LARGE ORGANIZATIONS: HIERARCHIES OR CLUSTERS?

Tony Ostrom

February 1975   WP-75-21

Formal Hierarchy and Informal Communications

Large organizations tend to have a formal, structured hierarchy, yet a considerable amount of decision-making is done using an 'informal' communication structure which is often too complex to be shown on the traditional organization chart (See Fig. 1).

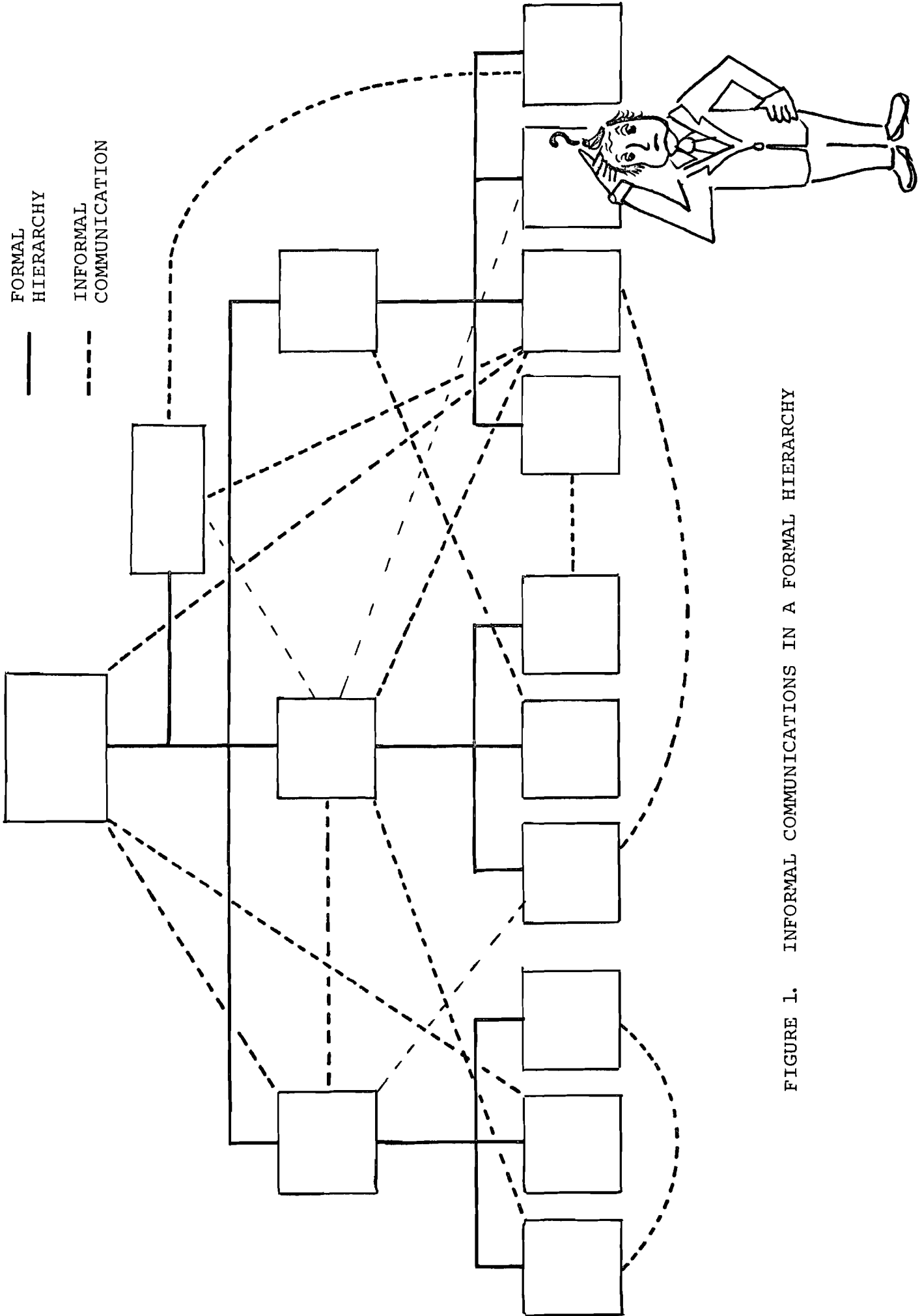What is the nature of this informal communication structure, and is it useful to understand it better?

FIGURE 1.   INFORMAL COMMUNICATIONS IN A FORMAL HIERARCHY

FORMAL
HIERARCHY

INFORMAL
COMMUNICATION

Communications Requirements

Suppose each individual in the organization is asked, by means of a questionnaire, to name all individuals with whom he needs face-to-face communication[1] in order to work effectively.[2] Using a numerical 'frequency' scale (many times per day = 5, once per day = 3, and so on) or an 'importance' scale (Vital = 5, very important = 3, not important = 0, etc.) we end up with a matrix $\tilde{A}'$ of communications a'_{ij} between individuals i and j.

With human judgment there is often conflict so that some sort of conflict resolution technique must be used where $a'_{ij} \neq a'_{ji}$ to produce a symmetric $\tilde{A}$. The purpose and method for producing a symmetric matrix will be seen later.

Objective Function

Having resolved all conflicts we have a symmetric matrix ($a_{ji} = a_{ji}$) of affinities between individuals. It is now desired to form groups or 'clusters' of individuals so that, as far as possible, those with close affinities lie in the same cluster. This is done (BEALE 1969), with observations represented in terms of orthogonal coordinates by minimising the sum of squares of the deviations of the observations from their respective cluster centres. It can be shown that this objective func-

---

[1] Telephone and written communications are expressly ignored to concentrate on face-to-face communication, probably the most influential means of inter-personal contact.

[2] This approach was originally developed as part of a space- and office planning system by Ostrom and Thomas, 1972. This paper looks at the problem again, but from the different viewpoint of structuring large organizations.

tion may be defined in terms of the inter-point distance matrix in the following way:

Define a distance matrix $d_{ij} = (M - a_{ij})$, whose M is some number larger than the largest a_{ij}. If n_c denotes the number of individuals in cluster $\overset{.}{c}$ than the sum of squares of the deviation of the individuals from their respective cluster center may be shown to be:

$$S = \sum_c \left[\frac{1}{n_c} \sum_{\substack{i \\ i,j \varepsilon c \\ i < j}} \sum_j d_{ij} \right] .$$

Substituting $(M - a_{ij})$ for d_{ij} we see that:

$$S = \sum_c \frac{1}{n_c} \frac{\left[n_c (n_c - 1) \ M \right]}{2} - \sum_c \frac{1}{n_c} \sum_{\substack{i \\ i,j \varepsilon c \\ i < j}} \sum_j a_{ij}$$

$$= \frac{1}{2} M \left(\sum_c (n_c - 1) \right) - \sum_c \frac{1}{n_c} \sum_{\substack{i \\ i,j \varepsilon c \\ i < j}} \sum_j a_{ij}$$

$$= \frac{1}{2} M(N - C) - \sum_c \frac{1}{n_c} \sum_{\substack{i \\ i,j \varepsilon c \\ i < j}} \sum_j a_{ij}$$

where N is the total number of individuals being clustered and C is the total number of clusters.

Thus to minimise S for any particular value of C we must maximise:

$$T = \sum_c \frac{1}{n_c} \sum_{\substack{i \\ i,j \in c \\ i < j}} \sum_j a_{ij} \quad .$$

A local optimal solution can be found for quite large N if most of the a_{ij} vanish.

A Heuristic Clustering Algorithm

1. For a starting solution, assign individuals to clusters arbitrarily.

2. Calculate the contribution of each cluster c to the objective function i.e.,

$$V_c = \frac{1}{n_c} \sum_{\substack{i \\ i,j \in c \\ i < j}} \sum_j a_{ij} \quad .$$

3. Take each cluster in turn. Each member is tested in turn for movement to all other clusters by calculating the change in the objective function entailed by

the move. If the gain is positive the move is made.
Otherwise, carry on to the next individual in the
cluster.

4. Step 3 is repeated for each cluster until no further
improvement in the objective function is possible.

The result is a local optimum clustering of individuals
for a given c. An improvement towards global optima can be
made by 'block moves' of two or more individuals from one
cluster to another. Such a procedure, however, greatly
increases computing time (see, for example, Cluster Analysis,
1971) and a trade off must be made between additional
computing costs and the extra benefits of global optima
that may result. Given the clustering results from such
a method it is natural to ask if the number of clusters is
statistically significant. Beale (1969) derived an F-test
to discover if a change in c will reduce the residual sum of
squares for a particular solution.

The Informal Cluster

Our once unwieldy organization is no longer characterized
by a formal organization chart, rather by a number of clusters
with specific numerical links between their members. An example cluster
is shown in Fig. $\underline{2}$. Each cluster represents a group of individuals who
need face-to-face contact and would rather work with each other
than with individuals in other clusters.

The 'strength' or value of each cluster c is represented by:

$$V_c = \frac{1}{n_c} \underset{\substack{i,j \varepsilon c \\ i < j}}{\sum_i \sum_j} a_{ij}$$

and the 'internal' contribution of each individual j to his

cluster c is:

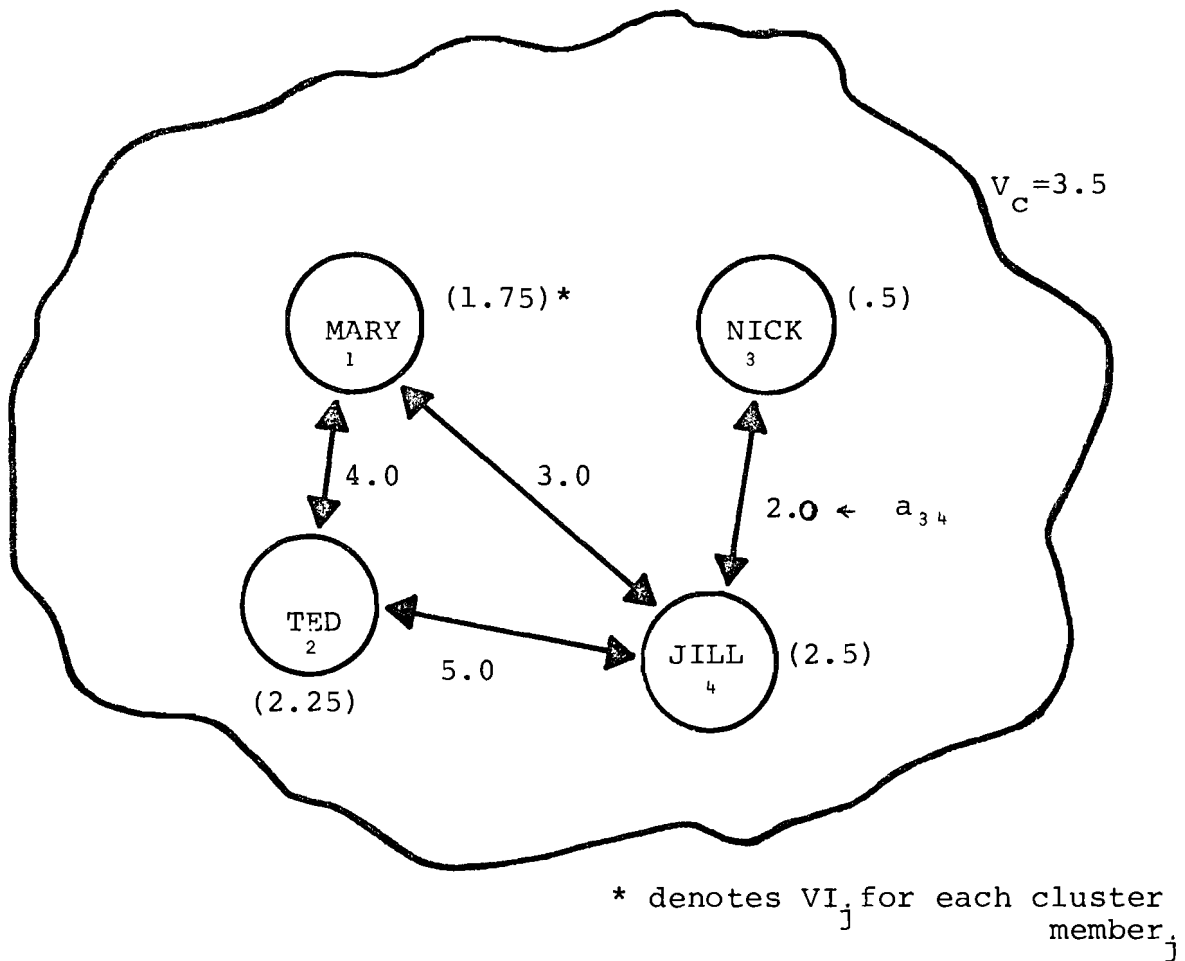$$VI_j = \frac{1}{n_c} \sum_{\substack{i \varepsilon c \\ i \neq j}} a_{ij} \qquad .$$

FIGURE 2.   AN INFORMAL CLUSTER

With V_c and VI_j we can understand how well the members

of each cluster communicate between themselves.

Inter-Cluster Affinities

We know the affinities between individuals, but what about relationships between clusters? As some members of a cluster have some outside communications, we shall define the communication values of each individual j 'external' to his cluster c as:

$$VE_j = \sum_{i \notin c} a_{ij} \quad .$$

Define the affinity between two clusters c_1 and c_2 with n_{c_1} and n_{c_2} members respectively as:

$$Aff_{c_1 c_2} = \frac{1}{n_{c_1} + n_{c_2}} \sum_{i \varepsilon c_1} \sum_{j \varepsilon c_2} a_{ij} \quad .$$

We have now fully specified the inter-cluster relationships as external communications by the cluster as a whole and by each individual within it. (See Fig. 3).

Analysis of Large Organizations

The clusters can be useful for manpower planning through a study of the functional role of each person in a cluster and the relationship between clusters. Clusters with similar communication patterns can be analyzed in detail to ensure that they do not have unnecessarily overlapping roles. To an outsider a clustering might reveal an informal hierarchy which is vastly different from the formal one, and thus expose the 'real' chain of command and power.
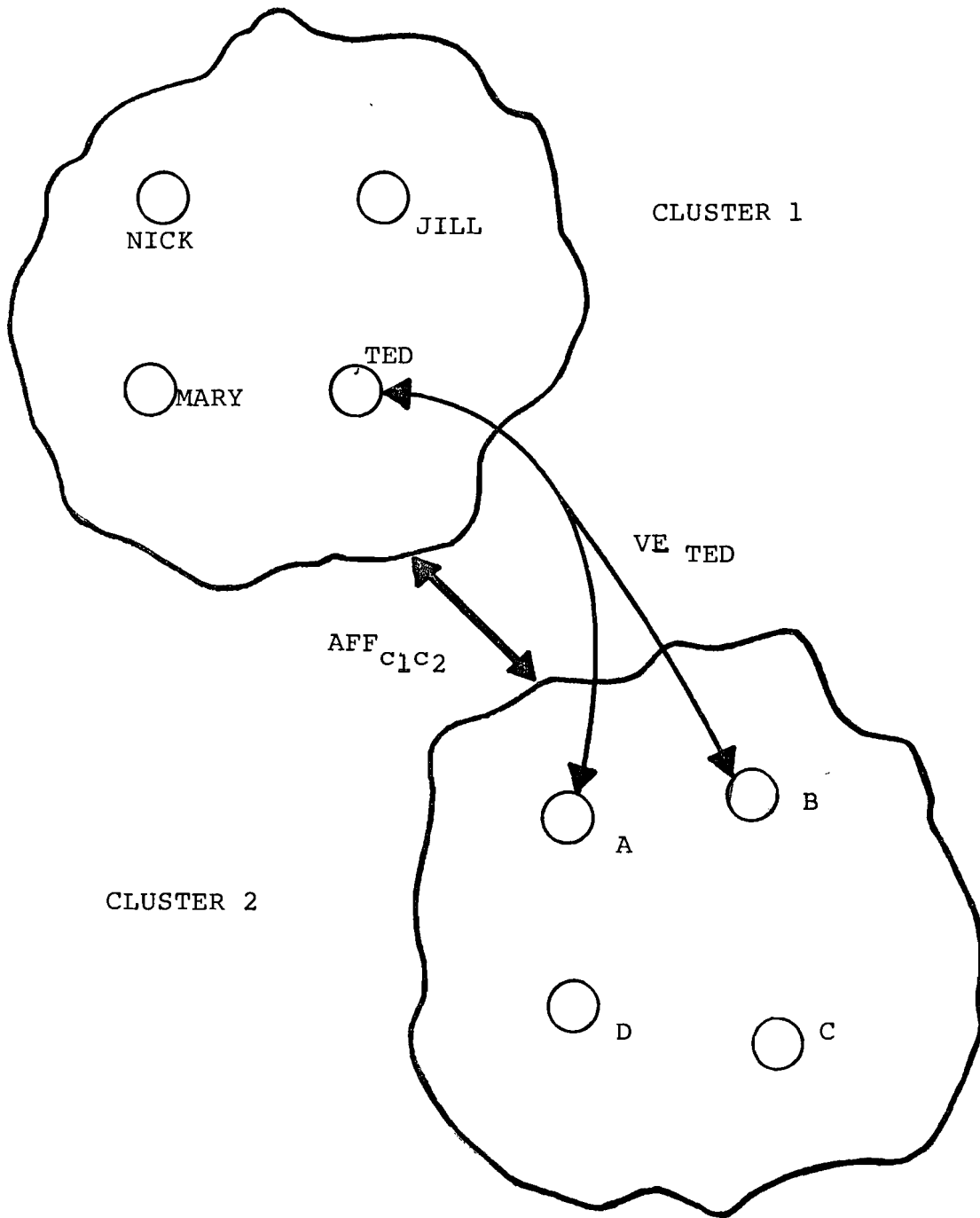
The clustering, of course, can be carried out on any

FIGURE 3.   INTER-CLUSTER RELATIONSHIPS

level of the organization. For very large government struc-
tures, for example, the clustering might be done at the branch
level[3]to identify actual (as opposed to theoretical) relation-
ships between functional units. Periodic reclustering could
identify the evolution of the organization, changing communi-
cations requirements and future space needs.

Bureaucracy Control

The often overwhelming size of many bureaucracies pre-
cludes any detailed understanding of the relationships between
functional groups, let alone those between individuals. By
clustering at any level the relationships between functional
units can be described from the dynamic communication patterns
of individuals within and between clusters. The complex maze
of communication links in Figure 1 can be succinctly compressed
into the more manageable 'inter' and 'intra'-cluster affinities,
so that the workings of the organization can thus be quickly under-
stood by the outsider, and better controlled by managment.

Design of Organizations

Most research (Hage, 1974) has focused on analysis of
existing communication patterns. As a descriptive tool our
model allows us to recognize and analyze communication needs
within an organization. In a normative sense, however, the
clustering shows us an ideal grouping of individuals so as
to maximize the opportunity for face-to-face contact.

---

3. For example, in 1971, 191 branches of the Department of
the Environment in London, England, were clustered to
identify overlapping roles.

Now suppose the 'clustered' organization differs widely
from the organization chart. What we have done is expose the
very innards of the working organization, and how it operates
and communicates. How then can the cluster structure provide
information on restructuring the organization? Assuming that
maximizing the opportunity for essential communication is the
objective, then the obvious answer is to physically locate the
clusters with high affinities near each other. Since some
clusters may be composed of individuals from various levels
and functions of the hierarchy, a new concept of organization
is necessary. The vertical and horizontal concepts would be
replaced by a multi-dimensional relationship (somewhat simpli-
fied in Fig. 4).

Difficulties

Some further questions arise:

(i) there are phychol0gical problems in the transition from a
     'hierarchial' to a 'clustered' organization (i.e., loss of
     identity, status).
(ii) since seniority within the cluster and between

     clusters is not explicit, a formal chain of com-

     mand may need to be specified.

(iii) a dynamic organization may change so quickly that

     frequent reclustering is necessary.

(iv) there is an unclear relationship between communi-

     cation requirements and proximity .[4]

---

[4]See, for example, the work on the communication-distance relation-
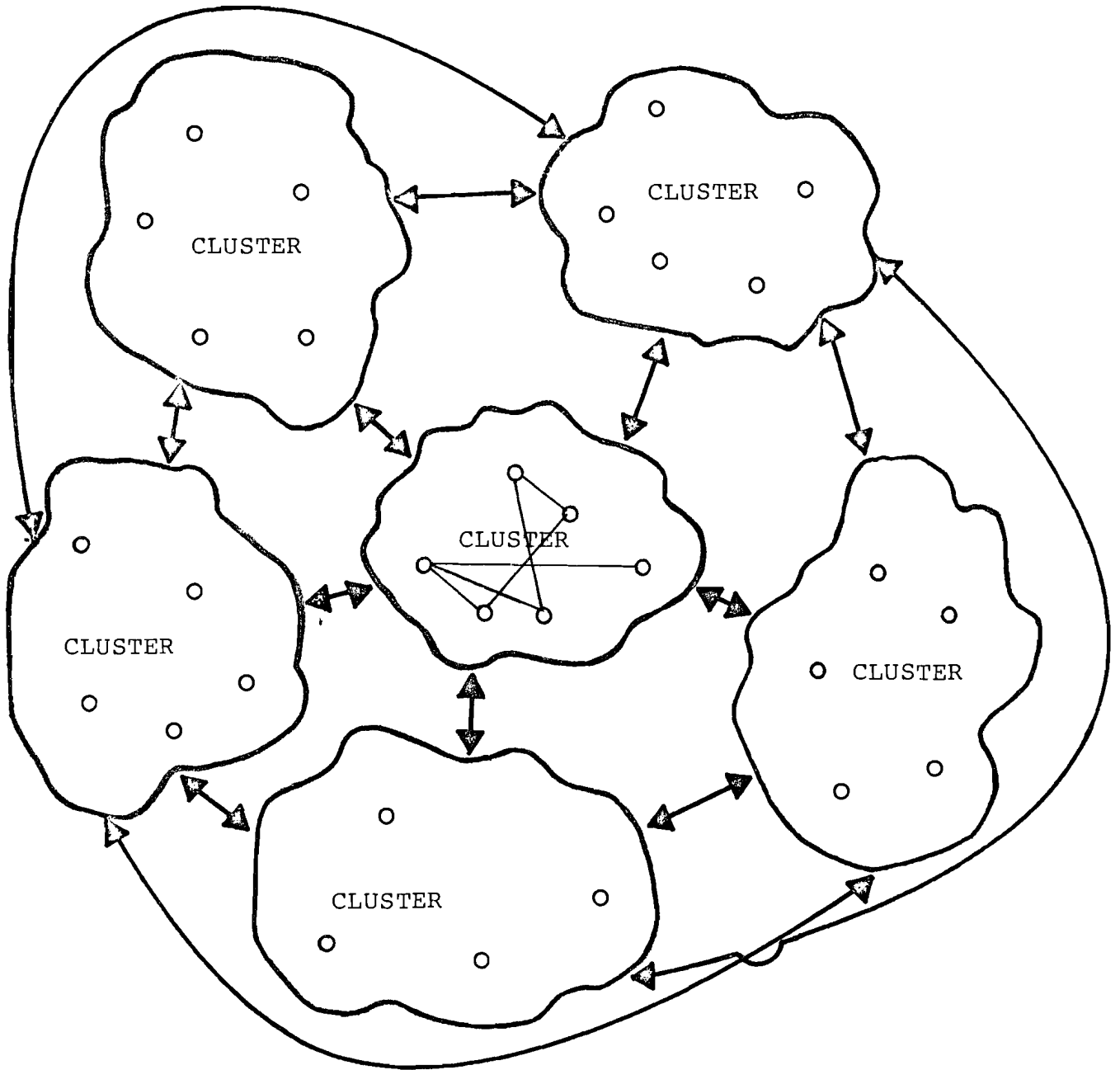ship in Ostrom and Thomas, 1972.

FIGURE 4.   MULTIDIMENSIONAL ORGANIZATIONAL CLUSTERS

(v)   there is difficulty in deciding on the number of clusters c. Trivially, the best solution for all c is to put everyone in the same cluster. While the F-test mentioned earlier may help somewhat to find the correct c there is a need for organization-specific bounds on c. These bounds would depend upon the size if the organization, type of work carried out and

some feeling as to the maximum number of communications that an individual can handle.

(vi)   There is perhaps a need for manual 'updating' of clusters as people come and go. Clearly, the complete questionnaire and reclustering process cannot be carried out for each individual that enters or leaves the organization. Hence some manual process is required, which although based upon incomplete information, still provides an adequate assignment until the next major clustering.

(vii)   there is difficulty with visual representation of the clusters in n dimensions. Some combination of visual and numerical relationships is required.

(viii)   the communication 'needs' upon which the clustering is based may have subjective elements in the personal opinions of each individual; some screening in the questionnaire process may thus be necessary. There is no doubt that the validity and accuracy of the raw data is a key issue and much depends upon the questionnaire techniques and consistency checks.

(ix)   the conflict resolution process (i.e., resolving

conflict when $a_{ij}^1 \neq a_{ji}^1$) is as yet undefined; the final a_{ij} used could be the mean communication between 2 individuals, or determined by some other technique e.g. 'truth table', below:

TRUTH TABLE FOR CONFLICT RESOLUTION

$$a_{ji}^1$$

	0	1	2	3	4
0	N/A	1	2	3	4
1	1	N/A	2	3	4
2	2	2	N/A	3	4
3	3	3	3	N/A	4
4	4	4	4	4	N/A

a_{ij}^1 (row label)

(x)   there is some question if 'desired face-to-face communication' is the only criterion upon which to cluster; other criteria might include 'common interests', 'ease of communication', 'common facility' (i.e. two individuals may use the same computer terminal). In the last case we have introduced inanimate facilities into our clustering with humans.

(xi) another consideration is the 'importance' of communication. In one study, this criterion was weighted with 'likely unavailability', i.e., a communication was downgraded if an individual were frequently absent from his workplace, and particularly so if another individual could take the message and act on it.

(xii) does a high VI for an individual within a cluster indicate strong value or strong redundancy? An ambiguous questionnaire might not differentiate between meaningful contact between two individuals and communication between a redundant go-between and an individual.

(xiii) individuals may belong to more than one cluster according to different functional roles of the organization. Should the organization be clustered independently for each role?

(xiv) how do we compare clusterings in terms of structure, flexibility, adaptability, etc.? Is there a 'measure of hierarchy'[5] with which we can compare our objective function T.

(xv) how good should the clustering be? Here, we are faced with a tradeoff between computing time and sub-optionality as mentioned above when introducing block moves.

·Conclusions

The formal organizational hierarchy and its complex communication structure can be better understood using cluster analysis techniques. The method introduced in this paper

---

[5] 'Entropy' has been proposed as a measure of disorder in an organisation and Clough (1964) has shown how to employ this parameter in hypothesis testing.

allocates individuals to clusters so that as far as possible, those people with close affinities lie in the same cluster. The resultant clusters and the links between them allow us to quickly understand and control the functional relationship between communicating units of an organization. Additional work, however, is required on the concept and design of a 'clustered' organization, and the problems of transition to it.

REFERENCES

[1]   Beale, E. M. L., "Euclidean Cluster Analysis," Bulletin
      of the International Statistical Institute, Pro-
      ceedings of the 37th Session, Vol. 43, Book 2, 1969.

[2]   Ostrom, A. R. and Thomas, G. S., "Office Layout," Sci-
      entific Control Systems Ltd., London, U.K., Sept.
      1972.

[3]   Cluster Analysis, Scientific Control Systems Ltd.,
      London, 1971.

[4]   Hage, J., Communication and Organizational Control,
      Wiley, New York, 1974, p. 163-189.

[5]   Clough, D.J., "Application of the Principle of Maximizing
      Entropy in the Formulation of Hypotheses," CORS
      Journal, Vol. 2, No. 2, December 1964, pp. 53-70.