36 • Computer-Mediated Communication: A Network-Based Content Analysis Using a CBBS Conference

JAMES A. DANOWSKI

University of Wisconsin—Madison

COMPUTER-MEDIATED communication (CMC) systems such as private messaging, public conferencing, bulletin boards, and several variations in between are proliferating. Researchers are increasingly interested in measuring and evaluating CMC. On the one hand, some of the features of CMC provide a testbed for theoretical propositions about human communication. CMC provides a highly controlled, time-sensitive, and automatically audited environment. This not only enables testing of hypotheses about CMC in relation to other variables, but also enables testing of more general propositions about human communication that are not dependent on the type of medium used. On the other hand, the implementation of CMC systems in organizations can benefit from practically guided formative and summative evaluation research.

Based on these dual research needs, the objective of this chapter is to develop and present a methodology for both theoretically and practically oriented research that is particularly attuned to the nature of CMC. This methodology integrates communication *network* analysis perspectives and procedures, performs *content analysis* on the relationships among

AUTHOR'S NOTE: The author acknowledges the research assistance of Joanne Taeuffer in development of the methodology and of Ilya Adler, Rosa Elizondo, Charles Steinfield, and Eric Wendler for data gathering and processing. This research was supported by the National Science Foundation, Division of Information Science and Technology, IST-801088. Views expressed do not necessarily reflect those of NSF.

Correspondence and requests for reprints: James A. Danowski, School of Journalism and Mass Communication, University of Wisconsin, Madison, WI 53706.

concepts in message pairs, represents the aggregate message content relationships with a *distance* model, and enables derivation of *optimal communication management strategies* over time. Application of the methodology is illustrated with data obtained from a public computer conference, a CBBS in Boston.

SPECIAL FEATURES OF CMC

As Rice and Danowski (1981) discuss, the research techniques generally appropriate for evaluating CMC are no different from those appropriate for evaluating most other human activity. Depending on the stakeholders involved and their perspectives, an adequate mix of methods can be chosen from a well-stocked master toolkit of social science and evaluation methods. Why, then, propose a special CMC methodology? Some important CMC features point to research needs not easily met by "off the shelf" methods. This chapter shows that several aspects of existing methods may be linked, resulting in an enhanced research capability. This methodology, however, is not intended to replace others. In contrast, it may be one element of a larger methodological constellation. Four key features of CMC motivate the development of the enhanced method.

(1) *Communication networks.* It is widely known that a network perspective analyzes the structure of message exchange among a set of nodes — for example, individuals or groups. Separate network analyses can be performed by topics, by media used, by strength of links, and by other factors. Given the general value of network perspectives, one may wonder why network aspects of CMC are singled out as a basis for further methodological development. The main reason is that network traffic can be efficiently gathered on most systems. This reduces some of the barriers to network analysis of other communication behaviors, such as those via face-to-face modes, for which data are typically difficult to obtain, hard to code and clean, and filled with error. In contrast, conference network traffic can be captured in an automated fashion, at low cost, virtually error-free, with time-sensitivity, and without extensive manual coding and data entry. One particular approach to this will be discussed shortly.

(2) *Message content.* It can be argued that the most central aspects of human communication processes are the *messages* people exchange and the meanings they attach to symbolic message elements. Although networks of message traffic are important, they may be viewed as the accumulated traces of repeated message *content* exchange. In addition, even though the medium used is important, the main reason for its use is typically to exchange message content. And, while participants' individual differences are often significant, it is message content that communicates and bridges these differences.

Perhaps CMC message content has enhanced importance for several reasons. Each message typically has a standard header including such things as sender, receiver, time, date, and subject. These fixed formats may heighten user content awareness. Furthermore, CMC messages are visible, retrievable, and controllable by users, more so than with media such as newspapers, letters, and memos. Moreover, CMC messages are typically more personal than those in mass media. In short, for both general theoretical and specific CMC-related reasons, an appropriate methodology should focus on message content.

(3) *Time-sensitivity.* Users comment that some forms of CMC, particularly public conferencing, seem to have different "life-cycles." Speaking metaphorically, a conference can be viewed as wiggling, stumbling, and crawling about at first, then growing rapidly, experiencing identity crises, later maturing, and finally dying. While measurement of change over time in human processes is generally thought to be important, given these life-cycle notions, methods for CMC analysis should be particularly time-sensitive.

(4) *Leadership.* CMC participants have often commented that leadership is especially important. This may be because of greater coordination needs than for face-to-face communication, arising from asynchronous communication, users' reduced sensory engagement, their greater diversity, and other factors. These point to the need for a methodology that will enhance leaders' control over the course of conferences.

In sum, the methodology presented in this chapter is responsive to the four CMC features just described. It integrates *network* analysis perspectives and procedures, performs *content* analysis on the relationships among concepts in message pairs, represents the aggregate message content *relationships* with multidimensional scaling techniques, and enables derivation of *optimal communication management strategies* over time. The scope of application of the method is briefly discussed in the next section.

Scope

The utility of the current method ranges from basic communication research applications to practical conference management. It can address questions about and enable testing of scientific questions and hypotheses about change over time in CMC and related variables. For example: "How is change in the message content exchanged associated with change in the communication network structure?" Or, "What are the major life-stages of conferences and to what extent are these fixed by external factors?"

More practical conference management evaluations require two general kinds of applications. One is *formative* in nature: "How can the course of a computer conference be shaped as it occurs?" The second application is more *summative.* After a conference has lived out a normal life, one may ask: "How well did it accomplish its objectives?" Our methods enable these practical evaluation applications as well.

Basic Procedures

This CMC methodology has the following major components:

1. segmentation of communication activity by communication *network* structure
2. segmentation of communication activity by *time*
3. identification of message *content* elements; that is, *concepts*
4. identification of stimulus and response *message pairs*
5. tabulation of concept co-occurrence within message pairs, aggregated across all message pairs in a segment
6. multidimensional scaling of the aggregated co-occurrence matrix to identify the overall pattern of *relationships* among message elements

If derivation of communication strategies for changing the course of the CMC is desired, then additional steps follow:

7. identification of which concepts should be moved closer to or further from other concepts
8. derivation of *optimal messages* (combinations of concepts) to achieve the desired change
9. *entry* of optimal change messages into the conference

After step 7, the process begins again with step 1. Comments about some of the above steps are in order.

Segmentation. For many decades, it has been known that different social groups communicate differently. Because of this, communication participants, typically mass communication audience members, have been divided into subsets homogeneous within but different across. The first segmentations, starting in the 1930s, were based on *demographic* or structural locator variables such as income, education, age, sex, race, and so on. During the 1960s and 1970s, however, as communication participants appeared to develop increasing lifestyle and attitudinal differences that cut across demographic factors, *psychographic* segmentation gained prominence.

An even more refined *infographic* segmentation (Danowski, 1976) can be performed according to actual communication behaviors. A range of these may be used, including such things as the *network* structure of

communicators, the *media* used, *message* variables, and *information processing styles.*

In a particular research situation, the choice of specific demographic, psychographic, or infographic segmentation strategies should depend on the evaluation objectives. For example, the more a communication program is concerned with disseminating information, the more useful *infographic* segmentation is. If participants are grouped and analyzed according to their information exposure, processing styles, and subsequent "second stage" dissemination potential, then it is more likely that overall program objectives will be achieved. In evaluating CMC, infographic segmentation according to communication network variables appears particularly useful. Most CMC systems enable users to engage in series of overlapping dyadic relationships. These define a dynamic communication network among users.

An additional segmentation variable with particular relevance to CMC is *time.* Along with conference "life-cycle" aspects, reasons for time segmentation include the more general value placed on "over time" analysis in the social sciences. Measuring variables over time can reveal underlying causal sequencing among variables and more accurately reflect *processual* dynamics. CMC presents unique possibilities to segment both by time and network variables because CMC software codes each message entry according to time, sender, and receiver.

Message Content Parsing. Content analysis of message concepts is a focal point of our methodology. Particular approaches to isolating message concepts can be tailored to the research objectives. Various computerized and manual procedures exist for performing textual content analysis.

Concept Co-occurrence. Unlike traditional content analysis, our method does not simply identify the atomistic occurrence of concepts, that is, treat concepts as independent and separate. Rather, it maps the *relationships* among concepts by indexing the *co*-occurrence of concept *pairs.* Moreover, rather than selecting messages as the units within which to observe concept pairs, we select *pairs* of messages. This choice is based on the assumption that communication requires at least two participants, and that a communication event is constituted by a message sent and the response it triggers.

Consider the following hypothetical message pair. User A enters a public conference message about an upcoming user group meeting and offers a new software package he wrote. Subsequently, User B responds by asking User A to send him or her the software, but also requests information about User A's disk drives. There are two concepts in message A of the pair: (1) user group meeting information, and (2) offer of software. Two additional concepts are in message B: (3) request for software, and (4) request for hardware information.

Message A	Message B
Concept 1 (c1)	Concept 3 (c5)
Concept 2 (c2)	Concept 6 (c6)
Concept 3 (c3)	Concept 7 (c7)
Concept 4 (c4)	
Concept 5 (c5)	

CO-OCCURRENCE MATRIX

	c1	c2	c3	c4	c5	c6	c7
c1	0	0	1	0	0	1	1
c2		0	1	0	0	1	1
c3			1	1	1	1	1
c4				0	0	1	1
c5					0	1	1
c6						0	1
c7							0

Figure 36.1. An illustration of computation of message concept co-occurrence. The example analysis is for a single message pair. Note that the co-occurrence scoring is performed for each pair of messages from person A to person B and an aggregate matrix is created across the entire network or network segment within a time segment.

Consider the co-occurrence of these concepts across the two messages in the pair. Concept 1 co-occurs with concept 3 and concept 4; concept 2 likewise co-occurs. Each of these concept pairs (1-3, 1-4, 2-3, 2-4) receives a co-occurrence score of 1. If these pairs co-occur in other message pairs in the conference segment, their scores would be incremented accordingly. Note that co-occurrence of concepts within one message is not counted. If it were, then individual-level content structures, rather than group-level structures, would be measured. Figure 36.1 graphically presents the basic co-occurrence procedure with another, more abstract message pair.

Concept co-occurrence mapping, aggregated across message pairs, represents two aspects of the communication process. One, it reveals the manifest conversational structure among participants as it appears to an external observer. Second, to some extent it indirectly represents the collective cognitive structure among participants. This second aspect merits further discussion. First, the concepts co-occurring are not necessarily those in direct response to concepts initiated in the first message of a pair. This is exemplified by concept 4 in the earlier example. User B added the concept "disk drives." The ability of our methodology to measure these indirect relations presents opportunities to observe aspects of the underlying psychological structure among participants.

Consider that over a number of message pairs, the appearance of the same concept co-occurrences, even if at first glance they seem unrelated, indicates that a regularity exists in participants' underlying psychological processes. On one hand, there may be a kind of facilitative semantic trigger effect; one concept tends to positively elicit another concept. Alternatively, there may be a kind of compensatory trigger effect; one concept appears in response to another because the first does not create sufficient positive feelings. Hence, seemingly unrelated concepts emerge as the participants change an undesirable subject to a more pleasant one. Such underlying psychological processes can be investigated by linking the present methodology with self-report-based cognitive mapping.

MDS. Multidimensional scaling (MDS) of concept co-occurrence matrices, aggregated across message pairs in a segment, effectively represents the overall relationships among concepts. There are numerous nonmetric and metric scaling algorithms from which to choose (Kruscal & Wish, 1978). If, however, an evaluation objective is for leaders to extract optimal messages, as indicated earlier by steps 7-9, then it is advisable to use one particular kind of metric scaling procedure known as GALILEO (Woelfel & Fink, 1980). This approach has the Automatic Message Generator (AMG) algorithm programmed to select optimal combinations of concepts for inclusion in subsequent messages. Application of AMG will be illustrated shortly.

ILLUSTRATION OF THE CMC EVALUATION METHODOLOGY: BOSTON CBBS CASE

CBBS

The CMC system we chose to study as an illustration of the methodology is a public computer conference operating in the Boston area using a Computerized Bulletin Board System (CBBS) conferencing software

```
#1 Terminal need nulls?  Hit control-N while this types:

            ***  Welcome to CBBS/Boston  ***
*** New Englands 1st Computerized Bulletin Board System ***
                 [System up since 12/2/78]

----> Control characters accepted by this system:

   CTL-H/DEL  Erases last character typed. (And echos it)
   CTL-C      Cancel current printing
   CTL-K      'Kills' current function, returns to menu
   CTL-N      Send 5 nulls after CR/LF
   CTL-R      Retypes current input line (after DEL)
   CTL-S      Stop/start output (for video terminal)
   CTL-U      Erases current input line

Problems?  Try calling the following numbers:
[illegible]:  (61 ) 75[illegible] Rm. [illegible], [illegible]
[illegible]:  (61 ) 9[illegible]

Bulletins: Last updated 04/28/79, 14 lines.
(Hit multiple control-c's to skip this...)
]--> 04/28/79 Thanks to CBBS user LEO KENEN for solving a perplexing
              CPM problem... We now are running 48K CP/M!
]--> 03/26/79 New IDS modem installed, while other IDS modem is
              out being repaired...
]--> 03/10/79 CBBS phone numbers moved into messages and out of
              Bulletins (were too long..)
]--> 02/24/79 Second Shugart SA800 now online.. We'll now be able
              to handle up to 540 online messages!
]--> 01/25/79 Now running with SD Systems 48K ExpandoRAM.
]--> 01/09/79 We thank Tarbell Electronics for their donation of a
              disk controller.
Note: When we say C/R, we mean your return or newline key!

Y/N: IS THIS YOUR FIRST TIME ON THE SYSTEM?N
What is your first name?ROSA

Logging name to disk...
Next msg # will be 228
You are caller #  3234

FUNCTION: B,C,D,E,G,H,K,N,P,Q,R,S,W,X (OR ? IF NOT KNOWN)
```

Figure 36.2. Boston CBBS sample transcript: log-on and commands.

package (Christensen & Suess, 1978). CBBS systems are basically very similar to other conferencing systems such as EIES (Hiltz & Turoff, 1978), CONFER (Parnes, 1980), and others. Users log into the conference and read earlier entries, make entries, list summary header information on prior ones, search for and retrieve them. Sample CBBS transcripts appear in Figures 36.2 and 36.3.


```
MSG 115 IS 08 LINE(S) ON 03/04/79 FROM ROLF ROSENGREN TO ALL
ABOUT BACKGAMMON FOR N.S. OR SOL 20

PLEASE SEND FOR A FLYER BACKGAMMON FOR N,S.OR SOL-20 TO:
RR ELECTRONICS P.O. BOX 384 PARK RIDGE N.J. 07656
WILL RUN ON A CRT OR PRINTER DELUX GAME
PLACE YOUR CHIPS ANYWHERE YOU WANT UP TO 50.
THE COMPUTER PLAYS AGAINST YOU.
THE COMPUTER OR YOU CAN RULE THE DICE
MANY FEATURES THANK YOU FOR YOUR INTEREST.

MSG 116 IS 04 LINE(S) ON 03/06/79 FROM [illegible] TO ALL
ABOUT WANT GIRLFRIEND

DESPERATELY LONELY mathematician 33 wants compatible woman
18-35.  No smoking, minimal drinking.  call [illegible] or
write to [illegible], PO BOX [illegible], [illegible], CA [illegible]

MSG 117 is 02 LINE(S) ON 03/07/79 FROM CHARLIE STROM TO ALL
ABOUT EXIDY SORCERER

I AM INTERESTED IN EXCHANGING EXPERIENCES, IDEAS, ETC.
WITH USERS OF THE EXIDY SORCERER.

MSG 118 is 10 LINE(S) ON 3/9/79 FROM STEVE BROWN TO APPLE
USERS ABOUT PROGRAMS & IDEAS

     WE ARE A SMALL GROUP OF APPLE USERS IN LITTLE ROCK
ARKANSAS. WE ARE ABOUT 15 STRONG NOW.  WE ARE INTERESTED
IN IDEAS & PROGRAM EXCHANGE.  WE ARE NOT INTERESTED IN
PIRATING PROGRAMS.
CONTACT:   [illegible]
          %DATACODE
          [illegible]
           [illegible]
           [illegible]
MENTION THE BULLETIN BOARD

MSG 119 IS 06 LINE(S) ON 3/9/79 FROM ALDWEN OF THYMESWOOD TO
[illegible] ABOUT GIRLFRIEND

Aren't you looking a bit far from home for the love of your
life?  You might have better luck if you look in California...
Of course, there is the case of the two computer hackers who
were married via computer (over the ONTYME network) by this
weirdo
```

Figure 36.3. Boston CBBS sample transcript: message entries.

Most CBBSs use "Ward and Randy's" software, which is available for approximately $50. Ward and Randy's Bulletin Board, located in Chicago, was the first CBBS operating in the United States. It is still the largest and serves as a sort of informal national headquarters for over 200 CBBSs around the country. Most of the latter are accessed mainly by users in their local telephone dialing area, because calling CBBSs in other areas requires long-distance telephone charges.

The hardware necessary to operate a CBBS is quite basic: a small "home computer" with a dual floppy disk drive, a modem, and a normal telephone line. These simple requirements have aided CBBS proliferation. Someone with a home computer and the software package who wants to start a conference simply announces their telephone number, and anyone can call in and begin conferencing. The only cost to users is for the phone call. Some CBBSs are operated by user groups and other organizations, others by individuals, and yet others by computer-related merchandisers who use the CBBSs as a promotional vehicle. The message content varies some across CBBSs, but overall there is high similarity. We observed this as we read the entries on all the conferences operating at the time we designed this research, late in 1978.

There are several specific reasons that we selected a CBBS conference for the present research.

(1) *CBBS conferences represent "natural" forces in the developing "information society."* One reason is that users have no particular occupational or organizational affiliation motivating their participation. Conferencers are primarily hobbyists, the rapidly growing home computer user segment. They are motivated by personal interest. Another reason for naturalness is that CBBS use is essentially free, particularly if users reside in the local telephone calling areas of the conference; there is no "artificial" use stimulation or dampening such as might occur with systems funded by government agencies or dedicated to in-house organizational or corporate use.

(2) *Multiple conferences are occurring using the same CBBS software.* This enables rich possibilities for studying sets of conferences rather than solely individual users within a particular conference. Investigators can treat each CBBS conference, currently numbering more than 200 across the United States, as a distinct unit of analysis. There is sufficient sample size to make system-level generalizations and also to study regional and other variations in use. In contrast, most other CMC occurs on "one of a kind" systems. Many factors of these are often unique: operating software and user interfaces, dedicated purposes, cost structures, user characteristics, and so on. This makes generalizing evaluation results particularly troublesome, although less so with CBBS.

(3) *CBBSs are public conferences.* This is advantageous for the current research. We are able to access conference entries without privacy problems. Other situations may require negotiation with users before capturing their message content. Furthermore, it is not even necessary to contact the CBBS conference managers to obtain access. There are therefore reduced chances for researchers to contaminate the CMC system.

Boston CBBS

For the current research application, we chose a Boston-area CBBS. We selected this particular one because we recognized the highly developed information infrastructure in the Boston area. As a result, we expected this conference to have a sufficiently high message content diversity for a challenging test of the methodology. At the same time, our observation of all other CBBS conferences operating when data were gathered in 1978 and 1979 revealed that the Boston CBBS was representative. Another reason we chose it is that it had recently begun operation. We could therefore capture the conference in its earliest "life-stage." At the outset, message content diversity is probably higher before more routine message patterns develop. Moreover, conference managers would have undertaken less message packing, i.e., deletion of messages from the conference records.

On CBBSs, most packing appears to occur on the basis of a time criterion. For example, messages are deleted offering equipment for sale that has subsequently been sold, or announcing dates and times for user meetings that have already been held. To represent the actual, in contrast to the packed version of the conference, we logged in daily and recorded messages. Messages later packed were thus retained for our analysis, thus circumventing the packing process. Nevertheless, interesting research questions about packing processes abound with respect to alternative editing rules, message "throwaways," and other aspects.

Procedures

In analyzing the Boston CBBS, we executed the methods as follows:

(1) *Network segmentation.* Because the main purpose of the present research is to illustrate the kernel procedures — the content analysis and co-occurrence scaling — we analyzed the aggregate network structure rather than separate network groups. Moreover, network patterns were important to the identification of message pairs, as discussed in (3) below.

(2) *Time segmentation.* We selected the first 161 messages entered into the Boston conference. These began with its first operation on December 2, 1978; the 161st message was entered on February 23, 1979. Thus we had a bit more than the first ten weeks of the conference life. Again, because of our initial objective, we chose to first test the technique with one time segment.

(3) *Message pair identification.* To identify message pairs, we took each message, beginning with the last message in the series, then searched backward through the earlier 160 messages to see if the message was a response to a prior stimulus message. If so, then these two messages met a necessary but not sufficient condition to be a pair for further analysis. The identification of a potential stimulus message resulted under two conditions. One, a response message, person A's, was explicitly addressed to a particular person B. If so, we searched back and located the prior message that person B had sent that triggered A's response. Two, person A may have responded to a message person B had addressed to all conference users, not specifically to person A.

Once a potential message pair was so located, an additional sufficiency criterion was applied, namely that the two message pair candidates had to have at least one common concept. This fits conceptual definitions that for communication to occur there must be some minimal commonality in the code participants use. Moreover, it also makes the analysis of conference CMC, in contrast to private mail CMC, possible. Consider the conference situation in which an earlier message is addressed to all or a group of users rather than to a specific person. If person A's message did not reference a specific earlier sender, then without the common concept criterion, all of the generally addressed messages would be paired with person A's message. Many concept co-occurrences would be identified in error, and the resulting analysis would be highly misleading. Generally addressed messages would therefore have to be eliminated from such an analysis. Yet this is the very kind of message that distinguishes conferencing from private messaging.

In the Boston CBBS twenty-two message pairs were identified among the first 161 messages. Thirty-eight different messages were involved in these twenty-two pairs. This may seem to be a rather low degree of "networking," i.e., of users responding to the messages entered by others. Perhaps this is because this was the first series of messages entered. As a conference matures, the proportion of networking may increase up to a peak during the mid-life of the conference, then decline as it approaches the latter part of its life-cycle. These notions suggest interesting hypotheses for future research.

(4) *Identification of message concepts.* Because this is the first application of our procedures, we thought it best to use human coders in identifying concepts within messages. A coder read each of the thirty-eight messages in the twenty-two message pairs and partitioned them into the smallest meaningful concept units. A second coder repeated this process, and agreement was above 90 percent. Forty-three distinct concepts appeared.

Table 36.1
Message Concept Elements

1.	CBBS procedures
2.	Modems/couplers
3.	Request help/information
4.	Give help/information
5.	Offer information at future date
6.	Greetings/salutations
7.	Give name/address/phone number
8.	Computer software
9.	Discuss user groups
10.	Offer computer-related service/software free
11.	Computer games
12.	Leave message on computer bulletin board (this or other)
13.	Refer to earlier message
14.	Computer for the blind
15.	Express interest
16.	Source listing
17.	Computer system (other than CBBS)
18.	Hard copy
19.	Thank you
20.	Acknowledge receipt of message
21.	Discuss problems with own computer
22.	Delete this message
23.	Fantasy
24.	Ask for participation in discussing topic
25.	Will send information by other means (telephone, mail)

(5) *Computation of concept co-occurrence.* The concepts identified created a 43 × 43 concept matrix. Each cell of the matrix represents a particular concept pair. In filling the matrix, a coder took each of the twenty-two message pairs one at a time and tabulated co-occurrence scores for concept pairs within it. Each time a concept pair co-occurred in a message pair, a value of 1 was added to that concept pair's cell in the master matrix. Again, Figure 36.1 illustrates this process.

After we formed the 43 × 43 aggregate matrix, we packed it. This was necessary because the MDS program we used was limited to thirty-nine concepts. Hence we were required to remove the lowest-frequency concepts. We examined the matrix rows and columns on a concept-by-concept basis, rather than a concept pair basis, and looked for concepts that had only one co-occurrence with only one other concept. There were sixteen such concepts. These were removed, resulting in a revised 25 × 25 concept matrix. The twenty-five concepts appear in Table 36.1.

Before factoring the aggregate matrix, we reversed the cell scores so that higher numbers meant less co-occurrence and smaller numbers

Table 36.2
Distances Among Pairs of Message Concepts

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1																									
2																									
3	49	64																							
4		64	25																						
5		64	64																						
6	49	64	01	49	64																				
7		49	25	64		16																			
8		49	64			36	36																		
9			49				64																		
10			36			64																			
11			04				04	09	49	36															
12			09			16	25	49	64	64	09														
13			64	64		49	64	64	64	64	64														
14							49					64													
15		64	64			49	36	36				64		64											
16			25	25		25	16	64	49	49	04	25													
17			64		64	49							64												
18					64	49							64				64								
19			64		64	49		64					64					64							
20	64				64	49		64					64		64			64							
21	49	64	36			25	64	64	49		64	49			49	64	64		64						
22			49	64		49	64		64	49	64	64									64				
23			64			64				64		64										64			
24			49			64															64	64	64		
25							49	49							64					64	64	64	64		

Note: Cells with a value of 640, indicating no co-occurrence, are left blank for presentational convenience. The actual matrix scaled, however, includes the appropriate 640 cell values.

Table 36.3
Coordinates of Message Concepts on Three Dimensions

Concept	*Dimension 1*	*Dimension 2*	*Dimension 3*
1	125.5	6.6	−137.5
2	77.0	198.4	−52.3
3	−79.2	−140.9	5.7
4	−37.7	−56.4	87.0
5	380.0	−77.5	−9.6
6	157.5	−99.9	−109.7
7	−226.1	233.3	119.7
8	−3.0	282.1	137.4
9	−251.2	−59.1	213.4
10	−159.1	−298.6	12.7
11	−282.5	−62.6	258.4
12	−306.7	13.1	−46.5
13	146.0	−127.7	355.0
14	−59.4	188.0	−89.2
15	7.7	393.3	−162.4
16	−238.2	−29.7	201.5
17	260.9	−147.8	63.4
18	407.4	−121.8	107.9
19	260.2	−59.4	120.3
20	335.0	165.5	−31.8
21	−120.7	137.9	−69.0
22	−264.7	−170.3	−184.4
23	−95.0	−290.9	−324.4
24	−31.4	−189.9	−334.7
25	−2.2	314.4	−130.9

higher co-occurrence, just as in physical distance measurement. After reversal, the highest co-occurrence score, originally an 8, was converted to 1 and the original 1 became 8. At this point we made a further transformation by squaring the scores. We noted that in physics the mutual attraction of two objects (of equal mass) is related to the square of the distance between them. We think a parallel concept attraction function is plausible, particularly because we used a distance model. After squaring, a score of 1 remained 1, while the highest score became 64.

The final data preparation step was to assign a very large number to the pairs not co-occurring at all, those having zero cell values before transformation. Finding no standard yet accepted in the literature, we chose a rule: Multiply the highest cell entry times ten and assign this number to the unrelated pairs. In the current data, this resulted in cell values of 640 for the unrelated concepts. This transformed co-occurrence matrix, which was then multidimensionally scaled, appears in Table 36.2.

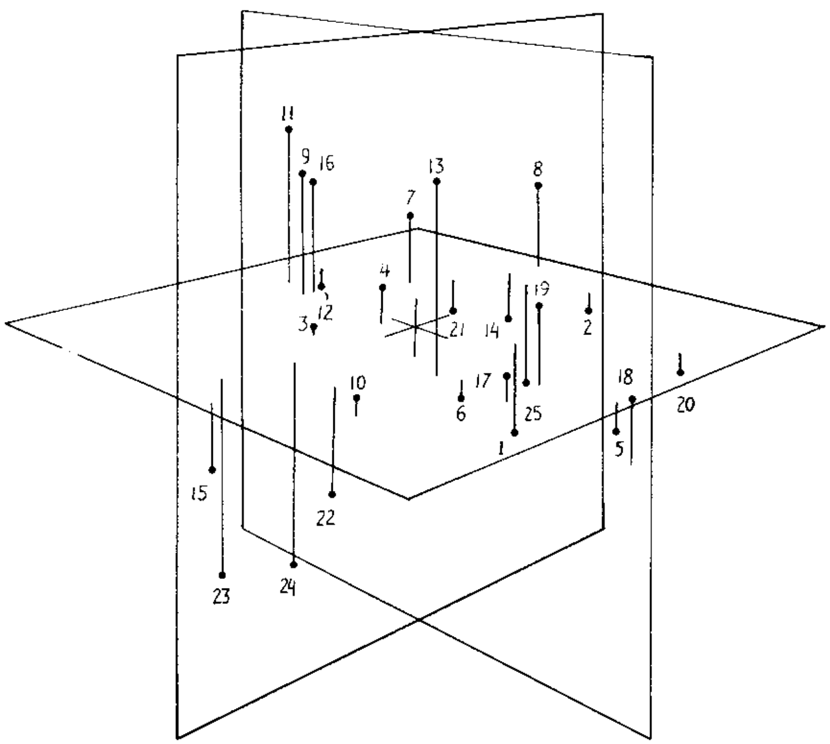

Figure 36.4. Plot of message concepts in three dimensions.

(6) *Multidimensional scaling.* We factored the matrix using the GALILEO metric multidimensional scaling procedure referenced earlier. Three dimensions accounted for meaningful variance. The coordinate projections of each of the twenty-five concepts are represented in Figure 36.4. It shows the x-y, x-z, and y-z planes and the twenty-five concepts positioned within these dimensions. The names of these concepts are listed in Table 36.1. More detailed MDS results are available from the author upon request.

(7) *Optimal message generation.* As discussed earlier, for some evaluation objectives it is useful to formatively evaluate communication management strategies to enable conference leaders to shape the course of subsequent conferencing, bringing it more closely in line with purposes and objectives. To illustrate application of the present method for this purpose, we used the Automatic Message Generator (AMG) function of the GALILEO multidimensional scaling program. AMG works according

to standard vector algebra procedures, determining the length and direction of the resultant vector for the concept to be moved as vectors of additional concepts are added. The program determines the efficiency of each possible combination so that users can select the particular combination with the highest likelihood of repositioning the concept within the larger concept space. The details as to how the algorithm operates are well documented in Woelfel and Fink (1980).

To apply AMG, the investigator first selects a focal concept to move and a target toward which to move it. An informal analogy as to how AMG works is the game of billiards. A player selects a ball to move toward a target, usually a pocket. Then he or she considers alternative angles and forces with which to strike the focal ball with other balls. The player then selects the most likely combination and executes.

Because the current research is illustrative, the choice of concepts to be moved and target concepts is arbitrary. We selected as the target the concept closest to the centroid of the space. This was *giving information.* The most central concept was selected as the target because it may often be the case that conference leaders wish to move a concept closer to the center of discussion. Nevertheless, movement of concepts away from targets is just as easily analyzed. We chose *user groups* as the concept to move closest to the center. Many CBBS conferences are operated by user groups, and there may be advantages to increasing the centrality of discussion about user group concepts.

We instructed AMG to select the best two-concept AMG message set. The two optimal message concepts the AMG routine selected were *offer information* and *source listing.* After entry of the optimal messages, including these two concepts with *user groups,* the actual effects could be measured by performing steps 1-6 with the next time segment. Furthermore, experimentation could then determine how many repetitions of the message are needed to achieve the desired movement of a concept toward or away from the target.

DISCUSSION

Paths to Refining the Methodology

The present case illustration demonstrates that the CMC evaluation methodology's core network analysis, content analysis, and multidimensional scaling procedures can be meaningfully applied to actual CMC. There are now several directions for refinement and elaboration. One is to extend the present analysis to the time-series case. A stream of conference messages can be segmented into time intervals based on when messages are entered. Then a series of content analyses and multidimensional scaling routines can be performed. The results can reveal change over time in the conceptual space of the conference.

These procedures enable the conduct of refined field experiments that can serve a range of basic questions, hypotheses, and practical applications. For example, does increasing message content diversity lead to a less densely connected communication network structure? Might the reverse be the case depending on the kind of content (Danowski, 1980)? A more practical example would be that if one were interested in the effects of leaders' use of optimal (AMG) messages or other messages in a conference, one could systematically enter these messages and observe their effects over time on the conference content space.

Consider another example that ties the current method to a survey method: If one were interested in experimenting with the effects of message content on the cognitive structures of users, one could apply our methodology to the conference messages, survey users to directly measure cognitive structures with self-report concept proximities data, and then analyze these using the same MDS program. In so doing, one could examine the effects of changing message content on users' direct conceptual structures over a time-series.

As we discussed earlier, our procedures alone, without self-report surveys, measure the *underlying* conceptual structure to the extent that psychological dynamics are translated into overt messages. This may be a major portion of cognitive structure. But there may be interesting aspects of it that do not get translated into overt messages to others — for example, perceived attributes of a computer or conferencing system as a whole. Users may have many attitudes and cognitions that they have no motive to express unless asked by someone, yet these psychological factors may indirectly affect other variables of interest.

A second, future extension of the method also concerns segmentation, but of a different sort. Distinct communication network groups can be identified according to structural criteria — for example, nodes who share a majority of links among themselves compared to the total set of nodes in the conference. If separate network groups exist, then the basic methodology presented here can be used within each segment. Such *infographic* segmentation can be useful for a variety of evaluation purposes. One example is a possible need to develop different optimal communication management strategies for the various network groups. Or at a more basic theory level, one might hypothesize that within network groups varying in structural features — for example, in internal density or connectivity or in the diversity of environmental linkage — different patterns of message content may be exchanged.

A third extension of the methodology, one we are currently exploring, is to use automated content analysis procedures. Once an appropriate program is selected and tested, the application of the present methodology could be extended to virtual real time application. To do this, all that would be required is for the user to read prior CMC messages into a file, select the segments for which separate analyses are desired, then call up the content analysis program and the multidimensional scaling program. At that point the statistical analysis would be complete. The user could observe the graphic and/or tabular results of the MDS on his or her terminal, perhaps select the AMG optimal message option and identify the optimal message, log back into the conference, and enter the message. After entering optimal messages for a time, the user could then repeat the analysis in the same way to see what effects the optimal messages have actually had on the overall concept space of the conference. A program called CATPAK(TM) (Woelfel & Holmes, 1982), developed after the current research was done, seems useful for such purposes.

Issues

The nature of the CMC evaluation methodology may raise some critical issues. One issue concerns the social control aspects of the optimal message applications. Some may view the techiques as too "Orwellian." Privacy per se, of course, is not technically an issue, provided that one applies the methodology as we have here to *public* conferences, not to private electronic mail. Still, some may feel that the analysis, selection, and entry of optimal messages is excessively manipulative. Nevertheless, a counter-argument can be raised that people naturally attempt to influence the course of their communication with others, regardless of whether it is face-to-face communication, telephone communication, computer communication, and so on. Attempting to influence the course of computer conferencing is qualitatively no different from influencing day-to-day communication as it has been occurring for millenia. Furthermore, people expect control to be exercised by their leaders, provided it is not excessive and in their best interests.

Nevertheless, such techniques could be made available to anyone who wished to use them in a conference. Yet some have informally suggested that this may result in "message wars" among communicators, each of whom is analyzing and entering optimal messages. While the images this suggests may be entertaining, message wars are unlikely to become day-to-day practice. One may expect that the degree of message-optimizing that might go on would be similar to what occurs for other kinds of communication within a particular social community. The basic personalities of people govern the overall contours of communication experiences.

All things considered, the method is merely a tool for measuring CMC, one tool in a large assortment. Its uses and implications depend fully on the research or evaluation stakeholders, their objectives, their applications in conjunction with other methods, and their results in achieving these objectives. In short, the method presented here is an enhanced set of procedures linking together several bodies of methods — network analysis, content analysis, and multidimensional scaling — in such a way that these are responsive and sensitive to CMC's special features.

REFERENCES

Christensen, W., & Suess, R. Hobbyist computerized bulletin board. *Byte,* November, 1978, pp. 150-157.

Danowski, J. A. An infographic model of media access for elderly advocate organizations. In P.A. Kerschner (Ed.), *Advocacy and age.* Los Angeles: University of Southern California Press, 1976.

Danowski, J. A. Group attitude uniformity and connectivity of organizational communication networks for production, innovation, and maintenance content. *Human Communication Research,* 1980, 299-308.

Hiltz, S. R., & Turoff, M. *Network nation: Human communication via computer.* Reading, MA: Addison-Wesley, 1978.

Kruscal, J. B., & Wish, M. *Multidimensional scaling.* Beverly Hills, CA: Sage, 1978.

Parnes, R. *The CONFER conferencing system.* Paper presented at the conference on Evaluating Computer Conferencing, Newark, New Jersey, 1980.

Rice, R. E., & Danowski, J. A. Issues in computer conferencing evaluation and research. In S. R. Hiltz and E. B. Kerr (Eds.), *Studies of computer mediated communication: A synthesis of findings.* Final Report in the National Science Foundation, NSF IST-801077, 1981.

Woelfel, J., & Fink, E. *The measurement of communication processes: Galileo theory and method.* New York: Academic Press, 1980.

Woelfel, J., & Holmes, R. *CATPAK® demonstration.* Paper presented at the meeting of the International Communication Association, Boston, 1982.