MARKETING MANAGEMENT

"COMPUTER MEDIATED COMMUNICATION IN ENTERPRISES, NEED FOR THEIR CLASSIFICATION AND ALGORITHMIC COMPLEXITIES"

PROF. PRITI KULKARNI
ASST. PROFESSOR
SYMBIOSIS INST OF COMPUTER STUDIES AND RESEARCH,
SYMBIOSIS INTERNATIONAL UNIVERSITY, PUNE

DR . HARIDASA S. ACHARYA
PROFESSOR
AIMS, PUNE

I. INTRODUCTION

Computer-mediated communication (CMC) has become very common in work life and had replaced need for face to face (F2F) communication to a large extent. Email communication is one dominant component of CMC. Chat and voice and video mails are also important. In the business domain e-mail is most frequently used in an information-gathering and communication function (Kettinger & Grover 1997). Email has also led to the emergence of on-line communities by supporting asynchronous communication (Sproul and Kiesler 1991). Kidd (1994) , Landsdale (1988) and Malone (1983) have portrayed email as a **virtual file cabinet**. It is a means of storage to use, and organise messages. Classifying the messages in terms of their contents, for effective retrieval for proper business use of the messages is the next step, hence there is a need for proper structured filing. Malone, Grant and Turbak (1986) developed a computer based system which helps employees to filter, sort and prioritize messages at individual levels.

Mackay et al., (1989) revealed that newer computer users can effectively sort and prioritize emails to some extent. Landsdale (1988) emphasized that two problems faced by uses persisted :

a) How to define categories to use.

b) How to remember these categories later, at the time of retrieval.

He concluded that information retrieval based on recall-directed search is always followed by reorganisation based scanning. Another dilemma for classification is about sorting and categorizing email is that if user spends

time to categorize an item less likely categorization will be done. If this process is automated the fewer users will able to recall it. This suggest that automatic filing and message folders two ubiquitous features in contemporary email software and intended to help email overload problem.

In this paper, evidences from few prominent international surveys are given to indicate the extent to which enterprises rely on emails, next different purposes and approaches to categorization are discussed. There are a few network infrastructure related intricacies which cannot be ignored when Enterprises go for dedicated applications to be deployed for the purpose of monitoring and categorization of emails. Most of the researchers seem to have ignored this aspect while discussing the categorization problems. We focus on this aspect of email categorization in the later part of this article.

2. INTERNATIONAL SURVEYS

Radical group (Anonymous, 2012, 2013) has conducted survey in mid of 2011 to study corporate and business user preferences, deployments, attitudes and behaviour with regards to email with 100 businesses and a total of 228,068 email users. The survey indicated that

a. 44.8 billion Emails are delivered each day.

b. 112 daily emails are received and sent on average by business professional causes an overflowing inbox.

c. ***42% of corporate email is ignored***. (Due to inbox is overflow).

d. ***50% of emails are misunderstood***. (Without these visual/auditory cues)

e. People switch between applications 30 times an hour. (Leeds to loss of focus and working time).

3. SURVEY BY INFOCOMM

According to a survey conducted by Infocomm In 2012, the usage of computers, Internet, broadband has increased. The proportion of enterprises that used computers and the Internet increased from 79% in 2010 to 84% in 2012 and from 77% in 2010 to 82% in 2012 respectively. The most common Internet activities were sending and receiving emails (95%) and information search (92%).

Fig 1 : Internet activities on computer (Annual Survey on Usage by Enterprises 2012)

The results show that usage of internet for communication dominates practically every other usage.

4. TRENDS AND PREDICTIONS

According to statistics published by Radicati Group, estimates for future indicate(Table1) :

An average annual growth rate of 6% over the next four years is indicated.

a. ESTIMATES OF SPAM EMAILS

Spam mails are a real problem. Statistics as on Week ending November 18, 2012 by **Trust wave** indicated 67.8% Spam , and **Spam cop** reports 2679044 spam messages per week. Hence the top level categorization is always spam vs no spam.

b. EMAILS ON A BUSINESS STRATEGIC NETWORK

A study by Markus (1994) based on 549 emails sent from the hub of the strategic network during six months in 2002 and 2003 , are classified in Table 2. The respondents have reported that they have not deleted any job-related emails, the content of their out boxes can be expected to be a good representation of typical communication behaviour in the strategic network. As most emails have been classified as having more than one communication reason, the total number of registered reasons exceeds the number of emails.

	Daily Traffic emails per day			Volume (Number of Email Accounts)		
Year	2013	2014	2017	2013	2014	2017
Worldwide	182.9	191.4	206.6	3.9billion	4.1billion	4.9 billion
Business	100.5	116.2	132.1	24%(929)	24%	23%
Consumer	82.5	80.2	74.5	76%	76%	77%

Table: 1 Email traffic and Accounts 2013-2017 (Extracted from: Radicati Group survey -2013-17)

The use of email in the hub of the strategic network

	Project leader	Admini-strator	Marketer	Network members	TOTAL
Number of emails	413	115	6	15	549
Content Reasons					
• To convey confidential, private or delicate information	2	-	-	-	2
• To describe a complicated situation or proposal	6	1	-	-	7
• To influence, persuade, delegate or sell an idea	13	4	2	1	20
• To express feelings or emotions	78	18	-	2	98
• To keep someone informed	323	84	6	7	420
• To follow-up earlier communication	54	23	1	5	83
• To question	53	27	1	3	84
Situational Reasons					
• To respond to a straightforward telephone message	9	-	-	-	9
• To respond to a complicated email message	18	16	-	8	42
• To communicate the same thing to many people	8	13	2	5	28
C. Symbolic Reasons					
• To be casual, informal	37	5	-	-	42
• To convey urgency	5	10	-	3	18
• To convey personal concern or interest	48	4	-	-	52
• To obtain and immediate response, action	48	26	1	5	80
• To show authority, status, position	-	1	-	-	1
• To show that the communication is official	-	-	-	-	-
Average number of registered reasons per email	1.7	2.0	2.2	2.6	1.8

Table 2: Category wise analysis of emails from a Business Strategic Hub

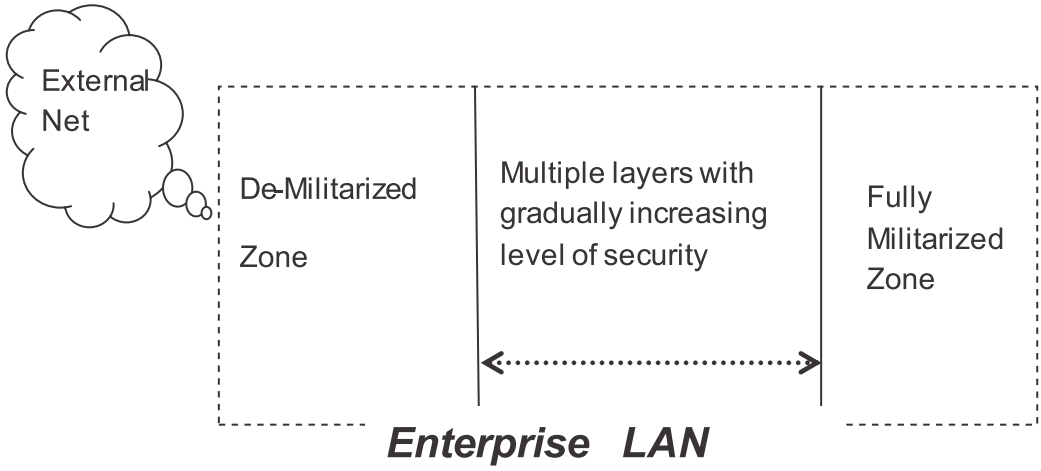

Fig 2 The Layered Concept of Security in an Enterprise Network

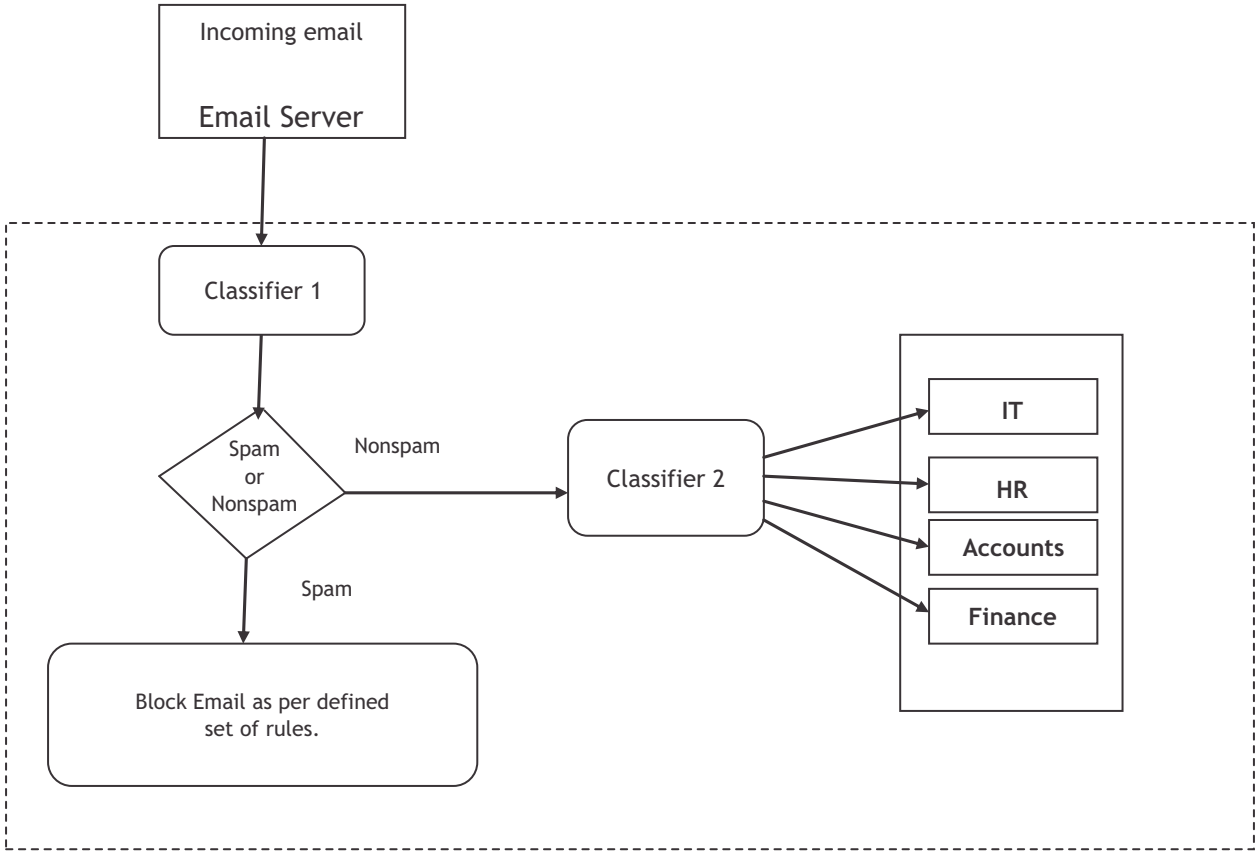

Fig.3 Typical (simplified) Email classification framework at enterprise level

5. EMAIL CLASSIFICATION AT ENTERPRISE LEVEL

According to AMA 43% of companies monitor email, 96% track incoming and outgoing messages. Only 58% monitor internal traffic sent from internal staff. Employees assigned to read and review employee e-mail are employed by the following departments: IT (73%), HR (34%), legal (18%), compliance (17%), outside third-party (4%), other (17%).

In order to classify the emails, first decide on useful categories that represented the whole of emails. When email hit email server first it is checked for spam or non spam , and is discarded. If email is non-spam it is classified on basis of different users.

6. NEED FOR MULTI LEVEL CLASSIFICATION

As we mentioned earlier need for classification though is well recognized most of the surveys have focussed on identification of business reasons for categorization. However it is necessary to recognise that dedicated classifications and monitoring require huge cost to be incurred. Most of the time Infrastructure designers do not understand the needs and hence do not create appropriate provisions for multi-level . Let us have a look at the security setup in a typical Enterprise level network. For simplicity we have shown a LAN, but the principles remain the same when we are looking at a global network.

An Enterprise network will have varied security setup, possibly separated by multiple firewalls. The information need will be different in different layers.

Email monitoring and categorization needs at each layer certainly would be different. The stringency of censorship, importance of decision making needs, risk associated with actions associated with each layer will dictate the type of email classification applications to be used in each layer. Most of the surveys have focussed either on total volume and traffic of email communications. However a strong coupling of the network design and dedicated categorization applications customizable to varies levels of accuracy of classifications would be needed in any enterprise.

The quick access to records and other documents helps for sound decision-making. According to the Australian Standard for Records Management, AS ISO 15489, best practice in records management classification is on a rigorous analysis of business functions and activities. Similar standards would exists in most countries as this is essentially a business requirement.

A natural corollary of these standards specification is that such policies should have natural extention to Email categorization and monitoring as well.

The process of information classification describes, organize and control information. It creates order in understanding what an organization does and how it does it. Government and other organization have invested heavily to protect their system from virus, worms and other threads. Users often have cluttered in boxes containing hundreds of messages, including spam, Unsolicited marketing messages, outstanding tasks and conversational threads. Also the exponential increase in the volume of e-mails can make the processing of e-mails tedious and time consuming. The e-mail user spends most of the time on organizing these e-mails to reduce their size of inbox. Sometimes important messages get overlooked, or "lost" in archives. Also another risk includes loss of proprietary information violation of record retention and privacy laws. Email is considered as a weak link in organizational security. This rapid growing email requires proper classification with less user interference.

An E-mail filter is personalized and the knowledge used by each personal filter is subjective. Therefore, classifying personal E-mail messages is more challenging. A rule-based system provides a way to semi-automate e-mail classification. Such a system requires users to define a set of instructions for the e-mail application to sort incoming messages into existing folder. (Bekkerman et al., 2004) showed applications in spam filtering of automatic e-mail classification extraction of e-mail threads and automatic e-mail foldering, as per user-defined folders. ***Ishmail*** automatically sorts messages into folders and orders them by importance, based on user-defined classification filters. ***Procmail***, ***Elm Filter*** works on same concept. Most popular commercial e-mail clients such as ***Eudora***, ***Mozilla Thunderbird***, ***and Microsoft's Outlook and Outlook Express*** also support message filing according to user-defined rule sets.

These types of systems are challenging for nontechnical users because writing the rules requires some level of programming experience, folders defined by individual users can be a tiresome and time consuming task if it is not

done regularly. Email data normally does not follow any fixed structure. The use of machine learning techniques would certainly be required. Examples of such systems are ***PEA-a Personal mail Assistant with Evolutionary Adaptation*** (Werner Winiwarter,1999), MailCat(Richard et al,1999), ***Re:Agent*** (Gary Boone,1998).

7. NEED FOR DEDICATED EMAIL CLASSIFIERS

Personal classification , organization and sorting can never be eliminated at the employee level. But an enterprise should go for better dedicated and standardised classifiers to eliminate business risks. The situation is almost similar to what we observe in the form of Spread Sheet Usage Risk , which is common in case of financial data processing .

8. CONCLUSIONS

The evidences about and the discussions on the network designs indicate that

a. Emails are heavily used by business as a Computer Mediated mode of communication.

b. Problems of lost mails, ignored mails and misunderstood mails are a reality.

c. Proper dedicated classification softwares will certainly reduce dependency on employees using personal classification strategies and hence reduce these problems.

d. Generally a multi layered security setup would be available in any enterprise network. Email categorization policies should also be parallel .

9. REFERENCES

1. Anonymous (2012)ANNUAL SURVEY ON INFOCOMM USAGE BY ENTERPRISES FOR 2012.
2. Anonymous (2012),The radicati group, inc. A technology market research firm, “Survey : Corporate Email, 2011-2012” http://www.radicati.com
3. Anonymous (2013) The radicati group, inc. A technology market research firm, “Survey: Corporate Email, 2013-2017” http://www.radicati.com
4. Anonymous(2007), The 2007 Electronic Monitoring & Surveillance Survey from American Management Association (AMA) and The ePolicy Institute.
5. Bekkerman, R., McCallum, A., and Huang, G.:(2004) Automatic Categorization of Email into Folders: Benchmark Experiments on Enron and SRI Corpora. Center for Intelligent Information Retrieval, Technical Report IR-418, 2004
6. Derks, D., & Bakker, A. (2010). The Impact of E-mail Communication on Organizational Life. Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 4(1), article 1.
7. Edith Andresen, Anette Bergeman, and Lars Hallen (2003), ELECTRONIC COMMUNICATION IN STRATEGIC NETWORKS , 19th Annual IMP Conference 4th-6th September 2003
8. Kidd, A. (1994). The marks are on the knowledge worker. Proceedings of the CHI 94 Conference on Human Factors in Computing Systems. New York: ACM.
9. Landsdale, M. (1988). The psychology of personal information management. Applied Ergonomics, 19, 55-66
10. Malone, T. W. (1983). How do people organize their desks? Implications for the design of office information systems. ACM Transactions on Office Information Systems, 1,99-112
11. Mackay, W. E., Malone, T. W., Crowston, K., Rao, R., Rosenblitt, D., & Card, S. K.(1989). How do experienced information lens users use rules? Proceedings of the CHI 89 Conference on Human Factors in Computing Systems. New York: ACM.
12. N. Ducheneaut, L. Watts, In search of coherence: a review of e-mail research, Human-Computer Interaction 20 (2004) 11-48.
13. Nicolas D.,and Palo A., “In Search of Coherence : A Review of E-Mail Research,” HUMAN-COMPUTER INTERACTION, 2005, Volume 20, pp. 11-48
14. Richard B. Segal and Jeffrey O. Kephart, ”MailCat : An Intelligent Assistant for Organizing E-Mail.”, In Proc. of the Sixteenth National Conference on Artificial Intelligence, pages 925-926, Orlando, Florida,USA, 1999
15. Rennie, Jason (2000) “ifile: An Application of Machine Learning to Mail Filtering.” Proceedings of the KDD-2000 Workshop on Text Mining, Boston, USA, August 20, 2000.
16. Radicati Group,(2013) The Ineffectiveness of Email Communication, Radicati Group, Message Mind Survey, Journal of Personality and Social Psychology, Communications of the ACM, Marketing Sherpa https://vingapp.com/ving-improved-communication-technology/5-shocking-facts-email-communication/#more-6717
17. Sproull, L., and Kiesler, S. (1991) ***Connections***. MIT Press, Cambridge,Mass., 1991
18. Walther, J. B., & Tidwell, L. C.(1995). Nonverbal cues in computer-mediated communication. Journal of Organizational Computing, 5, 355-378
19. W. Kettinger and V. Grover, 1997. "The Use of Computer-Mediated Communication in an Inter organizational Context," Decision Sciences, volume 28, number 3, pp. 513-555.