Computer-Based Message Services
H.T. Smith (Editor)
Elsevier Science Publlshers B.V. (North-Holland)

You Have 134 Unread Mail!
Do You Want To Reed Them Now?

Jacob Palme
QZ Computer Center
Pox 27322, 102 54 Stockholm
Sweden

Electronic mail system can, if used by many people, cause severe information overload problems. The cause of this problem is that it is so easy to send a message to a large number of people, and that systems are often designed to give the sender too much control of the communication process, and the receiver too little control. The solution to the problem must be too increase the control of the receiver. To do this, structure Is needed on the set of messages. Electronic mail systems thus need to be more data base oriented, like some computer conference systems already are.

The problem

In many large computer message systems or networks, one of the major problems already is that people get too many messages, which they do not have time to read. This also means that the really important messages are difficult to find in a large flow of less important messages.

In the future, when we get larger and larger message systems, and these systems get more and more interconnected, this will be a problem for almost all users of these systems.

If electronic message systems are to succeed, we must find a way of overcoming this problem. This paper discusses the problem, and ways of overcoming it in different existing electronic message systems.

The Cause

In order to handle this problem, we must first understand its cause.

The average time of writing a message (according to statistics on our COM system, see Palme [1] is 3.6 minutes, and the average time of reading a message is 0.47 minutes. Thus, if every written message was sent to one receiver, people would spend eight times more writing messages than reading them. Some very few very popular people would in such a situation get too many messages, but the average user would certainly not be overloaded with messages.

If, however, the message system allows the sender of a message to send copies of the same message to many receivers, the odds will change. In many systems, the time to write a message to one hundred receivers is not any longer than the time to write a message to one single receiver. Thus, with only 3.6 minutes of work to write a message, its author can cause 0.47 minutes of reading time for one hundred receivers, or a total of 47 minutes of reading time for all its receivers. Obviously, this will easily mean that receivers get more messages than they can cope with.

Thus, the problem of people getting too many messages is closely connected with the facilities of message systems to easily distribute the same message to a large number of receivers. In ARPANET and CSNET, for example, this problem is severe because these networks have a large number of

distribution lists where a message sent to a distribution will be sent to all people on the distribution list. USENET has the same problems for its distributed conferences.

Another way of explaining the problem is to say that many mail systems give too much control over the communication to the senders of messages, too little control to the receivers:

All control with the sender - Balance of control - All control with the receiver

Electronic mail system | Computer conferencing system | Typical information retrieval system

By designing CPMS-es to shift the control more to receivers, less to the senders, the information overload problem can be overcome.

Do not forbid multi-receiver messages

Since the cause of the problem is that it is so easy to send messages to many receivers, one solution might be to forbid messages to many receivers. This is however a bad solution. There is a need for messages sent to many receivers. Many systems have a facility called "distribution list" or "bulletin board" or "computer conference", through which the sender only needs to give the name of a group of receivers, in order to get a message sent to all members of the group. In the rest of this paper, the word "group communication" will be used for this facility. It is very is popular and widely used.

By sending messages to many receivers, a communication process involving many people Is created. And computer message systems can, with better design, be very useful for communication between many people.

Suppose you have a need to communicate in a group of 12 people. The total time for all 12 participants is shown below (Turoff [2], Palme [6]):

Computer message system: Longer writing time but shorter reading time

Writing	Reading	Total time
3.6 min.	11 times 0.47 = 5.2 min.	3.6 + 5.2 = 8.8 min.

Face-to face or telephone/video meeting: You talk faster than you write, but you listen slower than you read:

Total time for talking and listening:
12 times 1.7 = 20.4 minutes.

Communication through a computer message system is thus more efficient with time, and this will be more pronounced as the group size increases. If the time and cost of travel is included, the message system is of course even more efficient.

The reason why the reading time is shorter in the computer message system is not only because people read faster than they listen, but also because a computer message system allows every participant to decide how much time to spend on each message. You can read carefully items of importance and skip items with information you already know or which is of no interest to you.

This difference is not only an efficiency factor. It is also important psychologically. With twelve participants, as in the example above, every person uses about a third of his/her time giving information and about two thirds of the time receiving information, in the computer message system. In an ordinary face-to-face meeting with 12 participants, they would on average talk 8 % of the time and listen g2 % of the time (Palme [6]). Communication can work psychologically better with computer message systems, because you are not forced to be a passive listener as much as in face-to-face meetings. This also means that computer message systems can work well even in group sizes of 30-100 people which would be very difficult to manage in face-to-face meetings, provided the problem with information overload can be solved.

A typical situation in a face-to-face meeting with 12 participants is that one person is talking. Some other persons are listening very impatiently, saying to themselves: "Does he

have to say the same things I have heard ten times before. The meeting is already late, and I have other things to do." Put at the same time, other participants at the same meeting may find the same presentation very valuable - they have not heard it before.

Another side of the same coin is that a talker may not say what he wants to say, because he knows that some participants have heard it before and want to go somewhere else. Put this may mean that other participants do not get information which is important to them and which they have not heard before.

Very common in face-to-face meetings is that time is not enough to take up all you want to discuss, and people have to suppress comments which might have been very valuable. This seldom happens in computer communication systems.

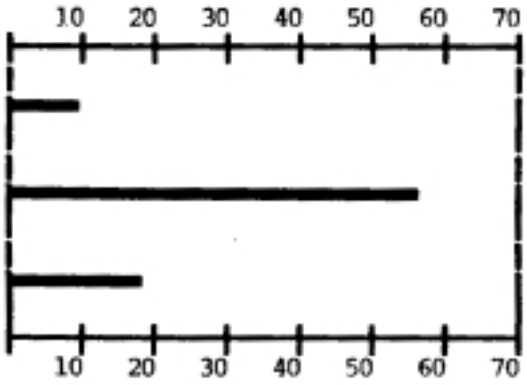
Thus, because of the shorter reading time and that you easily can skip messages you are not interested in, computer message systems can be very efficient media in larger groups.

Compare the following times to communicate the same amount of information to all the participants in a group (Palme [6]):

To 5 people in a face-to-face meeting: 9 minutes

To 33 people in a face-to-face meeting: 56 minutes

To 33 people with a computer message system: 16 minutes



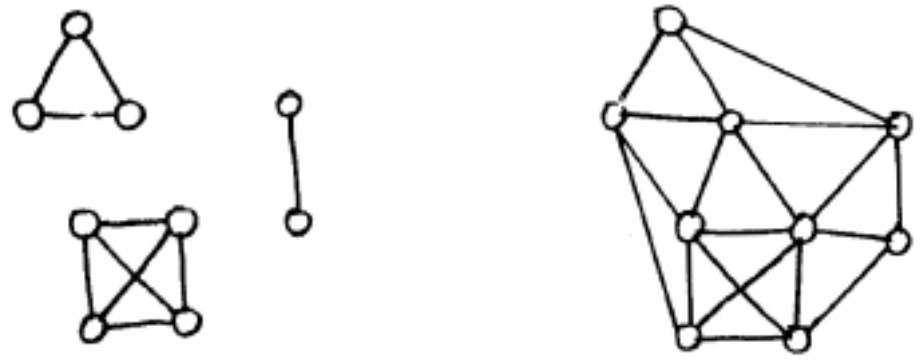
Increasing the group size to 33 people becomes Prohibitively inefficient with a face-to-face meeting, but not so with a computer message system - if we can solve the information overload problem.

This table shows how much of the communication in a large research institute using the COM computer conference system which went between people who were close and distant in the organization (Palme [61] :

	Using the mail facility	Using the conference facility
Communication between people within one department	77 %	36 %
Communication between people in different departments	23 %	62 %

These result shows that there is a difference between who communicates with whom using the mail and the conference facility in the system. The mail facility gives more communication between people who are close geographically or organizationally and who know each other well. The group communication facility gives more communication between people who are far away and do not know each other. The reason for this is that the sender of a conference entry does not have to think of the names of all the people who are to receive the entry.

A CPMS with a group communication facility provides an environment where people can "meet" and exchange ideas much more freely than in a pure mail system (Hiltz [2]). Contacts between people who did not know each other before are much easier to establish with a group communication facility than in a pure mail system. A system with group communication facilities will much more easily provide cooperation and a feeling of togetherness between widely dispersed people. People who regularly use computer conferencing say that a whole new dimension of contacts and communication has opened up for them, and that they cannot understand how they were able to live in the seclusion before they started using the system.



Mail system Conference system

What then, is the effect of the increased number of contacts at large organizational and geographical distances, which systems with group communication facilities give. This has been studied in sociological research (Allen [1]). This research shows that people having such contacts are more successful. They tend to be less conservative, they will easier accept new ideas and they are less prone to get stuck with bad or suboptimal solutions to their problems.

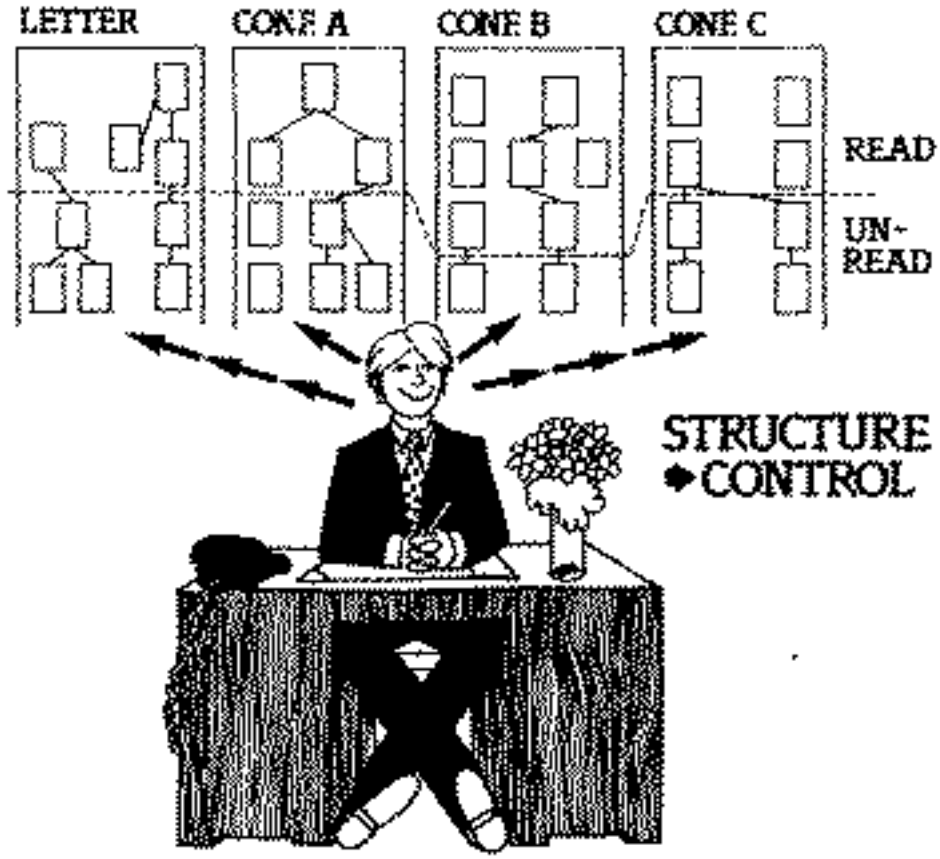
Thus, even though sending the same message to many receivers can cause information overload problems, it is also a very valuable and useful facility. We should find ways of overcoming the overload problems but still permitting larger groups.

Control by conferences

One solution to the information overload problem is to put a structure on the set of incoming messages. Instead of delivering an unordered heap of messages, the system should deliver a neatly structured data base of incoming messages. The structure of this data base should be such that the reader can easily find those messages he finds important. It should allow the reader to decide which messages to read immediately, which to save to another time, and which not to read at all. It should be easy for a user to change these decisions as new information comes in.

There is an example of how a user of the COM computer conference system is greeted by the system (Palme [7]]:

```
You have 5 unseen letters
You have 2 unseen entries in GILT open meeting
You have 13 unseen entries in Supercomputers
You have 5 unseen entries in English language
You have 6 unseen entries in Announcement [of new) conferences
You have 19 unseen entries in Speakers corner
You have 19 unseen entries in Presentation (of new) COM users
You have 1 unseen entries in Fifth generation computer project
You have 18 unseen entries in Packet-switched network use
You have 11 unseen entries in Microcomputers
You have 5 unseen entries in TeX inter-network mailing list
You have 1 unseen entries in KEPMIT experience
You have 34 unseen entries in TOPS-10/20 SIG
You have 134 unseen entries
```



Every message which was sent to the user via a computer conference (distribution list, bulletin board) is also delivered to him as an entry in that conference. The user normally will read one conference at a time. The users decide themselves in which order to read the conferences, and they can save some conferences to read at a later time. If they get

too much information, they can also withdraw completely from conferences they are not interested in, or skip part of the discussion in that conference but still stay as a member.

A computer conference system also allows messages which are not sent as conference entries. Put the systems usually intentionally are designed to make it difficult to send such a message to more than a few receivers.

Thus, the amount of messages which are not sorted into conferences is kept small. Instead, the systems make it easy to create new conferences as the need arises.

A computer conference has an organizer, who can remove messages which do not fit to the subject of the conference. Thus, the organizer helps the participants to control what they receive by ensuring that they get messages on the subject they have chosen when participating in the conference. In the COM system, the organizer can not delete text entries. The organizer can only remove the link between the entry and the conference, and optionally add a link to another conference more suitable to the contents of the entry.

Control by comment trees

Another way of structuring messages is by comment trees. A system can be designed to store relations between messages, where one of them can be a comment or a reply to another message. Thus, a set of messages which refer to each other directly or indirectly can be identified automatically by the system. Such a set of related messages can be called a "comment tree". It is a grouping of messages, just as a computer conference. And in the same way, the receivers of messages can be given the facility of choosing in which order to read the different comment trees, and to skip messages in a comment tree of less interest.

In the COM computer conference system (Palme [7]), comment trees are used to structure those messages which do not belong to conferences, and comment trees are also used as a sub-structure within conferences.

Control by keywords

Yet another solution is to affix keywords to messages. The system can then be told to deliver messages according to their

keywords, thus giving the reader more control of what to read and not to read. A problem with this solution is that it can be difficult to get the senders of messages to assign well-chosen keywords to their messages.

Note that keywords and computer conferences are very similar concepts. This is especially so in the COM and PortaCOM computer conference systems, since in those systems one and the same entry can be linked to more than one conference (Palme [7]). Thus, the set of conferences for a CON message is very similar to the set of keywords in a keyword-based system. In other systems, like the ETES system, keywords and conferences are kept as two separate concepts which can both be used by readers to select which messages they want to read.

Control by subject

Another way of controlling communication is to select messages by subject. Again, this is rather similar to computer conferencing, where all messages with a certain subject can be seen as a kind of conference.

Just like in conference systems, it would be valuable to be able to read all messages on a certain subject before continuing with a new subject.

In the EIES computer conference system, there is a facility called TOPICS in which every new subject taken up in a conference becomes a new sub-conference on that subject. Every member of the main conference decides whether or not to participate in the subconference.

The experience from EIES is that this facility is very efficient in reducing communication. In fact, it is so efficient that it can easily kill a conference by splitting the participants into too many small subconferences, and thus reducing communication so much that people

stop participating in the main conference. Thus, the people behind EIES recommend use of the TOPICS facility for very large and too active conferences where too much is written for each member, but they do not recommend this facility normally for normal-size conferences.

In COM, comment trees as sub-conferences work in a similar manner, but every member of the main conference becomes a "member" of the subconference unless they explicitly give a command to skip that subconference. This design will not

reduce communication so much as the TOPICS facility in EIES.

Control by selection

One way of controlling communication is to have some people select messages. Other people can then read only the selected messages. This is thus similar to editors in magazines who select what to publish.

In the COM computer conference system we have write-protected conferences, where only certain people can link entries to the conference. Other people must thus first send their entries to one of the editors, or to a conference for submitted papers. The editors then decide which messages can be linked to the write-protected conference.

Write-protected conferences can also be used to contain a selection of the most important entries from ordinary open conferences. COM also has a special kind of conference to which no one can send messages directly, but anyone can link messages indirectly. Anyone can link a message s/he reads which is especially interesting to such a conference. They have been very useful.

Both these kinds of conferences get very few messages compared to ordinary open conferences, and are thus a good selection criteria for those who only want to read a small selection of the most important items.

Control by author

Finally, messages can be selected by author. This can be done in several ways. One crude way would be to allow a person to tell his system "I do not want to read any more messages written by John Smith".

Other ways of selecting by author is to divide the user population into groups, so that a reader can select only messages by authors within certain groups. The write-protected conference, as described above, can be seen as such a facility, since the editors can write directly to the conference, but no other users.

In COM, a comment on an entry in a write-protected conference is automatically furthered to a specially designated super-conference which is not write-protected.

Group selection can also be used so that only experts in a

certain field can write in a conference, but other people may read their discussions and comment on them in a super-conference. COM has such a facility. Or one might select by CBMS. For example, ARPANET-CSNET might for some of their mailing list allow people outside ARPANET-CSNET to read, but not to enter messages to that particular mailing list.

Instead of just skipping messages by certain authors, it is better to further these messages to special structures, so that those who want to

read them can read them there. For example, the so-called "postmaster" conference often gets many messages, which can automatically be sorted by categories and sent along to different conferences depending on who is interested in reading them.

Selection by abstract writing

Finally, some people could abstract the discussions in voluminous open conferences into write-protected conferences containing only the abstracts. Such abstracts have been very useful in the ARPANET-CSNET community. In the EIES TOPICS system, an abstract of each subconference is meant to be entered into the main conference.

User interface aspects

To reduce information overload, we need structuring on the message set. This structuring must be based on information input by the writer of the messages, by someone else (e.g. assigning keywords) or automatically by the system. Having special people assigning keywords to all messages in a large mail system is not practical.

Important is therefore to use such information which we can easily get the writers of message to input in a reliable manner. Useful is also if someone else can correct mistakes by the writers, like the conference organizer who moves entries to another conference when needed.

The COM system is intentionally designed to make it easier to input a comment on a previous entry than to input a non-comment, just because the comment link is useful structuring information. This is an example of how the user interface can be designed to further structuring.

Future development

In the future, we can expect larger and larger systems and networks of systems. New structuring facilities will then be needed. Probably what will have to be introduced is a facility to divide large discussion groups into subgroups, where only selected messages or abstracts of the discussions in the subgroups are made available to all participants in all groups.

Conclusion

Computer-based message systems are especially good for communication in large groups, where they can widen horizons and give more people more information and contacts. Efficient methods of allowing the readers of messages to control what they get will actually enable communication in larger groups than without such methods, and will thus make the message systems more valuable.

References

[1] Allen, Thomas J.: Managing the Flow of Technology. MIT Press 1977.

[2] Hiltz, Starr Roxanne and Turoff, Murray: The Network Nation: Human Communication via Computer. Addison-Wesley 1978.

[3] Hiltz, Starr Roxanne and Kerr, Elaine B.: Studies of computer mediated communications systems: A synthesis of the findings, New Jersey Institute of Technology, August 1981.

[4] Hiltz, Starr Roxanne, Johnson, K., Aronovitch, C., Turoff, M.: Face-to-face vs. computerized conference: A controlled experiment. New Jersey Institute of Technology, 1980.

[5] Kerr, Elaine, and Hiltz, Starr Roxanne: Computer-Mediated Communication Systems. Academic Press 1982.

[6] Palme, Jacob 1981: Experience with the Use of the COM Computerized Conferencing System. Swedish National Defense Research Institute, Pox 27322, S-102 54 Stockholm, Sweden, 1981.

[7] Palme, Jacob 1.983: COM/PortaCOM conference system: Design goals and principles. QE Computer Centre, Pox 27322, S-102 54 Stockholm, Sweden, 1983.

[8] Palme, Jacob 1983: Computer Conferencing is More than Electronic Mail. Transcripts of the EUTECO, European Teleinformatics Conference, North-Holland, 1983.

[9] Palme, Jacob 1984: Survey of computer-based message systems. Transcripts of the INTERACT '84 first IFIP conference on human-computer interaction. North-Holland, 1984.