

Empirical Evidence for Information Overload In Mass Interaction

Quentin Jones, Gilad Ravid, Sheizaf Rafaeli

Graduate School of Business

University of Haifa,

Mount Carmel, Haifa 31905, Israel

+972 (8) 972-5221 qgjones@acm.org

ABSTRACT

'Virtual publics' are computer-mediated discourse spaces created by using various technologies including email, the USENET, web based bulletin boards, IRC, MUDS, etc. [3]. This paper outlines ongoing field research into the stress zones or boundaries to interactive virtual public discourse produced by information overload. It describes initial findings, based on an examination of 2.65 million USENET messages, which suggest that information overload impacts on discourse structure. Findings include a higher proportional user turnover, and shorter messages in relation to group size. The research in progress addresses issues associated with group level usability of communication technologies.

Keywords

Virtual Communities, Virtual Publics, Computer Mediated Communication, Information Overload.

INTRODUCTION

Virtual publics are symbolically delineated, computer-mediated spaces, whose existence is relatively transparent and open, which allow groups of individuals to attend and contribute to a similar set of computer-mediated interpersonal interactions. The value of any virtual public will relate to the size of its user population, and both the quality and quantity of user contributions. Unfortunately, the expansion of virtual public discourse is not simply a matter of increasing user population. This is because the relationship between user population and discourse contributions is influenced by a variety of factors including, critical mass, social loafing and information overload [3]. While these points can be deduced from existing empirical and theoretical works, little data has been presented that concretely describes how these phenomena are related to various technologies. This research addresses this issue.

MODELING THE CONNECTION BETWEEN COMMUNICATION TECHNOLOGY & DISCOURSE

The degree to which information technologies can effectively control, or aid, computer-mediated communication (CMC) is limited by the finite capacity of

human cognition. The inability of users to process effectively certain message patterns will result in limitations to the possible forms of sustainable group-CMC. Beyond a particular communication-processing load, the behavioral stress zones encountered will make particular forms of group communication unsustainable. Communication load is the processing effort required to deal with a set of communications.

In cases of "mass interaction" [6], such as Usenet discourse, in theory it should be possible to observe empirically the combined effects of the average maximum communication load individuals are prepared to process (AvMaxCL). Mass interaction provides a unique opportunity to explore the impact of communication load on group discourse.

Communication-processing load relates to a number of message-system characteristics. Users generally have to make more of an effort to reply coherently to a thread [4] than to a single message. Therefore, higher interactivity correlates with higher communication-processing load. Interactive communication refers here to the extent to which messages in a sequence relate to each other, and especially the extent to which later messages recount the relatedness of earlier messages [5]. Similarly, a high frequency of postings will require more processing by group members. Therefore, message frequency will also covary with communication-processing load. It is also likely that a decrease in 'interactional coherence', not compensated for by a useable persistent record, will increase communication-processing load [1]. For example, disrupted turn adjacency may require increased user effort to track sequential exchanges. Disrupted turn adjacency is caused by the fact that CMC-systems, such as email lists, transmit messages in the order they are received. Thus in group-CMC a message may be separated from the previous message it is responding to by another message, or lags in message transmission may even result in reversed sequencing.

Different CMC-technologies have different typical message system characteristics. Therefore, the point at which a user population's interactions will typically result in information overload will relate to the CMC-tool used.

It is suggested that the potential number of users involved in 'a sustained pattern of interactive group discourse' via a

single virtual public will be constrained by cognitive processing limits. Further, that such limits will produce communication stress zones, or boundaries, related to: message posting frequency; average depth of threads; the quantity of discourse material; disrupted turn adjacency; and how the technology in question structures communication. These communication boundaries are not rigid, deterministic, instantaneous halt lines. Rather, they are zones, with an uncertain range of likelihood, within which the behavioral limitations become severe. The research described here, aims to identify such zones for Usenet discourse.

Cognitive Processing Limits Model and Usenet Discourse

The study aims to identify the stress zones produced by cognitive processing limits for USENET newsgroups by the analysis of data collected from large-scale field research.

From the introduction above, it can be concluded that if the structure of a newsgroup's discourse is close to AvMaxCL, then users will try to offset an increase in any one part of the cognitive load function (*CLf*) by decreases in other parts. Therefore, the *CLf* predicts that at AvMaxCL an increase in the number of interactive posters will typically be associated with one or more compensatory strategies such as: 1) user disengagement; 2) the length of messages (number of words or lines); 3) number or depth of discussion threads. In other words, there will be a richness-versus-reach tradeoff in the of high volume discourse.

RESEARCH METHODOLOGY

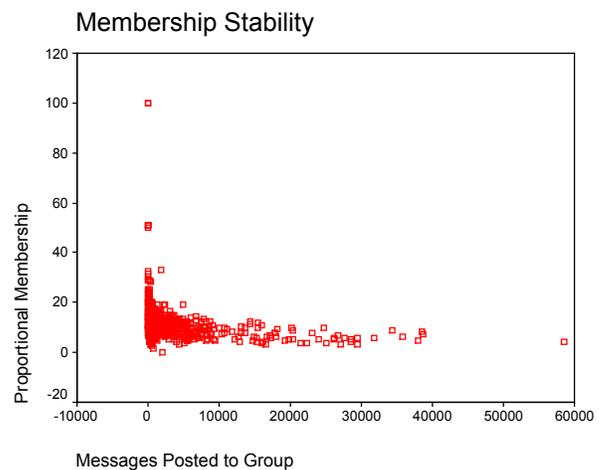
Data Collection and Sampling

For this project, data was collected from the 600 Usenet newsgroups based on the sampling technique of Whittaker et al's [13]. Over a 6-month period, 2.65 million USENET messages were collected.

RESULTS

Diagram one displays the number of messages posted to the 578 newsgroups that were active during the first 5 months of the study. Proportional membership is the percent of posters per-newsgroup per-month, who also posted to the next study month. On average only 11.5% of posters sent messages 2 months in a row. The diagram suggests that user dropout as a response to information over load can clearly be observed in Usenet interaction. The drop in the proportion of individuals involved in sustained discourse is quite strong with a Spearman's rank correlation coefficient of $-.42$ ($p < .000$).

Similar evidence and plots were found for all the other compensatory strategies in relationship to various measures of group size examined. For example, the larger the group, the smaller the number of words posted in interactive messages. Likewise, the relationship between number of threads and their depth supports the notion of a richness-versus-reach tradeoff strategy. Shorter messages were also more likely to generate discussion threads.



SIGNIFICANCE

This work in progress is of significance for two reasons. First, it should pave the way for comparative studies of the boundaries to sustainable interactive group discourse using a variety of CMC-tools. The authors have collected email list data for this purpose. Such comparative studies should help us start to come to terms with the link between types of CMC-tools, discourse, and discourse structure. This in turn will provide us with useful usability information at the group level. Second, the study offers hope that in the future, with sufficient modeling of the data obtained from mass interaction, we will be able to provide a formula that roughly estimates communication load in a variety of situations.

REFERENCES

1. Herring, S. C. Interactional coherence in CMC. In *Proceedings of the 32nd Hawaii International Conference on System Sciences*, (Hawaii 1999), IEEE Press.
2. Jones, Q., and Rafaeli, S. What do virtual 'Tells' tell? Placing cybersociety research into a hierarchy of social explanation. In *Proceedings of the 33rd HICSS*, (Hawaii 2000), IEEE Press.
3. Jones Q., & Rafaeli, S., "Time to Split, Virtually: 'Discourse Architecture' and 'Community Building' as means to Creating Vibrant Virtual Publics". *Electronic Markets, The International Journal of Electronic Commerce and Business Media*. Vol. 10, No. 4, 2000, Routledge, London.
4. Lewis, D., and Knowles, K. Threading electronic mail: A preliminary study. *Information Processing and Management* 33,2, 1997, 209-217.
5. Rafaeli, S., and Sudweeks, F. Networked interactivity. *J.C.M.C.* 2, 4, 1997, <http://www.ascusc.org/jcmc/vol2/issue4/rafaeli.sudweeks.html>
6. Whittaker, S., Terveen, L. et al. The dynamics of mass interaction. In *Proceedings of CSCW 98*, (Seattle 98), ACM Press.