 x_3 \end{cases} \quad (3)$$

where $k = 0,1,2,\ldots$

4.5 Example

We now show the example of best partitioning, calculated according to (3):

Table 2

k	x_1	x_2	x_3	N
0	1	1	1	1
1	3	5	13	195
2	5	13	85	5525
3	7	25	313	54775

For example, for $N = 195$ the best partitioning with regard to the number of links is 13 groups by 5 subgroups by 3 objects in each subgroup.

4.6 Step $C^{(n)}$ (induction step)

By looking at the relations above we notice that analogous relations could hold for the general case. Let us show that.

The following recurrent relations are true for recursively changing the level of clusterization:

$$\begin{aligned} & L^{(n)} = x_{n+1}L^{(n-1)} + \frac{x_{n+1}(x_{n+1}+1)}{2}, L^{(0)} = \frac{x_1(x_1+1)}{2} \\ & N^{(n)} = x_{n+1}N^{(n-1)} = N, N^{(0)} = N \\ & J^{(n)} = L^{(n)} + \lambda(N^{(n)} - N) \end{aligned} \quad (4)$$

Required conditions of function extremum for (4) are as follows:

$$\begin{cases} J_{x_i}^{(n)'} = L_{x_i}^{(n)'} + \lambda N_{x_i}^{(n)'} = 0; i = 1; 2; \dots \\ N^{(n)} = N \end{cases} \tag{5}$$

From (4) it follows:

$$L_{x_i}^{(n)'} = \begin{cases} x_{n+1} L_{x_i}^{(n-1)'}, & i = 1, 2, \dots, n \\ L^{(n-1)} + x_{n+1} + \dfrac{1}{2}, & i = n+1 \end{cases}$$
$$N_{x_i}^{(n)'} = \begin{cases} x_{n+1} N_{x_i}^{(n-1)'}, & i = 1, 2, \dots, n \\ N^{(n-1)}, & i = n+1 \end{cases} \tag{6}$$

Suppose that from relations (5) the following equations over $x_1, x_2, \dots, x_{n+1}$ follow:

$$\begin{cases} x_{i+1} = \dfrac{1}{2} x_i^2 + \dfrac{1}{2}; i = 1; 2; \dots; n \\ \prod\limits_{j=1}^{n+1} x_j = N \end{cases} \tag{7}$$

And for the additional parameter the following relation is true:

$$I^{(n)} = L^{(n)} + \lambda N^{(n)} = -\frac{1}{2} x_{n+1}^2 \tag{8}$$

Let us show that for *N*+1-level clusterization $C^{(n+1)}$ system of equations over the needed variables $x_i, i = 1, 2, \dots, n+2$ will look like this:

$$\begin{cases} x_{n+1} = \dfrac{1}{2} x_i^2 + \dfrac{1}{2}; i = 1; 2; \dots; n+1 \\ \prod\limits_{j=1}^{n+2} x_j = N \end{cases} \tag{9}$$

The required condition for the minimum in that case is like this:

$$J_{x_i}^{(n+1)'} = L_{x_i}^{(n+1)'} + \lambda N_{x_i}^{(n+1)'} = 0$$

or, by using the recurrent relations (4):

$$x_{n+1} L_{x_i}^{(n)'} + \lambda x_{n+1} N_{x_i}^{(n)} = 0, i = 1, 2, \ldots, n$$

where taking into account (4) we get the relations (9) for $x_1, x_2, \ldots, x_{n+1}$. Let us look at the relation for x_{n+2}. Taking into account the relation (6) we have:

$$L^{(n)} + x_{(n+2)} + \frac{1}{2} + \lambda N^{(n)} = 0$$

where given (8):

$$I^{(n)} + x_{n+2} - \frac{1}{2} = 0$$

or

$$x_{n+2} = \frac{1}{2} x_{n+1}^2 + \frac{1}{2}$$

So, the following relations are the necessary extremum conditions for the problem of *n*-level clusterization of distributed system, which includes *N* objects and minimizes the number of links:

$$\begin{cases} x_2 = \frac{1}{2} x_1^2 + \frac{1}{2} \\ x_3 = \frac{1}{2} x_2^2 + \frac{1}{2} \\ \ldots \\ x_{n+1} = \frac{1}{2} x_n^2 + \frac{1}{2} \\ x_1 x_2 x_3 \cdots x_n x_{n+1} = N \end{cases} \quad (10)$$

It is also easy to get the equation for the additional parameter:

$$I^{(n+1)} = -\frac{1}{2}x_{n+2}^2$$

Let us note that the sequence of x_i is non-decreasing:

$$x_{i+1} - x_i = \frac{1}{2}x_i^2 + \frac{1}{2} - x_i = \frac{1}{2}(x_i - 1)^2 \geq 0$$

This means that the number of subgroups is not less than the number of objects in subgroup.

5. Strategy 2

Let us consider another, different strategy.

In previous case the selected objects had a role of communication devices. But for the advanced systems it is required to appoint some objects to fulfil the management roles, commanding the clusters which are lower by hierarchy, and communicating with higher-level objects. Such situations arise, e.g., when objects are the robots which could make their own decisions, that is they constitute a multi-agent system. The same situation arises in social collectives. Let us make several notes:

1) Each of the clusters has its managing node. The set of managing nodes constitute the next-level cluster, which has its own managing node.
2) Clusters do not contain their managing nodes. So, x_i is the number of "employees" in a cluster, but without their "manager". x_2 is the number of managers, without their "high-level manager", and so on. Structure of such a system is presented in Figure 2.

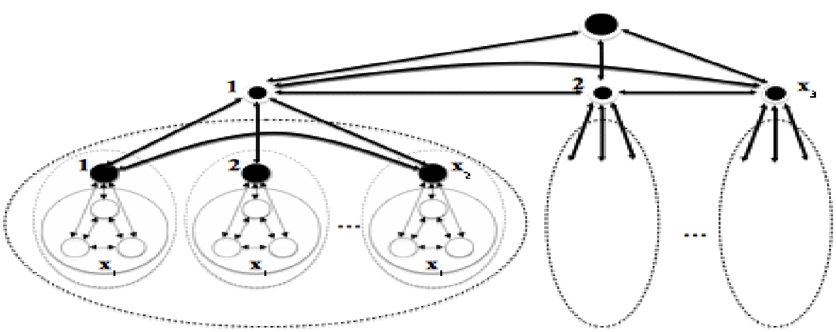

Figure 2. Structure of the "department"

As before, the problem is to choose such a structure that the number of links is minimum on a given level of structuring.

Method of solving this problem is mostly the same as in a previous strategy. So we shall omit all the calculations and present only the final results.

For then *n*-level clusterization we have the following series of recurrent relations:

The number of links, that is the function to be minimized, is:

$$L^{(n)} = x_{n+1} N^{(n-1)} + 1, N^{(0)} = x_1 + 1 \quad (11)$$

The total number of objects:

$$N^{(n)} = x_{n+1} N^{(n-1)} + 1, N^{(0)} = x_1 + 1 \quad (12)$$

Then then required conditions for the extremum are:

$$\begin{cases} L_{x_i}^{(n)'} + \lambda N_{x_i}^{(n)'} = 0; i = 1;2;\ldots;n+1 \\ N^{(n)} = N \end{cases} \quad (13)$$

Relations (13) are *n*+2 equations with *n*+2 unknown variables $\lambda, x_1, x_2, \ldots, x_{n+1}$.

Taking into account the recurrent relations (11, 12), we get the following equations:

$$\begin{cases} L_{x_i}^{(n-1)'} + \lambda N_{x_i}^{(n-1)'} = 0; i = 1;2;\ldots;n \\ L^{(n-1)} + x_{n+1} + \frac{1}{2} + \lambda N^{(n-1)} = 0; i = n+1 \\ N^{(n)} = N \end{cases} \quad (14)$$

Equivalent system of equations over $x_i, i = 1,2,\ldots,n+1$ looks like this:

$$\begin{cases} x_2 = \frac{1}{2}x_1^2 + x_1 \\ x_3 = \frac{1}{2}x_2^2 + x_1 \\ \dots \\ x_{n+1} = \frac{1}{2}x_n^2 + x_1 \\ \sum_{i=1}^{n+1}(\prod_{j=i}^{n+1} x_j) + 1 = N \end{cases} \qquad (15)$$

Relations (15) are the final solution to the link minization problem using the Strategy 2.

Note: It is easy to see that the integer solutions to the system (15) have the following form:

$$\begin{cases} x_1 = 2k \\ x_2 = \frac{1}{2}x_1^2 + x_1 \\ x_3 = \frac{1}{2}x_2^2 + x_1 \\ \dots \\ x_{n+1} = \frac{1}{2}x_n^2 + x_1 \\ \sum_{i=1}^{n+1}(\prod_{j=i}^{n+1} x_j) + 1 = N \end{cases} \qquad (16)$$

where k=1,2,3,...

5.1 Example

For example, here are the best integer partitions for structurization level n=3, using the Strategy 2: