



PERGAMON

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Acta Astronautica 59 (2006) 1086–1092

ACTA
ASTRONAUTICA

www.elsevier.com/locate/actaastro

A proposed model for small-world structural organization of mission teams and tasks in order to optimize efficiency and minimize costs

André S. Ribeiro^{a, b, *}, Miguel Almeida^c

^a*Institute for Biocomplexity and Informatics, Department of Physics & Astronomy, University of Calgary, Canada*

^b*Center for Computational Physics, University of Coimbra, P-3004-516 Coimbra, Portugal*

^c*ESA/ESTEC, SCI-SB, Keplerlaan 1, NL-2201 AZ Noordwijk, The Netherlands*

Available online 10 November 2005

Abstract

We propose a model of structural organization and intercommunication between all elements of every team involved in the development of a space probe to improve efficiency.

Such structure is built to minimize path between any two elements, allowing fast information flow in the structure.

Structures are usually very clustered inside each task team but only the heads of departments, or occasional meetings, usually assure the links between team elements. This is responsible for a lack of information exchange between staff members of each team.

We propose the establishment of permanent small working groups of staff elements from different teams, in a random but permanent basis. The elements chosen for such connections establishment can be chosen on a temporary basis, but the connections must exist permanently because only with permanent connections can information flow when needed.

A few of such random connections between staff members will diminish the average path length, between any two elements of any team, for information exchange.

A small world structure will emerge with low internal energy costs, which is the structure used by biological neuronal systems.

© 2005 Published by Elsevier Ltd.

1. Introduction

Several teams are part of a mission organization, each with specific tasks, performed by their task members.

Team heads periodically exchange information between each other and then transmit it to their staff members, assuring coordination between teams.

Periodical meetings between other staff members occur to discuss specific questions to help coordinating tasks.

Less formal forms of communication include e-mail exchange, phones and occasional meetings at locations such as the cafeteria. The problems associated with

* Corresponding author. Tel.: +1 403 220 2425;

fax: +1 403 210 8655.

E-mail addresses: ARibeiro@ucalgary.ca (A.S. Ribeiro), malmeida@rssd.esa.int (M. Almeida).

these communication processes are the lack of time to trade relevant information and relies on casualty to bring relevant subjects to conversation.

These structures are somehow inefficient in the sense that each staff element does not have a clear view of the global developments in other teams, because information usually does not flow directly between him and staff members of other teams.

When communication occurs, the “right questions” might not be placed, that is, information held by one staff member might not be by him considered relevant to anyone else, although it might be crucial to someone in a completely different task.

This is a consequence of the lack of knowledge that staff members usually have about other teams tasks and achievements, even in a general sense. It causes delays in information exchange and consequently in goals achievement or unnecessary repetition of tasks by different teams.

One solution to this problem is the establishment of a free access database, with all teams works. This is not the best solution although it helps the information flow between all.

Unfortunately, such database rapidly becomes unreadable due to the enormous amount of information that it contains, making each team member to reduce his reading almost only to his own work related contributions to such database. Other problem is due to the necessity of restricted access and usage of these databases, limiting also information sharing.

The use of these databases allow access to documents related to someone work but, usually, someone not directly involved in a task, takes more time to learn by studying documents about a subject outside the scope of his specialty than it would take by trading information directly with someone involved in that task.

Another problem arises due to human psychology. It is easier to share information with someone “known” to us, specially the “state of the art” of a work, than to share it in a public database.

All the processes of information exchange described above make the process somehow sparse in time. Such is the reason of its lack of efficiency.

Since the system properties lie in the structure, the system improvement of efficiency will be achieved with structural changes.

In the case of the structure studied here, since we cannot simply change the entire system structure due

to the costs involved and because the existing structure has many advantages, we propose the establishment of a few permanent connections between some of the staff members, from different teams, without altering the existing structure.

The nature of such connections is permanent working relations. A few staff members would work at the same time in their own teams and with a very small group of staff members from other teams.

Such “long range” connections allow permanent information exchange, resulting in a faster information flow due to the decrease of path length between staff members of any team.

As we shall see, the elements with such “long range” connections can be chosen randomly and on a temporary basis, but connections between staff members of different teams must exist permanently because only permanent connections allow information flow when needed.

2. A geometrical model of the system

From systems theory [1] we know that systems properties lie in connections and not in its elements properties [2].

Therefore we propose a model of system, where staff members are the elements of such system, with no specific inner properties, and, direct work relations between any two staff members are the interactions, represented by lines connecting staff members.

Elements’ spatial position is irrelevant and is made only for representation purposes. The only relevant information about the system structure is to whom each element is connected.

Connections between elements represent direct work relations. In order for two staff members to have a direct work relation, they must work in the same project. This leads to considering staff members from the same team always directly connected.

Using this model, we are able to represent the system structure as a graph of points and lines, where “distance” between staff members is measured only by the minimum number of connections needed to reach from one to another [3–5].

Because within a staff, all elements are connected, we can represent only the team as the system element although reminding that the connections are actually between staff members and not between teams.

Between two teams only one connection is, at most, considered since a line represents any kind of interaction, making more than one connection between two elements redundant.

This model allows the computation of the cluster coefficient, the diameter and average path length [3] from the knowledge of the connections structure.

It also allows the computation of a measure of the systems internal energy, that is, the necessary energy need to be spent when creating the system's structure.

These quantities characterize the system efficiency in terms of information flux as new connections are added.

Those new connections will be between two staff members only for sake of simplicity. Yet there is another reason. Although a group of, for example, five staff members from five different teams decreases path length approximately by the same amount that five groups of two staff members each, it would have a disadvantage. If such connection fails, the path length is strongly affected, increasing rapidly. Also, too much information would have to be transmitted in the same channel, diminishing efficiency.

Moreover, the number of connections created between k members, would be equal to $[k*(k-1)/2]$. The more staff members involved in the new connection system, the more internal energy needs to be spent, as we shall see.

The solution of using a few small groups provides to the system some redundancy in the possible paths between staff members. That is, if a certain connection "fails", the others can still be used, not affecting path length by a great amount. It also requires less information exchange using one single channel, thereby allowing a faster information flux.

Thus, the use of this technique requires minimal internal energy costs.

3. Internal energy

Because spatial distance is not considered in the model of system used, the internal energy, U , necessary to create the system, is an exclusive function of the number of elements, m , and connections, n :

$$U = U(n, m). \quad (1)$$

To create each connection and each element a certain amount of energy is needed.

It is then possible to establish a formula to compute the internal energy and its variation with the variation of the number of elements and connections.

Such relation must respect the following conditions:

$$\frac{dU}{dn} > 0, \quad (2)$$

$$\frac{dU}{dm} > 0. \quad (3)$$

The problem consists in knowing the relative importance of the variation of the number of elements and number of connections, for the variation of the internal energy.

Since spatial distance is irrelevant each connection has the same amount of internal energy. Since elements have no inner properties they all have the same amount of internal energy.

Therefore two constants with energy units must be associated, respectively, to connections and to elements.

For simplicity, but respecting the laws of variation of the internal energy with the variation of the number of elements and connections, the internal energy necessary to create the system can be defined as

$$U(m, n) = k_1 \cdot m + k_2 \cdot n. \quad (4)$$

The positive constants, k_1 and k_2 , represent the average energy required to create an element and the average energy required to create a connection. Considering that the constants k_1 and k_2 represent energy "quantum" quantities, they must be positive: $k_1, k_2 > 0$. Therefore:

$$\frac{\partial U}{\partial m} = k_1 > 0 \quad \forall m > 0, \quad (5)$$

$$\frac{\partial U}{\partial n} = k_2 > 0 \quad \forall 0 \leq n \leq \frac{m \cdot (m - 1)}{2}. \quad (6)$$

We conclude that the internal energy increases with the number of connections, for all the possible values of n . The same conclusion is valid for the increase of the number of elements.

The internal energy increase, in a system of constant number of elements, due to the increase of connections, is the limitation to connections establishment.

Such limitation is caused by the necessity of system structure to obey to the principle of least energy.

The process of connections' adding to decrease path length between elements is thereby limited by the principle of least energy.

Therefore, our search for an optimal information exchange between elements must lie in the connections disposition, rather than in the total number of connections [6,7].

4. The process of inter-staff working groups establishment

We shall consider that two staff members with direct work relations have a reasonable knowledge of the work each other are developing. This is a necessary condition of the model used here. Moreover, we consider that elements from the same team have a direct work relation.

Teams are designated by letters A, B, C, . . . , and staff members of those teams, designated by $A_1 \dots A_k$. Imagine someone, element B_1 from staff B, wanting to know information on a certain document produced by element A_1 , but he does not know that such information is in that document.

From the existing structure, B_1 has two processes of searching for the information he requires. Either he searches a library (virtual or real) or he asks someone. If none of the people he asks has a direct work relation with A_1 , they probably would not know his work and will only be able to give general guidance of where the information might exist.

Both these search processes might take a lot of time until A_1 is reached. For example, B_1 asks B_7 , who knows C_3 , who knows A_4 , who knows A_1 .

We call "path" between B_1 and A_1 to be the set of connections used to reach from B_1 to A_1 . The number of connections of that path defines the path length.

The path used might not even be, in the example, the path of minimum length between B_1 and A_1 .

From these considerations we realize that an optimal structure for knowledge sharing must have minimal average path length, between any two elements, for the existing number of connections.

From graph theory [4,5], we know that the structure that minimizes path length between elements is the random distribution of connections.

We cannot apply such structure because teams would lose cohesion, and consequently its ability to develop group projects.

We must work with the already existing structure and try, using the least number of extra connections possible, to minimize average path length.

The solution then is to maintain the existing structure and add random connections between elements of different teams.

As we shall see, a very small number of such connections will diminish the path length significantly. The cost is a small increase of internal energy due to increasing the number of connections.

5. Path length, diameter and clustering coefficient

The variables that characterize a system structure are the average path length, L , the diameter D , and the clustering coefficient, C [3].

Given two elements of a system, s_i and s_j , the smallest path between them, L_{ij} , is defined as the smallest number of connections needed to be crossed by a signal from s_i to s_j . The average path length can be computed as the average of the sizes of all paths. If m is the total number of elements, the number of smallest paths between all m elements is C_2^m [11]:

$$\langle L \rangle = \sum_{i,j} \frac{L_{ij}}{C_2^m}. \tag{7}$$

We can estimate the average number of steps needed for an element, in a system of m elements and n connections where each element has k connections in average, to send a signal to any other element of the same system [8,9]. For a signal starting in one element to reach each one of the others ($m - 1$) elements, with $m = k^l$, it is required, in average, a number of steps given by [6]

$$\langle L \rangle \approx \frac{\text{Log}(m)}{\text{Log}(k)} \quad \text{with } k = \frac{2.n}{m}. \tag{8}$$

This quantity, that measures the degree of separation of the graph, is always smaller in graphs where connections are totally random than for regular or semi-regular graphs, supposing all graphs with m elements and n connections.

Another quantity measures the network capability to transmit information between all elements. Such

quantity is the graph diameter and is equal to the largest path length found in the graph

$$D = \text{Max}(L_{ij}) \quad \forall i, j. \tag{9}$$

To measure defects on graphs regularity, Watts and Strogatz defined a clustering coefficient, $C(p)$, which is the number of nearest neighbors of each element, that are also first order neighbors between themselves. From the cluster coefficient definition [3,9,10], with E_i representing the number of connections between the k_i elements connected to a single element, it results that $k_i \cdot (k_i - 1) / 2$ is the maximum number of connections possible between those elements [11]. Thus we have:

$$C(p) = \langle C_i \rangle_i = \left\langle \frac{2 \cdot E_i}{k_i \cdot (k_i - 1)} \right\rangle. \tag{10}$$

Using these quantities we can study the effects of the new connection introduction in the system.

6. The effects of random connections introduction

Let us suppose an initial typical structure and add a few connections randomly. We can, from this example, compute the variation of path length, diameter and cluster coefficient to determine the effects of new connections randomly introduced in the structure.

We can now observe the effect caused by the introduction of a few random connections due to the creation of inter teams work groups using only a few elements of each team.

Supposing we had added each new connection one at a time, we can compute, using (6), the variation of L as the new connections are added.

The initial structure had a small value of L . Yet its variation is relevant.

Notice that, although the decrease seems “small” we only added 5 connections, and caused a significant change.

In fact, since there are 21 elements, there are C_2^{21} possible paths. For each path, in average, its distance went from 2,6, before random connections adding, to 2,3 (after those connections are introduced).

To compute in time spent, let us suppose that, on average, one relevant message is transferred between

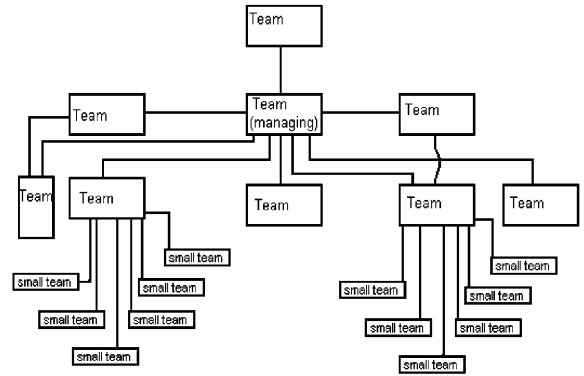


Fig. 1. A typical structure where teams are elements, and direct work relations and dependencies are connections. L is equal to 2,6.

two elements, with a direct connection, in 1 h. Since there are C_2^{21} , that is, 210 possible paths, the time for every element to send a message to every other element in the system is reduced by a significant amount:

$$T_i = 2, 6 \cdot 210 = 546 \text{ h,}$$

$$T_f = 2, 3 \cdot 210 = 483 \text{ h.}$$

We had 22 connections, and added 5. This is not a large internal energy increase.

If we wished to maintain the internal energy, we could have removed some redundant regular connections without affecting significantly the path length. Our only restriction is keeping all teams linked in a single cluster whatever is the number of connections between them.

The new connections diminish significantly, the path length and diameter because they connect elements from teams with large path length between them [12].

The diameter variation is also an indication of the improvement of efficiency

$$\Delta D = D_f - D_i = 4 - 3 = 1. \tag{11}$$

In this case, a significant 25% variation is observed.

As for the clustering coefficient, no significant changes were observed. The initial clustering coefficient is, computed from the initial structure (Fig. 1) using (2)

$$C(p)^{\text{initial}} = 0.148. \tag{12}$$

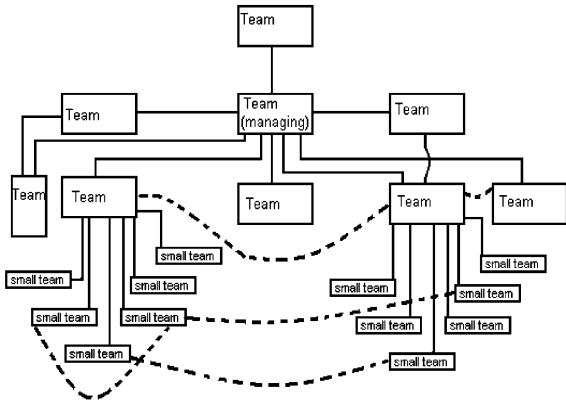


Fig. 2. A structure originated from the introduction of a few random connections between elements of different teams. L is equal to 2,3.

The final clustering coefficient is, computed from the final structure (Fig. 2) using (2):

$$C(p)^{\text{final}} = 0.149. \tag{13}$$

The variation caused by the new connections is irrelevant thus indicating that the structure has not been significantly altered in terms of teams structure, although the path length of the system as a whole has diminished.

7. Characterizing the resulting structure. The emergence of a small-world

Structures vary from regular to random, according to the rule used for placing connections.

The random structure is characterized by a small clustering coefficient and small average path length. For a system of m elements and n connections randomly placed, the clustering coefficient is given by

$$C_{\text{random}} = \frac{k}{m} = \frac{2n}{m^2} = 0.122 < C(p)^{\text{initial}}, C(p)^{\text{final}}. \tag{14}$$

The values observed by us showed that our example of system has larger clustering coefficient values than those exhibited by a random network of the same number of elements and connections.

For a completely regular structure, all elements have the same number of connections. In our case this is not true, therefore, the system is also not a regular graph.

As for the average path length, our system exhibits a small value, characteristic of random graphs (Fig. 3).

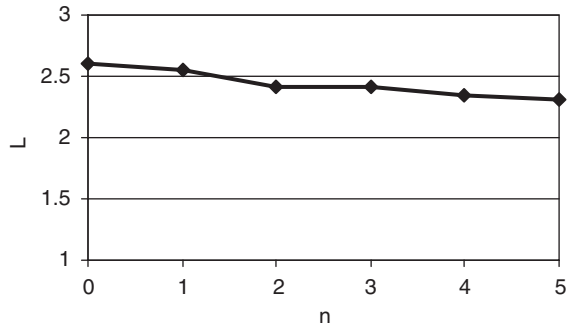


Fig. 3. Path length decrease as random connections are added to the system elements.

A small world structure is characterized for having a small average path length, allowing fast communication between all elements ($\langle L \rangle_{\text{SW}} \cong \langle L \rangle_{\text{random}}$) and a high clustering coefficient ($C_{\text{SW}} > C_{\text{random}}$) which means that if two elements are connected to a third one, they are most probably connected between themselves.

Such are the characteristics of the structure generated by us with the random connections adding process, so that a small world structure has been generated.

We believe such structure is the most efficient possible, in this case, in providing fast information exchange between all elements and if one element is by some reason removed, maintaining the structure efficiency.

Such capacity is due to the high clustering coefficient value and does not occur in random graphs, where the removal of a single element can cause a significant increase of path length.

In fact, the small world structure is the structure used by the majority of biological neuronal networks and others [13–16], which, in some way, validates our results.

Biological neuronal networks were generated through an evolution process so we believe their structure is optimal.

These networks have high clustering, allowing to maintain the teams’ structure, with the advantage of reducing the path length between any two elements of the system, when compared with regular networks, thereby reducing information exchange delays.

Because it needs only a few random connections to reduce path length, the increase of internal energy

necessary to create the extra connections is almost insignificant comparing with the energy already used to create the initial structure.

8. Conclusions

We proposed the establishment of permanent working groups of staff members from different teams in order to improve information flow between different teams, working on distinct projects but with common goals and dependencies. We represented such work relations by lines connecting staff members, the elements of the system.

By adding only a few connections [5] we maintained the clustering coefficient but decreased significantly the average path length, thereby improving the efficiency of information exchange between all staff members of all teams.

Because of the decrease in average path length, the average time for information flow between any two elements of the system diminishes thereby improving the system efficiency. The diameter decrease is another evidence of such improvement.

Using this method, where random connections are introduced between elements of different teams, a small world structure was generated.

Such structure is the structure observed in the majority of the biological neuronal networks [13–15] and other biological networks [6,12,16]. Such is an evidence of the model efficiency.

References

- [1] L. Von Bertalanffy, *General System Theory*, Penguin Books, Harmondsworth, 1973.
- [2] F. Carvalho Rodrigues, J. Dockery, Defining systems based on information exchange: structure from dynamics, *BioSystems* 38 (1996) 229–234.
- [3] D.J. Watts, S.H. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature* 393 (1998) 440–442.
- [4] B. Bela, *Modern Graph Theory*, in: *Graduate Texts in Mathematics*, vol. 184, Springer, New York, 1998.
- [5] D. Reinhard, *Graph Theory*, in: *Graduate Texts in Mathematics*, vol. 173, Springer, New York, 1997.
- [6] M.E.J. Newman, Models of the small world: a review, *Journal of Statistical Physics* 101 (2000) 819–841.
- [7] G. Strang, Random shortcuts make it a small world indeed, *SIAM News* 32 (10).
- [8] A. Barabasi, *Linked: The New Science of Networks*, Perseus Publishing, Cambridge, MA, 2002.
- [9] R. Albert, H. Jeong, A.-L. Barabasi, Diameter of the World-Wide Web, *Nature* 410 (9) (1999) 130–131.
- [10] K. Klemm, V. Eguíluz, cond-mat/0107607.
- [11] P. Erdos, A. Rényi, On the evolution of random graphs, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5 (1960) 17–61.
- [12] M.S. Granovetter, The strength of weak ties, *American Journal of Sociology* 78 (6) (1973) 1360–1380.
- [13] J.G. White, E. Southgate, J.N. Thomson, S. Brenner, The structure of the nervous system of *Caenorhabditis elegans*, *Philosophical Transactions of the Royal Society of London, Series B* 314 (1986) 1–340.
- [14] T. Achacoso, W. Yamamoto, *AY’s Neuroanatomy of C. elegans for Computation*, CRC, Boca Raton, FL, 1992.
- [15] D.J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness*, Princeton University Press, Princeton, NJ, 1999.
- [16] H. Pette, Characteristics of Small World Networks, 20th April 2001.