# Analysis on Information Overload in Mass Interaction

**Kalpik Patel[1] Prof. Dr. Dhaval Kathiriya[2]**

[1]Research Scholar, Mewar University, Chittorgarh, Rajasthan, India

[2](Director) - Information Technology (Professor) - Computer Science, Anand Agricultural University

***Abstract – Usenet is a popular distributed messaging and files sharing service: servers in Usenet flood articles over an overlay network to fully replicate articles across all servers. This paper presents the design and implementation of Usenet MHT with removal technique. From 1979 and beyond, it has seen a near exponential expansion in the mass of data transported, which has been a strain on bandwidth and storage. There has been a spacious range of academic research with focus on the WWW, but Usenet has been neglected. Instead, Usenet's evolution has been dominated by practical solutions.***

***Keyword: NNTP, UUCP, Usenet***

----------♦----------

## 1. INTRODUCTION

News is a distributed platform for group communication — mainly between humans —based on a network of servers all around the world. "Usenet" is an abbreviation for "Unix User Network", but is also known under other names, specifically "Netnews", simply "News" [Hardy, 1993] or "Usenet News". News is a slightly misleading name for what Usenet is meant for: asynchronous communication between people, as opposed to news items distributed by mass media. 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 (Hardy, 1993). 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. The exponential growth in recent years of telecommunication technologies has resulted in a new era of interpersonal communication (Saxena, Saxena, 2011).

Computer mediated communication (CMC) tools have altered both one-to-one and one-to-many communication. The growth of discourse systems where the audience is a significant source of media content as well as its primary receiver has resulted in what has been described as "mass interaction" (Rheingold, 1993), shared discourse between hundreds, thousands or more individuals. The term virtual community is commonly associated with such large-scale discourse. However, no single dominant definition of the term exists (Jones, Rafaeli, 2000b). In fact, a number of authors dispute the existence of virtual communities (Whittaker, Terveen, 1998). Despite the significance of the phenomena commonly labeled *virtual community*, this term is problematic.

It consists of a set of "newsgroups" with names that are classified hierarchically by subject. "Articles" or "messages" are "posted" to these newsgroups by people on computers with the suitable software these articles are then broadcast to other interrelated computer systems via a wide variety of networks.

In simpler way we say this electronic bulletin system. The paper can be propose the idea to better management of that articles using minimum storage as possible as and also removal technique for the old articles.

## 2. REVIEW OF LITERATURES:

The Usenet is a system of electronic bulletin boards, referred to as newsgroups. It is not a computer network, but rather a network of bilateral agreements among system administrators to cooperate on bulletin board management (Sproull, Faraj, 1997). Representative sampling of Usenet discourse is difficult; Whittaker et al's (Lewis, Knowles, 1997). solution was to produce a randomly stratified sample, of English text based Usenet newsgroups. They extracted 500 newsgroups from a subset of then active, widely distributed newsgroups, which contained predominately English language text, based conversational messages. For this project, data was collected from the 500 newsgroups studied by Whittaker et al enabling detailed historical comparisons. An additional 100 newsgroups were selected using Whittaker et al's approach with only minor modifications. This allowed for 100 moderated-groups to be selected. The full content of 3,293,995 postings were collected over eight months and stored in an Oracle database. The 2,652,552 messages collected over the 6months from 1st August 1999 to 29th February 2000, were used to conduct this study. Newsgroup servers are hosted by different organizations and institutions. Most Internet service providers host their own news servers, or rent access to one, for their subscribers. There are also a number of companies who sell access to premium news servers.

Each host of a news server maintains agreements with other news servers to frequently synchronize. In this way news servers form a network. When a user posts to one news server, the message is stored locally. That server then shares the message with the servers that are associated to it if both carry the newsgroup, and from those servers to servers that they are connected to, and so on. For newsgroups that are not commonly carried, sometimes a carrier group is used for cross redistribution to aid distribution. This is typically only useful for groups that have been removed or newer *alt.** groups. The main flow of Usenet is commonly through the Internet, using the Network News Transfer Protocol (NNTP) , a TCP based protocol for transmission. Most Internet standards are described in RFC, and the IETF is working on several new standards. Usenet's standards are described in RFCs. but there are de facto Usenet standards not included in the RFCs, although the IETF is working on standardizing these enhancements. Message format are logically divided into two separate parts, *head (also called headers) and body. The headers* contain meta-information about the article , such as who allegedly posted the article, from where, at what time, to which newsgroups, with what subject of discussion, a unique message ID, and the path through which servers the article has been passed to avoid re-relaying to those servers. Other headers may be used, but these are not relevant here. Body part can be used to store a data or complete article . once a article is send no one can change the article other can be read the article send to other to tag to other but can't change the article.

## 3. 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. 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 (Rafaeli, Sudweeks, 1997) 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 (Herring, 1999). 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 (Rigaut, *et. al.,* 1997). 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. The study aims to identify the stress zones produced by cognitive processing limits for USENET newsgroups by than lysis of data collected from large-scale field research.

From the introduction above, it can be concluded that if structure of a newsgroup's discourse is close to AvMaxCLthen users will try to offset an increase in any one part other 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 message(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.

## 4. ANALYSIS OF THREAD TRAPPING

The TAP method involves the fusion of the TAP tag (see below) to the target protein and the introduction of the construct into the host cell or organism. For optimal results, it is preferable to maintain expression of the fusion protein at, or close to, its natural level. Indeed, overexpression of the protein often induces its association with non-natural partners (heat shock proteins, pro-tea some; Ref. (Rafaeli, Sudweeks, 1997). Cell extracts are prepared and the fusion protein as well as associated partners is recovered by two specific affinity purification/elution steps.

The material recovered can be analyzed in several ways. For protein complex characterizations, proteins are con-cent rated, and eventually fractionated on a denaturing gel, before identification by mass spectrometry. (Alter-natively, Edman degradation or Western blot may be used.) Because the various TAP purification steps are performed in a gentle native manner, purified complexes may also be tested for their activities or used in structural analysis.

As noted above, trapping interactivity is essential. To achieve this end discussion threads need to be reconstructed reasonably accurately. Therefore, it is important to identify if a message is truly a “reply”, and if it is, to correctly identify its “parent” message.

## 5. ORGANISATION, DISTRIBUTION AND MASS INTERACTION

There is several level of structure in Usenet. Coactions of messages are clustered into newsgroups, and newsgroups themselves are organized into hierarchies. Each hierarchy is intended to address different conversational topics. There are over a hundred different hierarchies but the majority of newsgroups belong to one of eight main hierarchies ~’the big eight”). These eight are: *alt* for alternative topics; *comp* for computer issues; *humanitia; mkc* for misce~aneous discussions; *nws* for discussions about UseNet~ *rec* for recreational topics; soc for social issues; and *talk* for general conversations. Each hierarchy is then subdivided into more specific sub hierarchies. The name of each newsgroup begins with the relevant main hierarchy, and tem of increasing specificity are added to this. Typical newsgroup names are rec. music. dylan and rec. music. beatles, where both newsgroups are part of the music subhierarchy within rec. Twelve percent of newsgroups are moderated, and moderation is reflected in the newsgroup name: soc. russian. Culture. Moderated. Newsgroup distribution also varies from site to site. Newsgroups are distributed through an informal network of servers.

**Generating Simpler Responses in Situations of Overloaded Mass Interaction**

Two sub-hypotheses are examined with respect to the generation of simpler responses in situations of overloaded mass interaction. These are: 1) There will be a decrease in surrogate measures of complexity of interactive message communication, such as word count, as the size of the interactive group increases although this will approach asymptote; and 2) There will be a decrease in surrogate measures of message complexity (e.g. the number of new words per interactive message) as the number of discussion threads in the newsgroup increases, although this will also approach asymptote.

The reason for this hypothesized reduction in message complexity is due to the increased effort required by authors to create such messages. Clearly, there is no absolute measure of message simplicity/complexity although there should be a rough correlation between various message characteristics and the effort required to create and read them. There are a number of ways the notion of “message complexity” and “interactive group” can be operationalized. The first step therefore, is to examine some of these measures and assess their appropriateness. Common sense informs us that on average the effort required to create a Usenet message will correlate with a number of message characteristics, the most obvious being message length. For each Usenet message the following variables related to length were calculated:

- The number of words in the body of the message (words).
- The number of words on non-indented lines in the body of the message (new words).
- The number of lines according to the header field, this typically includes lines of attachments.
- The number of lines excluding those of attachments (lines)
- The number of non-indented lines excluding those of attachments (new lines).

## CONCLUSION

The emergence of mass interaction has presented new opportunities to learn about and understand human communication. The availability and persistence of such communications, and the scale at which it operates allows

us to explore various system effects on group discourse. At this point in time empirical research into the *systemic* nature of the patterning of social relationships in cyberspace has, despite its importance, been relatively rare. Overall, the results strongly support the assertion that individual 'information overload' coping strategies have an observable impact on mass-interaction discourse dynamics. Clear evidence was found for the hypotheses: that users are more likely to respond to simpler messages in overloaded mass interaction; and that users are more likely to end active participation as the overloading of mass-interaction increases. Evidence was also found for the hypothesis that users are more likely to generate simpler responses as the overloading of mass-interaction increases. Work which use a systems approach to examine internet based group communication, such as: the modeling of free riding using the Napster like Gnutella network modeling the inter-relationship between homepages exploring the self-organizing nature of email lists and showing the World Wide Web to be structured like a small world network have been undertaken in the last five years.

## REFERENCES:

[Hardy, 1993] Hardy, H. E. (1993). The Usenet System. *ITCA Yearbook*.

Deepika Saxena, Monika Saxena (January 2011). "Methods Used to Handle Overloading of Information in Usenet" [International Journal of Scientific & Engineering Research, Volume 2, Issue 1. 1 ISSN pp. 2229-5518] http://en.wikipedia.org/wiki/Network_News_Transfer_Protoco

Rheingold, H. (1993). The virtual community: Homesteading on the electronic frontier. Reading, MA, Addison-Wesley.

Jones Q. and Rafaeli S. (2000b). What do virtual 'Tells' tell? Placing cybersociety research into a hierarchy of social explanation. 33rd Hawaii International Conference on System Sciences, (Hawaii 2000), Hawaii, IEEE Press.

Whittaker, S., L. Terveen, et al. (1998). The dynamics of mass interaction. CSCW 98, Seattle, ACM Press.

Sproull, L. and S. Faraj (1997). Atheism, sex and databases: The Net as a social technology. Culture of the Internet. S. Kiesler. Mahwah, NJ, Lawrence Erlbaum Assoc, Inc.

Lewis, D., and Knowles, K. (1997). Threading electronic mail: A preliminary study. *Information Processing and Management* 33,2, pp. 209-217.

Rafaeli, S., and Sudweeks, F. (1997). Networked interactivity. *J.C.M.C.* 2, 4, http:// www.ascusc.org/jcmc/vol2/issue4/rafaeli.sudweeks . Html

Herring, S. C. (1999). Interactional coherence in CMC. In *Proceedings of the 32nd Hawaii International Conference on System Sciences.* IEEE Press.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Se ´ raphin, B. (1999). *Nat. Biotechnol.* 17, pp. 1030.