

# 國立中山大學八十八學年度博士班招生考試試題

科目：論文評述（資訊管理學系博士班）－管理組題目

1. 請先閱讀所附論文, “The relation between user satisfaction, usage of information systems and performance,”再詳細回答下列問題:  
有關調查<sup>(Survey)</sup>研究方面之論文寫作可能需考量那些重點項目? 請依你所列項目逐一評述所附論文之優/缺點, 並摘述該論文之總體表現。(50分)
  
2. 請先閱讀所附文章  
Using Intranets: Preliminary Results from a Socio-technical Field Study  
後, 再回答下列二個問題:
  - 2.1 請依您認為應有之考量, 評述本文之研究方法。(25%)
  - 2.2 假設本文的研究結果達到預期成果, 請說明一個您認為可能的後續研究之議題、重要性及研究方法。(25%)

(管理體系一篇)

## The relation between user satisfaction, usage of information systems and performance

---

### Abstract

This study investigates the validity of two commonly used measures for the success of information systems (is): usage and user satisfaction (us). A questionnaire survey among Dutch managers was used to assess the mutual relation between both measures and performance. The results indicate that us is significantly related to performance ( $r=0.42$ ). The relation between usage and performance is not significant. A partial correlation after correction for us is not significant either. This study provides empirical evidence for the popular assumption that us is the most appropriate measure for is success available.  
© 1998 Elsevier Science B.V. All rights reserved

*Keywords:* Information systems success; User satisfaction; Usage; Validation of measurement instruments; Survey

---

### 1. Introduction

The explanation of information systems (is) success has been called one of the main goals of is research [5]. Unfortunately, however, one of the prerequisites for measurement of is success has been the subject of much controversy [8, 15, 23]. The measurement of is success has been high on the research agenda for well over 15 years [20]. During these years papers have changed from theoretical discussions and rough measurement of is success (see e.g. [32] for a survey and discussion of measurement issues) to

empirical validation of measurement instruments [2, 11, 25, 26]. Within this period a tendency towards the application of more advanced psychometric methods can be seen. Reliability analysis using Cronbach's  $\alpha$  and exploratory factor analysis have gradually been replaced by confirmatory factor models [6, 10, 12].

In parallel with this development, a shift in the success measures can be observed. Laboratory research typically applied measures in which the contribution of is to performance is determined. Similarly, it has been proposed to compute the contribution of is to organizational performance in real world studies [1, 9]; indeed, attempts have been made to apply such measures in empirical research. Gallagher tried to determine the value of is in monetary terms, but his results were disappointing. Two dissertations at Ohio State University tried to assess the influence of is implementation on financial performance. However,

"[t]he influence of non-controllable variables prevented their reaching a conclusion" [16]. Apparently it is difficult to assess the contribution of IS to performance in a real world situation: a large portion of the costs and benefits will be qualitative or intangible [4, 24], the assessment of the value of unstructured or ad hoc decision making may be nearly impossible and organizations typically will not record these costs and benefits [24, 25].

Partially as a consequence of the difficulty in direct measurement of the contribution of IS to organizational performance, two alternative success measures gained acceptance: usage and user satisfaction ("the extent to which users believe the information system available to them meets their information requirements" [25]). Both are supposed to be proxies for the contribution of IS to organizational performance. The validity of research findings in which those measures are used to operationalize IS success ultimately depends on these measures' validity. Consequently, the development of theoretical and empirical foundations for their application deserves a high place on the research agenda.

## 2. Usage and us as success measures

A rationale for the application of usage as an IS success measure is the idea that it does not contribute to performance if it is not used (and will contribute to performance when it is). An alternative rationale states that users are able to assess the value of the IS and will use it if they conclude that the benefits (rewards) will outweigh the costs (efforts) [27, 31]. On similar grounds Ein-Dor and Segev assume that usage is highly correlated with other criteria (like profitability, application to major problems of the organization, quality of decisions, or performance and user satisfaction) as "a manager will use a system intensively only if it meets at least some of [these] criteria" [13]. Both rationales assume that more usage is better, which is not necessarily the case. Furthermore, application of usage as a success measure may suffer from the fact that a system will be used if managers feel that it facilitates their own goals. Thus, both perfect knowledge and goal-congruence between manager and organization are assumed. On another level, it is unclear what exactly is the amount usage of an IS. Also,

subjective measurement of usage may be influenced by social desirability and usage measurement may suffer from time-dependent noise. Finally, the application of usage as a success measure may lack sensitivity. Usage measurement will only identify the very unsuccessful systems [3] and whether managers will use an IS mainly will depend on negative aspects of the system. Provided that the benefits of using the IS outweigh the costs, it will be used.

The measurement of user satisfaction (us) will treat the very unsuccessful systems as non-existent, but is easier to differentiate between IS that are used. us measurement assumes that managers know their own information needs and this introduces the necessity of goal-congruence between the manager and the organization. Furthermore, it is assumed that improved performance will automatically follow if the system meets management information needs. This does not imply that satisfaction causes performance: performance and us are both caused by the extent to which information requirements are met. A possible shortcoming was noted by Melone [28], who doubts that users necessarily hold attitudes about their IS and, if they do, whether they are only formed when questions about us have to be answered; this would negatively affect the reliability of responses. She goes one step further and claims that attitudes that are not articulated will not influence perception, judgment, and behavior. This, however, presumes that us causes performance, which is not required: us is a reflection of the extent to which the information needs of the manager have been met and the assumption made in treating us as a success measure is that performance of managers will improve if their information requirements are met.

us also shares some shortcomings with usage: it may suffer from time-dependent noise [30] and may be influenced by social desirability. Furthermore, the problem of a valid us measure is apparent, but recently considerable progress has been made in developing an instrument and validating it by the application of more advanced psychometric methods.

## 3. Empirical evidence of the validity of us

Notwithstanding the apparent shortcomings of us as a success measure, the research community seems to

use it as the best proxy. us is increasingly employed in practice [7] and is the most commonly used dependent variable in is research [22, 33]. In a recent meta-analysis [18] 27 studies used some operationalization of us as the success measure, 17 employed usage, and 13 some other dependent variable. Furthermore, the meta-analysis found a significant, negative relation between effect sizes for usage and the year in which the studies had been carried out. This may well be a consequence of this measure's limited sensitivity.

Some studies apply multiple measures simultaneously and some attempts have been made to gain insight in the mutual relations between success measures and organizational performance. Gatian investigated the relation between us, 'decision-making performance' and 'efficiency.' Her research population consisted of two groups of university and college users of a financial accounting and accounts payable system: department heads and controllers. She found that there was a relatively strong positive relation between satisfaction and both decision performance (assessed by both user groups) and efficiency (only assessed by the controllers group). However, her results may be affected by the fact that her decision performance measure asked users about their perceptions of the contribution of the system to performance: the decision-performance measure may well be considered to be a us measure itself. The efficiency measures assess 'specifically, data processing correctness, report preparation and distribution timeliness' [17]; these variables do not seem particularly suited to assess the contribution of is to organizational performance. Furthermore, respondents may have tried to answer consistently: a respondent who first indicates that she is very satisfied with the is is unlikely to answer that the system has a negative influence on her performance.

Iivari and Ervasti investigated 21 different systems in a single municipal organization. For a group of users and a group of user-managers us scores were determined using a version of the Bailey and Pearson instrument that was adapted by the authors in order to be able to determine us with an individual system.<sup>2</sup> Furthermore, implementability of the system was

<sup>2</sup>The authors explicitly acknowledge that the Doll and Torkzadeh us instrument could have been used to assess the relation between performance and us.

assessed using a scale developed by the authors, and effectiveness of the organizational unit was determined using the Van de Ven and Ferry [34] organizational assessment framework. Iivari and Ervasti found a positive relation between us (in particular ease of use) and implementability. Further results are somewhat ambiguous, but point to a positive relation between us and unit performance.

Etezadi-Amoli and Farhoomand investigated the relation between a newly developed us instrument (similar to the Bailey and Pearson instrument) and a newly developed performance instrument. Their respondents were employed by 22 different organizations and 38% of the respondents occupied a managerial position. They find a strong relation between us and performance. However, the nature of the performance measures employed may have inflated the findings: users were asked about the contribution of the software to their performance.

Igbaria and Chan [21] investigated the influence of us – assessed by the Doll and Torkzadeh instrument – on system usage and 'individual impact.' They used questionnaire data provided by 371 employees of a large organization located in Singapore and found a significant, positive influence of us on both usage and individual impact and of usage on individual impact. However, their results should be interpreted with some caution. The authors did not define usage in terms of frequency of use, as the differences between respondents were rather small. Instead they chose to define usage by the number of computerized applications (e.g. word processor, spreadsheet) used by respondents and the number of business tasks (e.g. writing reports, communicating with others) for which they used the system. This measure might better be labeled 'computer experience.' The individual impact measure is similar to the success measures used by Gatian and Etezadi-Amoli and Farhoomand: users are asked whether the system contributes to "decision making quality, performance, productivity, and effectiveness of the job." The tendency of users to give consistent answers to the questionnaire may, again, have inflated the results.

In my study, unit performance was assessed using the Van de Ven and Ferry measures with two new questions concerning financial performance (revenues and profit). In this way the occurrence of spurious relations between us and performance found in pre-

vious studies is avoided. The Doll and Torkzadeh instrument will be used to assess us. This instrument has been validated extensively (e.g. [14, 19]) and measures satisfaction with an individual application; this eliminates the need for adaptation of the instrument. In order to counter the criticisms of Etezadi-Amoli and Farhoomand about two-item measures the first version of the instrument (before elimination of items deemed superfluous by Doll and Torkzadeh) was used. Analysis of the relation between the outcomes of this instrument and other performance measures is particularly interesting because it has been criticized for not including performance related variables.

#### 4. Research method

A questionnaire survey was sent to 1024 Dutch managers, information managers, and controllers.<sup>3</sup> A separate answer card was attached to the questionnaire; this could be used to obtain a booklet about us and the results of the survey. The respondents did receive a postage paid return envelope and a letter on university stationery. The letter asked for cooperation and guaranteed that answers would be private. Four weeks after the first mailing, a reminder was sent out in which respondents were thanked for their cooperation and in which the cooperation of people who had not yet responded was again solicited.<sup>4</sup> A final gross response rate of 20.7% ( $n=212$ ) was obtained. The net response rate was 16.6%, as 42 responses could not be used for analysis.<sup>5</sup>

The age of the respondents varied from 23 to 65 years, with an average of 44.9 years. On average respondents have worked 6.1 years in their current function and 11.2 years with their current employer. A large majority of the respondents (94.7%) was male. Of the respondents, 84% had at least a polytechnic,

<sup>3</sup>Administration of the questionnaire was made possible by financial support of Oasis Nieuwegein, which is gratefully acknowledged.

<sup>4</sup>The original intent of the answer card was to be able to keep track of respondents and non-respondents. However, the number of questionnaires received without any identification was quite large, and it was decided to send a reminder/thanks letter to all (non-)respondents.

<sup>5</sup>A large number of refusals consisted of a letter indicating company policy of non-cooperation in survey research.

Table 1  
Organizational function of respondents ( $n=170$ )

Function	%
Concern management	27.1
Division management	10.2
Business unit management	5.4
Line management	12.7
Staff member <sup>a</sup>	41.6
Other	3.0
Total <sup>b</sup>	100.0

<sup>a</sup>Of this group 23.2% indicated that their function was either information manager or controller.

<sup>b</sup>Due to rounding errors the sum of the individual items does not always equal 100%.

university, CPA or CMA degree.<sup>6</sup> An IS was available to 64.5% of the respondents, 26.7% used the system, only through an intermediary, and 11.4% did not use it at all. Table 1 presents descriptives of the function occupied by respondents.

The original Doll and Torkzadeh instrument consists of 18 questions, which are used to measure five dimensions of us: satisfaction with content (5 items), accuracy (4), format (4), and timeliness (2) of information and satisfaction with the system's ease of use (3). Two new indicators were added to the timeliness scale: "Are the data in the system updated *often* enough?" and "Are the data in the system updated *quickly* enough?" Both confirmatory factor analysis (CFA) and an expert panel were used to validate the resulting measurement instrument. They suggested the elimination of the fourth and fifth item of the content scale, the first newly added item of the timeliness scale and the third item of the ease of use scale. This reintroduces concerns that the number of indicators per construct is too low. However, the inclusion of only two items in the measure was preferred over the inclusion of a faulty item. The elimination of this item did result in an increase in Cronbach's  $\alpha$ . A possible explanation is that the formulation of this item ("Is the system efficient?") is ambiguous. After those analyses an extension of the CFA was used. In it a measurement model where all non-zero factor loadings were set equal to 1 was compared with a traditional measure-

<sup>6</sup>The subjects were Dutch managers, consequently the questions were phrased in terms of the Dutch educational system.

ment model where the non-zero factor loadings are left free. The difference in  $\chi^2$  between both models ( $\chi^2_{11} = 5.07$ ) is not significant. This indicates that in further analyses the sum of the individual item scores can be used. Finally, the reliability coefficients presented on the diagonal of Table 2 are quite satisfactory.

The second success measure, usage, was assessed in four different ways. The respondents were asked how many hours a week, and how many times a week they used their is. Both measures were also obtained for any indirect usage: the respondents were asked how many hours and times a week an assistant spend to get them results from the is.

Table 2

Pearson product moment correlations of is success measures with organizational performance. Underneath each correlation coefficient the number of cases it is based on the significance are presented. Reliability coefficients (Cronbach's  $\alpha$ ) for multiple item measures are presented on the diagonal of the matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Content	0.90 (103) n/a												
2. Accuracy	0.68 (103)	0.97 (103)											
3. Format	0.70 (103)	0.59 (103)	0.94 (103)										
4. Timeliness	0.59 (103)	0.76 (103)	0.56 (103)	0.91 (103)									
5. Ease of use	0.58 (103)	0.57 (103)	0.73 (103)	0.56 (103)	0.96 (103)								
6. us (1+2+3+4+5)	0.84 (103)	0.87 (103)	0.87 (103)	0.82 (103)	0.79 (103)	0.96 (103)							
7. Hours of direct usage	0.14 (100)	0.14 (100)	0.11 (100)	0.24 (100)	0.22 (100)	0.19 (100)	n/a (103)						
8. Hours of indirect usage	0.169 (93)	0.157 (93)	0.266 (93)	0.018 (93)	0.032 (93)	0.056 (93)	n/a (95)	n/a (96)					
9. Frequency of direct usage	0.169 (101)	0.133 (101)	0.906 (101)	0.116 (101)	0.724 (101)	0.275 (101)	0.000 (103)	0.08 (95)	n/a (104)				
10. Frequency of indirect usage	0.051 (98)	0.175 (98)	0.055 (98)	0.021 (98)	0.004 (98)	0.017 (98)	0.000 (99)	0.428 (96)	n/a (100)	n/a (101)			
11. Performance (Van de Ven and Ferry)	0.147 (100)	0.041 (100)	0.608 (100)	0.058 (100)	0.821 (100)	0.239 (100)	0.159 (100)	0.000 (94)	0.240 (101)	n/a (98)	n/a (162)		
12. Performance (new)	0.000 (92)	0.000 (92)	0.000 (92)	0.000 (92)	0.000 (92)	0.000 (92)	0.294 (92)	0.699 (88)	0.124 (93)	0.582 (91)	n/a (143)	n/a (146)	
13. Performance (11+12)	0.011 (91)	0.009 (91)	0.004 (91)	0.005 (91)	0.013 (91)	0.001 (91)	0.572 (91)	0.479 (87)	0.535 (92)	0.914 (90)	0.000 (143)	n/a (143)	n/a (143)
	0.38 (91)	0.36 (91)	0.42 (91)	0.38 (91)	0.37 (91)	0.46 (91)	0.09 (91)	0.01 (87)	0.17 (92)	0.04 (90)	0.95 (143)	0.70 (143)	0.84 (143)
	0.000	0.000	0.000	0.000	0.000	0.000	0.383	0.908	0.101	0.707	0.00	0.000	n/a

claim that *us* is an adequate proxy for the contribution of *is* to organizational performance is correct. However, the *us* instrument still needs some improvement. The ease of use component of the instrument consists of only two items and is relatively unreliable. Furthermore, reliabilities of the instrument may have been overestimated as a consequence of the tendency of respondents to give consistent answers. Thus the correlations between the subdimensions of *us* and the total *us* score are better estimates of reliability than the diagonal *as*.

The correlations between *us* and performance may be inflated by two artifacts of the research design. First, self-reports of performance were used. Second, the correlations may have been inflated by the fact that this study focused on *is* only; organizations that build better *is* are likely to perform better in other areas, too. All those areas contribute to organizational performance. As only *is* success is assessed, the relation between *is* success and organizational performance will partially reflect the contribution to performance of the other areas in which the organization performs better.

It should also be acknowledged that *us* is less suited to assess the success of an *is* that is used indirectly. However, in survey research it may not be obvious whether a manager is a user of the *is* or uses it only indirectly.

Finally, the low and insignificant correlations between usage measures and performance measures add weight to the doubts that already exist about the validity of usage as a success criterion. However, the results of this study should not be generalized to other kinds of systems. For some systems (e.g. Internet sites or other information systems aimed at a general public) usage may remain the most appropriate and most easily assessed success measure; for *is*, *us* measurement is more appropriate.<sup>3</sup>

<sup>3</sup>In a similar vein, the results of Chan, Huff, Barclay and Copeland indicate that using *us* as a success measure may underestimate systems' strategic importance: "The seven dimensions of *is* effectiveness formed two separate clusters - *uis* [user information satisfaction] and strategic impact - suggesting that *is* satisfaction measures were not just another way of assessing *is* strategic importance; these factors were very different. This in turn implies that studies investigating organizational *is* effectiveness may be well-advised not to rely solely on *uis*-based measures."

## Acknowledgements

This study has been carried out as part of my Ph.D research. I wish to thank the members of my reading committee: Gary Bamossy, Tom Groot, Cees van Halem, David Otley, Edu Spoor, and Berend Wierenga. This paper benefitted from comments by Ypke Hiemstra and both anonymous reviewers and from the detailed and accurate corrections of the editor.

## References

- [1] American Accounting Association. Report of the committee on management information systems. 1977.
- [2] James E. Bailey, Sammy W. Pearson. Development of a tool for measuring and analyzing computer user satisfaction. *Management Science* 29(5) 530-545. May 1983.
- [3] Henri Barki, Sid L. Huff. Change, attitude to change, and decision support system success. *Information and Management* 9(5) 261-268. December 1985.
- [4] Lloyd W. Belcher, Hugh Watson. Assessing the value of Conoco's *is*. *MIS Quarterly* 239-253. September 1993.
- [5] Robert P. Carveny, G. Lawrence Sanders. Implementation and structural variables. *Information and Management* 11. 1986, pp. 191-198.
- [6] Wynne W. Chin, Peter R. Newsted. The importance of specification in causal modeling: The case of end-user computing satisfaction. *Information Systems Research* 6(1). 1995, pp. 73-81.
- [7] David W. Conrath, Olivier P. Mignen. What is being done to measure user satisfaction with *spms*. *Information and Management* 19. 1990, pp. 7-19.
- [8] William H. DeLone, Ephraim R. McLean. Information systems success: The quest for the dependent variable. *Information Systems Research* 3(1). 1992, pp. 60-95.
- [9] Gary W. Dickson, James C. Wetherbe. *The Management of Information Systems*. McGraw-Hill. 1985.
- [10] William J. Doll, T.S. Raghunathan, Jeon-Su Lim, Yash P. Gupta. A confirmatory factor analysis of the user information satisfaction instrument. *Information Systems Research* 6(2). 1995, pp. 177-188.
- [11] William J. Doll, Gholamreza Torkzadeh. The measurement of end-user computing satisfaction. *MIS Quarterly* 12 259-274. June 1988.
- [12] William J. Doll, Weidong Xia, Gholamreza Torkzadeh. A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly* 453-461. December 1994.
- [13] Philip Ein-Dor, Eli Segev. Organizational context and the success of management information systems. *Management Science* 24(10) 1064-1077. June 1978.
- [14] Jamshid Etezadi-Amoli, Ali F. Farhoomand. On end-user computing satisfaction. *MIS Quarterly* 1-5. March 1991.

- [15] Jamshid Etezadi-Amoli, Ali F. Farhoomand, A structural model of end user computing satisfaction and user performance, *Information and Management* 30(2) 65-73, May 1996.
- [16] Charles A. Gallagher, Perceptions of the value of a management information system, *Academy of Management Journal* 17(1) 46-55, March 1974.
- [17] Amy W. Gatian, Is user satisfaction a valid measure of system effectiveness?, *Information and Management* 26, 1994, pp. 119-131.
- [18] Maarten Gelderman, Factors affecting the success of management support systems: Analysis and meta-analysis. Paper presented at the IS/MAS forum of the Annual Conference of the American Accounting Association, Orlando, FL, August 1995.
- [19] Maarten Gelderman, Translation and validation of the Doll and Torkzadeh end user computing satisfaction instrument. In: Hugh J. Watson (Ed.), *Proceedings of HICSS 31*, vol. VI. IEEE Computer Press, January 1998, pp. 537-546.
- [20] W.J. Haga, M. Zviran, Information systems effectiveness: Research designs for causal inference, *Journal of Information Systems* 4, 1994, pp. 141-166.
- [21] M. Igarria, M. Tan, The consequences of information technology acceptance on subsequent individual performance, *Information and Management* 32(3) 113-121, March 1997.
- [22] Magid Igarria, Sidney A. Nachman, Correlates of user satisfaction with end user computing: An exploratory study, *Information and Management* 19, 1990, pp. 73-82.
- [23] Juhani Iivari, Erja Ervasti, User information satisfaction: is implementability and effectiveness, *Information and Management* 27, 1994, pp. 205-220.
- [24] Blake Ives, Margrethe H. Olson, User involvement and MIS success: A review of research, *Management Science* 30(5) 586-603, May 1984.
- [25] Blake Ives, Margrethe H. Olson, Jack J. Baroudi, The measurement of user information satisfaction, *Communications of the ACM* 26(10) 785-793, October 1983.
- [26] Mo Adam Mahmood, Janet A. Sniezek, Defining decision support systems: An empirical assessment of end-user satisfaction, *INFOR Canada* 27(3) 253-271, August 1989.
- [27] M. Lynne Markus, Daniel Robey, The organizational validity of management information systems, *Human Relations* 36(3), 1983, pp. 203-226.
- [28] Nancy Paule Melone, A theoretical assessment of the user-satisfaction construct in information systems research, *Management Science* 36(1) 76-91, January 1990.
- [29] Jum C. Nunnally, *Psychometric Theory*, McGraw-Hill, 1967.
- [30] Margrethe H. Olson, Blake Ives, User involvement in systems design: An empirical test of alternative approaches, *Information and Management* 4(4), 1981, pp. 183-195.
- [31] Daniel Robey, User attitudes and management information system use, *Academy of Management Journal* 22(3), 1979, pp. 527-538.
- [32] Detmar W. Straub, Validating instruments in MIS research, *MIS Quarterly*, June 1989.
- [33] Ashok Subramanian, Sree Nilakanta, Measurement: A blueprint for theory-building in MIS, *Information and Management* 26, 1994, pp. 13-20.
- [34] Andrew H. Van de Ven, Diane L. Ferry, *Measuring and Assessing Organizations*, Series on Organizational Assessment and Change, Wiley, 1980.



( 管理 進 才 = 篇 )

## Using Intranets: Preliminary Results from a Socio-technical Field Study

### Abstract

*This paper describes research-in-progress that examines the relative influences of internal and external factors on the development and use of intranets, as a particular case of information and communication technologies (ICTs). This three-phase qualitative study integrates cross-sectional, comparative and traversal methodologies to identify where it could be effective to combine socio-technical theory (emphasizing individuals and social groups) with institutional approaches (emphasizing interaction through influence and regulation.) Preliminary results suggest that a balance of pressures are at work, with external factors motivating the use of intranets and other forms of web information systems within the constraints of local organizational contexts.*

### 1. Internal and external influences

We commonly assume that, within organizations, internal knowledge demands drive information technology use. These demands may be tactical, generating a need for information gathering, data analysis and knowledge management about core products and processes. Strategic demands may also generate a need for industry analysis and market data. They could even spur investment in new technologies to achieve greater production and communication efficiencies. When researchers examine organizational incentives to adopt and use information and communication technologies (ICTs), their studies concentrate attention on these internal knowledge demands [1, 2]. However, new research results indicate that external demands strongly drive ICT use. In a recent study, I

examined the use of online information by 26 California firms in three different industries. Data show that external environmental factors and interorganizational relationships strongly drive information investments and the use of online technologies [3]. Institutional demands, such as disclosure mandates and regulatory agency approval requirements, generate extensive data gathering activities. Interorganizational interactions further shape online use. Firm members frequently described their online use as a response to information demands arising from firm interactions with clients, competitors, partners and regulatory agencies. Some firms made much greater use of online information than others, and that had as much to do with external institutional and interorganizational factors as with any internal decisions of the firms' managers or staff. Do these external factors shape the organizational use of other information technologies as well? Do they shape the use of intranets?

### 2. Why study intranets?

This paper describes a study designed to examine whether the findings of my prior research can be extended to other industries and other ICTs, such as web information systems (WIS). I have chosen to study intranets for two reasons. First, intranets are an "inside-the-firm" technology configuration--unlike extranets or web-based electronic commerce networks that are intended for multi-organizational use. A study that examines the use of intranets, rather than these other new WIS, could characterize interorganizational influence less ambiguously. Thus, an intranet study might determine whether or not findings about the influence of external factors

could be extended to WIS in particular and to ICTs more broadly defined.

A second reason for choosing to study intranets when examining questions of technology shaping is that they are a new, highly configurable WIS that organizations are adopting at a rapid rate. According to a 1996 Forrester industry report, 40% of the firms surveyed had already implemented intranets, 25% were in the intranet planning stage, and 24% had begun to consider implementing one [4]. In the US, Forrester estimates intranet growth to be 60% annually, largely spurred by promises of ROIs exceeding 1000%. My own preliminary screening of midwest firms shows that in 1998, about 40% of the firms I contacted had intranets "up and running" or were just implementing them.

The use of intranets and internet technologies, is growing rapidly. Ironically, some information systems analysts express concern that they will be misused or underutilized [5]. Their concerns extend from firms that don't use intranets at all, to firms that use intranets simply as file servers, or as corporate web servers--not taking advantage of their potential capabilities. Some have suggested that intranets can be used as groupware, as LAN replacements, as local digital libraries, as document management systems, as Internet firewalls, and as web information systems [6, 7, 8]. But, at least for the moment, there is a gap between these anticipated uses of intranets and actual use. It would appear, then, that within many firms the potential usefulness of an intranet is not completely apparent, nor is actual use limited to a single style of implementation. Intranet implementation plans may very well be open to redirection and influence through local interorganizational interactions. A study of intranets at this point in time may provide a unique opportunity to examine what shapes the initial use(s) of a new ICT.

### **3. Organizations using technology**

For a long time now, researchers have tried to explain what shapes the organizational use of ICTs. A large body of research has focused on examining ICTs from an efficiency perspective. Often, this work looks at the fit between the task to be performed and the technology intervention [9, 10, 11]. These researchers try to measure how efficiently and effectively the technology aids an individual in executing a task. They concentrate

their investigations at the individual level and base their examinations on theories that describe how rational economic individuals make choices about tools and information resources [12, 13]. For these researchers, lower-than-expected use of a new technology often indicates problems with media accessibility and technology usability factors, and their prescriptions follow the ergonomic concepts of efficient and effective tool design [14, 15]. Their findings, however do not aggregate well, and they have not had much success in explaining organizational use of ICTs. Baldwin and Rice [16] confirm that individual preferences for ICTs have little impact within organizations.

### **3.1 CSCW studies**

Some researchers have tried to shift the level of focus of their efficiency and effectiveness studies toward individuals working within small groups and organizations. A wide-ranging set of research, loosely categorized as computer-supported cooperative work (CSCW) studies, have examined the situated use of ICTs within complex organizations. Their electronic mail studies, for example, have provided insights that explain the differential benefits of organization-level use of ICTs. Their work has shown the need for a "critical mass" of users, and has identified the wider social benefits of email use that may accrue to globally distributed organizations [17, 18]. Some CSCW researchers have also examined the incentives for using ICTs like groupware and internet-based help systems to augment organizational memory [19, 20, 21]. Their work shows that among different organizations, the same technologies will be used differently depending on the practices and incentives present in each organization. This line of research has not, however, had much luck in explaining what shapes new ICTs.

### **3.2 Information systems research**

Information systems (IS) researchers have had more success. Guided by theories about organizational change and workplace dynamics, they have tried to examine more specifically how organizational contexts can affect the use of a set of ICTs--an information system--and how these IS technologies may, in turn, present an occasion for

organizational restructuring. Some have compared the use of a single technology, by two or more organizations in the same industry [22, 23, 24]. They have characterized the organization-level incentives that lead to greater or lesser use of the technology, and how uses within some contexts can enhance the quality of worklife, while others can foster a "sweatshop" atmosphere. Their complex characterizations also show how an IS can trigger organizational change, and how reciprocally, the organizational context can spawn new ways of configuring ICTs and shaping their use.

### 3.3 Examining institutional factors

IS researchers, guided by institutional perspectives, have been particularly successful in making some connections between the use and the shaping of information technologies. These scholars have extended the scope of their studies to include the larger organizational environment--the social, economic, political and industry institutions that constrain and enable organizational action. Their work has identified the complex sets of interdependent resources needed to use information systems effectively within and among organizations [25]. Much of their work carefully examines the multi-faceted and often paradoxical incentives that shape information systems and information resource use in the public sector [26, 27, 28, 29, 30]. These studies highlight the interdependence of the public and private sectors, the strength of interorganizational ties, and the ways in which infrastructure supports particular types of interactions [31]. As King et al. [32] have noted, however, institutional and interorganizational contexts are rarely studied in conjunction with technology use. One criticism of institutional approaches has been that the individual is often portrayed as a passive recipient of new technology, rather than an active adopter or a creative interpreter.

Researchers guided by social constructionist and structurationist perspectives [33, 34], have also made some progress toward understanding what makes technologies more useful for some than others [35, 36]. Bijker [37] has begun developing a theory of the social construction of technology (SCOT) that explains how relevant social groups reshape the technologies they use based on what they know about the task domain and other technologies. IS

researchers have begun to explore the use of ICTs, such as groupware and management information systems, guided by socio-technical theories based on SCOT and structuration theory [38].

These institutional and constructionist perspectives also provide helpful explanations of complex adoption and use phenomena for WIS researchers. I believe that some interleaving of SCOT theory with institutional perspectives could help to situate individual and group activities within a dynamic, powerful, overarching industry and institutional setting. In this study of organizational intranet use, I intend to bring these two sets of concepts together to provide a more comprehensive explanation of how "information technology is situated in a web of interrelations that is intimately connected to other dimensions of the workplace" [25].

## 4. Methodology

To examine what shapes intranet use and development, I have designed a field study that involves three overlapping phases of data collection and analysis. (See Table 1.)

In the first phase, I will conduct a cross-sectional intranet status study within each of four industries: law, hospital care, real estate and general manufacturing. To ensure that the study encompasses firms with a wide-ranging set of external influences and conditions, I selected one industry from each quadrant of Scott's [39] technical and institutional industry environments framework. This industry selection will establish a basis for cross-case comparison [40]. My first task is to contact approximately 200 midwest organizations (50 in each industry) to determine their intranet status. I have chosen to conduct my research in Chicago and Cleveland, because I believe these cities are in a region which is representative of US organizational environments. The region is also convenient to my research facilities. I am using revenue-ranked Dun & Bradstreet listings of law firms, hospitals, chemical and metal products manufacturing companies, and real estate brokerages from which I pseudo-randomly select the sample firms of this study. I then telephone each organization to identify which firms currently have intranets and how long they have had them.

**Table 1: Phased Intranet Study Design**

	<b>Phase I</b>	<b>Phase II</b>	<b>Phase III</b>
<b>Analytic Focus</b>	Cross-Sectional	Comparative	Traversal
<b>Primary Industry</b>	Firms from Selected Industries	Manufacturing Company International Law Firm Hospital Real Estate Brokerage	Multiple Industries Firms & Agencies
<b>Participant Roles</b>	Information Ctr and MIS Mgrs	IC/MIS Dev/Mgrs Intranet Users/Non-Users ICT Strategists	Firm Contacts and Industry Coordinators
<b>Technology Focus</b>	Intranet Status	Intranets & Related ICT	Intranets & Related ICT

When this cross-sectional analysis is complete, I will have a snapshot of the status of intranet use within a set of firms that represent Scott's four categories of institutional and technical environments. The systematic screening process that I am using to select sites for this first phase will also provide the sampling criteria for my comparative study site selection. I will visit the sites that have used intranets for more than six months to determine whether they would be suitable for and willing to allow an in-depth, onsite study.

In the second phase, I will conduct the comparative case studies that form the basis of my theory building research [41]. I will select four sites from those I have visited--one in each industry--to participate in a qualitative study of intranet use. I intend to spend 4 to 6 weeks at each of the four sites, discussing and observing intranet use, in conjunction with other information systems and information resources that may be used for similar purposes. I will collect data from multiple sources within each organization, including direct examination of intranet content and usage statistics, observation of intranet use, semi-structured interviewing, and document collection [42, 43]. I will talk to intranet users, non-users, developers, managers and ICT strategists. An important part of the interview process will be a discussion of how these people coordinate interdepartmental and interorganizational activities. As I proceed from one industry case study to the next, I will use constant comparative methods to formulate data categories, identify cross-case patterns and develop theoretical leads for further investigation [44]. This data will allow for an analysis of the interplay between informational influences from both micro and macro environments, and an analysis of how these occasion the use of intranets in each firm. During this phase

of the study, I will record detailed descriptions of how each firm has developed and used its intranet(s), as well as the influences and incentives that firm members identify as motivators. I will chart the network of interorganizational relationships that anchor those influences, forming a list of organizational subsidiaries, other organizations and outside agencies that may influence each firm's intranet.

In the final phase of this research project, I will use Latour's methodology [45] to follow the links in these interorganizational networks. I will contact people at each organization that has a key relationship with one or more of the comparative study firms, and interview them about how information and information systems--including intranets and internet technologies--are used at their firm, and about the nature of their relationships with the comparative study firms. I expect these relationships to be of very different types. They may be formal or informal, and they may fulfill a variety of purposes. For example, some may be supplier-customer relationships, some may be competitor relationships, others relationships may involve regulatory oversight. I also want to identify which relationships influence intranet use and how they affect it. Some of these influences may be intentional and acknowledged, others may be implicit and unrecognized. Traversing the interorganizational network, as I follow up on each reference, will allow me to characterize the nature of these sets of technical and institutional influences on intranet use and development, and to identify the types of interorganizational interaction that shape intranets.

Sociotechnical studies are often comparative, sometimes cross-sectional, but only rarely traversal. I could have chosen other methods for examining the

interorganizational networks that influence intranet use and development at each of our comparative study firms. Graph theory, sociometry, clique detection, and centrality analysis have all been helpful in analyzing interfirm connections. These methods generally concentrate on measuring network density (the number of firms in the network that an organization is connected to) or network centrality (the path length that connects firms in a network, or subsets of firms in a network). They often assume that all interfirm connections are of the same type, for example in value chain alliances all links are between buyers and suppliers. I have chosen Latour's method for analyzing interfirm connections, because it allows me to richly characterize the informational dimension of each link type I encounter. Network traversal also promises to dovetail the constructionist and institutional underpinnings of this study--to provide a view of WIS shaping that incorporates both micro and macro environmental elements.

This coordinated, three-phase approach provides a powerful set of theory building tools, that each single phase cannot provide alone. To my knowledge, this process has not been used before, although the methodologies adopted are all well accepted by qualitative researchers. It is a tightly integrated process for collecting and examining data from individuals in organizations that focuses attention on the external environment and interorganizational relationships of the firm. It could establish a new way for WIS researchers to examine how networks of organizations use networking technologies.

## 5. Preliminary results

At the time of this writing, I have contacted over 50 MIS managers and Information Center directors in midwest firms. I have spoken to them about the use and development of intranets within their organizations, future intranet plans, and also the reasons for not developing intranets, at those firms that don't use and aren't planning to implement them. I have visited nine of the fourteen sites that reported using intranet technologies for longer than six months, and I have conducted one 6-week onsite study at a Fortune 500 manufacturing firm. (See Table 2.)

### 5.1 Intranets federations

One of the most interesting early results is that many firms implement several intranets within the same organization, *each one developed by a different group*. Such "federations of intranets" seem to result from multiple influences to use organizational networking technologies. This type of loosely connected intranet configuration has also been reported by other researchers [46].

The finding that intranets often develop in a federated configuration--with some intranet servers hosting a unique, heavily used application, clearly motivated by external factors--suggests that external context does indeed shape the use of 'inside-the-firm' technologies like intranets.

**Table 2: Current Status of Intranet Study**

Screening	Onsite Visits	Case Study
Locations: Cleveland, OH and Chicago, IL	Intranets Overviews	Comparative Intranet Study
Systematic Sampling from D&B Listings: Manufacturing Firms: 26 Legal Firms: 25	Firms with Intranets (6 mo. plus.): Have intranets: 14 Sites Visited: 9	Fortune 500 Manufacturing Firm: Interview Participants: 41 Intranets/External Sites: 12

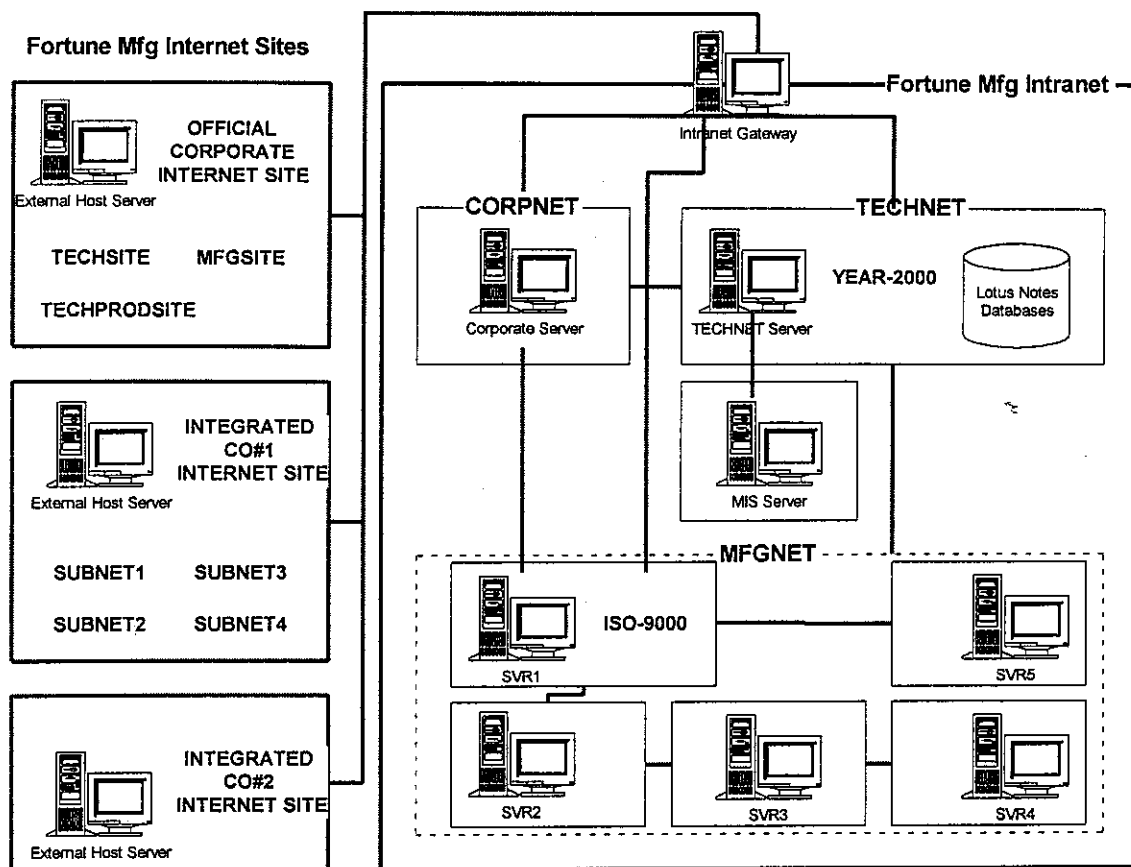


Figure 1: The Intranet/Internet Configuration of Fortune Manufacturing

**Common and uncommon uses.** In these firms, the intranets have developed independently, each reportedly influenced by the needs of the separate communities that implemented them, as Bijker's theory would predict. The intranet/internet configuration of Fortune Manufacturing (a pseudonym for the firm where I conducted the first comparative field study) serves as a good example of how a multi-server intranet federation might actually develop (see Figure 1.) In cases where business unit lines have been drawn along merger and acquisition boundaries (e.g. the separate SVRs associated within the MFGNET intranet in Figure 1), the intranets are more neatly separable as to who they serve and how they were influenced. Among these intranets, the redundancy of some applications, such as internal order forms, and services, like phone lists or online newsletters, can be explained by the fact that, while they may serve a need common to all members of the firm, the business units were originally unconnected to Fortune Manufacturing or to each other, and so did not always need to share this data or to pool

common concerns. But these common applications and services, while found on many intranet sites, have not been reported as motivating intranet development and active use. The unique applications motivate intranet development and use.

**Institutional and interorganizational influences.** Among the unique intranet applications and services at Fortune Manufacturing, two are well-developed and heavily used: the Year-2000 site on TECHNET, and the ISO-9000 site on MFGNET. ISO 9000 standardization clearly entails institutional factors. These standards define procedures and measures that help ensure the quality of manufacturing processes on an industry-wide basis. They are communicated within the firm, and made auditable by outside organizations, through a complex set of documentation. One business unit of Fortune Manufacturing has developed this documentation on their intranet server as a linked set of HTML pages and PDF forms. The ISO-9000 application serves the business unit quality management teams, as well as their customers--one

of which was noted as having been particularly influential in pushing for ISO 9000 compliance, and for setting even more stringent standards to guide industry-specific quality management processes.

The central piece of the Year-2000 site is a Lotus Notes database that tracks all the systems and software packages at Fortune Manufacturing, worldwide, and reports each system's status in terms of whether or not it can handle dates beyond December 31, 1999 correctly. It might be argued that bringing all systems to Year 2000 compliance status is in the best interests of the firm and would, therefore, be an internally motivated application. However, the recent demands that interacting organizations have placed on one another to certify their state of Year 2000 compliance in order to retain business has added a significant push to this effort. Although few firms can now certify that their systems are 100% Year 2000 compliant (or will be by January 1, 2000), managers at Fortune Manufacturing believe that it is important to be able to show that the firm knows its current status, and can demonstrate that it has a plan to become compliant. They speculate that, as the deadline approaches, the documentation that their Year-2000 site provides could help them retain customers who might otherwise give their business only to "certifiably 100% Year 2000 compliant" suppliers.

## 5.2 Intranet boundaries

Another interesting finding is that, because WIS technologies are highly configurable, it is difficult to define what is part of a firm's intranet, and what is not. At this point in the study, I define an intranet as "web-based technologies that are accessible inside the firewall." But that definition gets stretched a bit, because intranets routinely contain links to external internet sites, including a firm's own web presence sites.

**Definition.** "Mode of development" may help to additionally differentiate between what's an intranet and what's not. Among the firms of this study, external internet sites are almost exclusively developed and maintained by professional web developers. Intranets, on the other hand, are more often developed by people inside the firm, who are not primarily computer analysts or MIS developers, but who have a special interest in the technology and a specific need that it can serve. These folks may use consulting services to get started with the

technology, to solve specific implementation problems, or to create more professional looking web page artwork. But most intranet site development is done in-house and it's content-driven.

**Alignment of interests.** From a social constructionist point of view, it is fascinating that intranet sites are most often envisioned, developed and maintained by non-CIS/MIS authors--people who have no prior background and no formal training in systems development. Several people mentioned that, although initially they submitted their intranet content to someone who reformatted it into HTML pages, they had gradually assumed this responsibility themselves. They found that by using browser tools, simple editors, and the templates prepared by other firm members, they could just as easily manage their own content. Some of these self-taught novices are the "webmasters" of their intranet site. The affordances of the technology may explain some of these role-expansion instances. HTML pages can be developed in most of the word processors that people already have on their desktop computers. And an intranet server is not much more expensive than a typical desktop PC. These are technologies that many people are familiar with. However, they typically rely on their MIS group when it comes to some of the finer points of server monitoring or firewall protection.

Intranet authors include librarians, marketing analysts, human resource managers, research scientists, chemical engineers, and the occasional MIS person--and various implementations combine the efforts of these individuals. The ISO-9000 application at Fortune Manufacturing came together because the interests of a few key people aligned rather fortuitously--the quality manager needed to completely remaster his quality systems documentation and was open to the idea of providing it online, the technical writer wanted to learn HTML and was interested in tackling an intranet-based documentation project, and one of the MIS staff had been "fiddling around" with an intranet server and was willing to help implement a real application on that intranet.

These people may bring together a nicely integrated set of skills and local interests, but they may not share any wider concepts about WIS use and development. They may each be aligned with very different relevant social groups, and their conceptualizations about the uses of a new technology like intranets--what Bijker would call their "technological frames"--may differ greatly.

This could explain why, within a single company, the intranets that develop independently tend to look and feel very different from one another, even when the applications and services are common.

### 5.3. Intranet development

Until now, the way that intranets have developed at Fortune Manufacturing, and at some of the other study sites, has been through what might be characterized as the "grass roots" mobilization of local individuals motivated by external influences. Fortune management has just begun to think seriously about using their intranets strategically. The company has made some recent acquisitions, and it now needs to integrate those firms into Fortune. The corporate communications team has talked about using their intranets to share the company culture with these new firm members. Over the past year, a new executive management team has also taken the lead at Fortune, and they want to infuse the entire organization with their initiatives and their vision for the future. In fact, the chairman has issued a mandate to develop the company's intranets. To use Bijker's terms, executive management represents a new and powerful relevant social group that has begun to influence intranet development.

### 5.4. WIS forms

At Fortune Manufacturing, I can see some of the same external factors influencing intranet use that I saw influencing online information use in my prior research. These factors involve both institutional and technical pressures, such as the need to manage the documentation requirements of regulatory agencies, the need to document the processes of quality control regimens, and the need to provide proof of industry standards compliance; as well as the need to profile the firm's technical and technological competence to its clients. In my previous study, these factors influenced only the intensity of online use--whether firms used a lot or a little. However, with highly configurable WIS technologies, these factors also influence the *form* the technology takes on in use. Different groups within the firm adopt WIS technologies and shape them for their specific uses. When the people I have interviewed begin to talk about process control, for

example, they talk about their intranets. When they talk about displaying competence, they describe their external internet sites. And when they discuss the need to improve service to clients, they talk about their plans for an extranet.

### 6. Interim observations and next steps

At this point in my study, I am encouraged to see that I am finding some evidence to support my theoretical bases, and that my three-phase methodology is allowing me to collect and integrate meaningful data about how firms use intranets. As institutionalists would predict, it seems that within organizations, both technical and institutional influences separate groups and focus their perceptions about technologies and the uses they make of them. I can also see that the interorganizational connections that convey these influences are of very different types. At Fortune, I have identified at least three types of links: industry coordination (e.g. ISO 9000 quality control documentation), professional practice (e.g. models for knowledge sharing among librarians), and client interactions (e.g. demands for fast problem-solving information.) Some links appear to be direct, but others are indirect and unacknowledged. Therefore, my decision to spend time traversing and characterizing these links seems justified, as different types of interconnections may allow for different types of influences and different processes to shape the use of intranets.

My next steps in this study will be to determine whether the intranet observations that I have noted while at Fortune are helpful in explaining intranet use at another firm in a very different industry. I am currently screening the remaining manufacturing sample firms to determine their intranet status, and I am preparing for a second comparative study at a midwest-based international law firm. I am also beginning to follow the interorganizational network connections that Fortune Manufacturing informants have given me.

As I continue, my examination of intranets will explicitly focus on external factors and interorganizational relationships, as they interact with internal influences to shape intranet implementation. Bijker's SCOT theory suggests that relevant social groups--whether internal or external to a firm--have the ability to shape ICTs early in the technology lifecycle. If this is the case, an intranet



study like this one that focuses on these influences at this point in time, could document how the shaping process occurs. By fusing institutional and socio-technical concepts in my field study analyses, I hope to provide insights that extend constructionist theories about ICT use and development.

## 7. References

- [1] Davenport, Thomas H., Michael Hammer and Tauno J. Metsisto (1989). "How Executives Can Shape Their Company's Information Systems", *Harvard Business Review*, March-April, pp. 130-134.
- [2] Applegate, Lynda M., F. Warren McFarlan and James L. McKenney (1996). "IT Architecture Implementation Issues in the 1990's", in *Corporate Information Systems Management: Text and Cases, 4th Edition*, Chicago, IL: Irwin.
- [3] Lamb, Roberta (1997). "Interorganizational Relationships and Information Services: How Technical and Institutional Environments Influence Data Gathering Practices", Unpublished PhD Dissertation, University of California, Irvine.
- [4] Pincince, T. J., D. Goodtree and C. Barth (1996). "The Forrester Report: Network Strategies", Forrester Research, Inc.
- [5] Chellappa, Ramnath, Anitesh Barua and Andrew B. Whinston (1996). "Intranets: Looking Beyond Internal Corporate Web Servers" in Kalakota, Ravi and Andrew B. Whinston (Eds.) *Readings in Electronic Commerce*, Reading, MA: Addison-Wesley.
- [6] Stahl, Stephanie (1996). "Document management arrives on the Web", *InformationWeek*, n570, Mar 11, p.21.
- [7] Hills, Mellanie (1997). *Intranet as Groupware*, New York: John Wiley and Sons, Inc.
- [8] Gibbs, Mark (1997). "An intranet is an intranet, but is it also an extranet?", *Network World*, v14n4, Jan 27, p.66.
- [9] Bellardo, T. (1985). "An Investigation of Online Searcher Traits and Their Relationship to Search Outcome", *Journal of the American Society for Information Science*, 36(4):241-250.
- [10] Borgman, C. L. (1989). "All Users of Information Retrieval Systems Are Not Created Equal: An Exploration of Individual Differences", *Information Processing and Management*, 25(3):237-251.
- [11] Sugar, W. (1995). "User-Centered Perspective of Information Retrieval Research and Analysis Methods", in M. E. Williams (ed), *Annual Review of Information Science and Technology (ARIST)*, Volume 30, Medford, NJ: Information Today.
- [12] March, J. (1978.) "Bounded Rationality, Ambiguity, and the Engineering of Choice," *Bell Journal of Economics*, (Autumn), 587-608.
- [13] Simon, H. A. (1979). "Rational Decision Making in Business Organizations," *American Economic Review*, 69 (September), 493-513.
- [14] Fox, E., D. Hix, L. T. Nowell, D. J. Brunei, W. C. Wake, L. S. Heath and D. Rao (1993). "Users, User Interfaces and Objects", *Journal of the American Society for Information Science*, 44(8):480-491.
- [15] Norman, D. (1993). *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*, Reading, MA: Addison-Wesley.
- [16] Baldwin, N. S. and R. E. Rice (1997). "Information-Seeking Behavior of Securities Analysts: Individual and Institutional Influences, Information Sources and Channels, and Outcomes", *Journal of the American Society for Information Science*, 48(8):674-693.
- [17] Markus, M. L. (1990). "Toward a 'Critical Mass' Theory of Interactive Media", in J. Fulk & C. W. Steinfield (eds), *Organizations and Communication Technology*, Newbury Park, CA: Sage.
- [18] Sproull, L. and S. Kiesler (1991). *Connections: New Ways of Working in the Networked Organization*, Cambridge, MA: MIT Press.
- [19] Grudin, J. (1989). "Why Groupware Applications Fail: Problems in Design and Evaluation", *Office: Technology and People*, 4(3):245-264.
- [20] Ackerman, M. S. (1994). "Definitional and Contextual Issues in Organizational and Group Memories", *Proceedings of 27th Hawaii International Conference on System Sciences*.
- [21] Orlikowski, W. J. (1996). "Learning From Notes: Organizational Issues in Groupware Implementation", in Rob Kling (ed), *Computerization and Controversy: Value Conflicts and Social Choices, 2nd edition*, San Diego, CA: Academic Press.
- [22] Barley, S. (1986). "Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments", *Administrative Science Quarterly*, Vol. 31, pp.78-108.

- [23] Zuboff, S. (1988). *In the Age of the Smart Machine: The Future of Work and Power*. New York: Basic Books.
- [24] Kling, R. and S. Iacono (1989). "Desktop Computerization and the Organization of Work", *Office: Technology and People*, 5(1):7-28.
- [25] Kling, Rob (1988). "Defining the Boundaries of Computing Across Complex Organizations", in R. Boland and R. Hirschheim (Eds.) *Critical Issues in Information Systems*, New York: John Wiley.
- [26] Kraemer, K. L., S. Dickhoven, S.F. Tierney and J. L. King (1987). *Datawars: The Politics of Modeling in Federal Policymaking*. New York: Columbia University Press.
- [27] Danziger, J. D. (1991). "Management information systems and interorganizational relations within the American governmental system", *Informatization and the Public Sector*, 1(1):169-187.
- [28] Czarniawska-Joerges, Barbara (1992). *Exploring Complex Organizations*, Newbury Park, CA: Sage Publications.
- [29] Feldman, Martha S. (1989). *Order Without Design: Information Production and Policy Making*, Stanford, CA: Stanford University Press.
- [30] Powell, W. W., K. W. Koput and L. Smith-Doerr (1996). "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology", *Administrative Science Quarterly*, Vol. 41, pp. 116-145.
- [31] Tolbert, P. S. (1995). "Institutional Environments and Resource Dependence: Sources of Administrative Structure in Institutions of Higher Education," *Administrative Science Quarterly*, (30): 1-13.
- [32] King, John Leslie, Vijay Gurbazani, Kenneth L. Kraemer, F. Warren McFarlan, C.S. Yap, and K.S. Raman (1994). "Institutional Factors in Information Technology Innovation", *Information Systems Research*, Vol. 5. No. 2 (July, 1994):139-169.
- [33] Giddens, Anthony (1984). *The Constitution of Society*, Berkeley: University of California Press.
- [34] Berger, P. L. and L. Luckman (1967). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*, New York: Doubleday.
- [35] Latour, Bruno and Steve Woolgar (1979). *Laboratory Life: The construction of scientific facts*, Princeton, NJ: Princeton University Press.
- [36] Bijker, W. E., T. P. Hughes and T. Pinch (eds.) (1987). *The Social Construction of Technological Systems*, Cambridge, MA: The MIT Press.
- [37] Bijker, W. E. (1995). *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*, Cambridge, MA: The MIT Press.
- [38] Orlikowski, Wanda J. and Debra C. Gash (1994). "Technological frames: making sense of information in organizations", *ACM Transactions on Information Systems*, 12(2), 174-207.
- [39] Scott, W. R. (1987). *Organizations: Rational, Natural, and Open Systems, Second Edition*, Englewood Cliffs, NJ: Prentice Hall.
- [40] Yin, R. K. (1989). *Case Study Research: Design and Methods*, Revised Edition, Newbury Park, CA: Sage Publications.
- [41] Eisenhardt, K. M. (1995). "Building Theories From Case Study Research", in G. P. Huber and A. H. Van de Ven (eds), *Longitudinal Field Research Methods: Studying Processes of Organizational Change*, Thousand Oaks: Sage Publications.
- [42] Miles, M. B. and A. M. Huberman (1994). *An Expanded Sourcebook: Qualitative Data Analysis, Second Edition*, Thousand Oaks, CA: Sage Publications.
- [43] Lofland, J. and L. H. Lofland (1995). *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis, Third Edition*, Belmont, CA: Wadsworth Publishing.
- [44] Strauss, A. and J. Corbin (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage Publications.
- [45] Latour, Bruno (1987). *Science in Action: How to follow scientists and engineers through society*, Cambridge, MA: Harvard University Press.
- [46] Hughes, D. (1996). "Internet use grows rapidly at leading aerospace firms", *Aviation Week & Space Technology*, v144 n8, Feb 19, p.50.

# 國立中山大學八十八學年度博士班招生考試試題

科目：論文評述（資訊管理學系博士班）—科技組題目

1. Please identify the research contributions of this paper. (15%)
2. What are “open objects” in this research? (15%)
3. Is the cost model employed in this research reasonable? What are additional cost components needed to be included in the cost model? Please develop a more realistic cost model. (25%)
4. Please explain READ algorithm (as shown in Figure 2) and WRITE algorithm (as shown in Figure 3). (20%)
5. What are the disadvantages of the proposed method? Please propose another replication scheme for mobile data management. Please also describe the information requirement essential to your replication scheme. (25%)

## An Active Replication Scheme for Mobile Data Management \*

## Abstract

*A replication scheme determines the number and location of replicas in a distributed system. Traditional static replication schemes do not perform well in mobile environment since the assumptions of fixed hosts and relatively static access patterns no longer hold. For effective data management in mobile environment, we propose a dynamic replication scheme which employs user profiles for recording users' mobility schedules, access behavior and read/write patterns, and actively reconfigures the replicas to adjust to the changes in user locations and system status. Simulation results demonstrate that the scheme can accurately predict the data requirement to facilitate effective replication, reduce response time, and increase data availability.*

## 1. Introduction

Data replication [5] is often employed to improve the availability and effectiveness of information services in distributed systems. A replication scheme determines the number (*replication level*) and location (*replication placement*) of replicas in a distributed system. Traditional replication schemes [3, 5, 7] are static in the sense that the number and placement of replicas are predetermined and fixed. Manual re-calculation of the access cost and redistribution of replicas are necessary to reflect new access patterns. This is acceptable in traditional distributed environment since the hosts are installed at fixed locations and the access patterns are relatively static. In mobile environment, however, static replication schemes do not perform well because the assumptions about fixed hosts and static access patterns no longer hold [2, 4]. Dynamic replication schemes

[1, 13], on the other hand, try to overcome the problem by continuously maintaining statistics about access patterns and system workload so as to dynamically recalculate access cost and reconfigure the replication structure to adjust to the changes in access patterns. This is particularly desirable for mobile computing environment [6, 10, 11]. Even with dynamic replication, these schemes are essentially *passive* in the sense that the actual reconfiguration can only happen after the changes in access patterns have been taken place for a period of time and reflected on the access statistics.

We improve upon existing dynamic replication algorithms and propose an active replication scheme which employs user profiles for recording mobile users' mobility schedules, access behavior and read/write patterns, and actively reconfigures the replicas to adjust to the changes in user behavior and system status. Our scheme is unique in several respects: (1) We maintain detail access statistics of each individual user in his/her profile such that the replication decision can be made to tailor the movement and data requirement of each user. (2) We allow a user to provide his/her daily schedule which is used to derive the user's mobility pattern and data requirement. This enables the system to provide a certain degree of predictive replication which allocates object replicas before the actual access. (3) We devise the concept of open objects to represent a user's current and near future data requirement. This leads to a more precise and responsive cost model to reflect the changes in access patterns. (4) We allow the declaration of emergency events and objects which are unconditionally replicated. This is targeting safety and time critical application domains.

For performance evaluation, we built a simulation environment and compared five representative schemes: (1) no replication; (2) static replication; (3) dynamic data allocation [14]; (4) adaptive data replication [15]; and (5) the proposed active replication. Our scheme can accurately predict the data require-

\*This work was supported in part by National Science Council under project number NSC 87-2213-E-259-004.

ment to facilitate effective replication, reduce response time, and increase data availability.

The rest of the paper is organized as follows. Section 2 provides a survey of related issues and research work. Section 3 presents our framework and system model. Section 4 introduces the active replication algorithms. In Section 5, we discuss our simulation method and the results of performance evaluation and comparison. Section 6 concludes the paper.

## 2. Related Work

The access cost of most dynamic replication schemes are calculated based on the accumulated read/write statistics and the chosen consistency control protocol [6, 11, 15]. Consider the case of deciding whether to replicate an object  $O$  on a site. Assume that  $R$  is the average number of local read requests per time unit to  $O$  on that site, and  $W$  is the average number of write requests to  $O$  made by all users in the entire system. Let  $\alpha$  be the cost that can be saved if the object were read locally instead of requested from a remote site, and  $\beta$  be the additional cost that must be spent for maintaining a replica of  $O$ . Obviously, during a time unit,  $\alpha \times R$  is the total cost that can be saved if a replica of  $O$  is allocated on the site, and  $\beta \times W$  is the total cost that must be spent to maintain this replica. Based on this cost model, we can determine that if  $\alpha \times R \geq \beta \times W$  then the allocation of a replica of  $O$  on the site is *judicious*. On the other hand, if  $\alpha \times R < \beta \times W$  then it is not cost worthy to do so.

Some other design dimensions have also been considered, such as the lower and/or upper limit on the number of replicas of an object [8], as well as the capacity of a site [11]. To set the limit on the number of replicas is to ensure certain level of availability with bounded overhead. Another interesting design is to require that all replicas of an object reside on neighboring sites [15]. This has the advantages of reducing replica allocation and consistency maintenance cost. For objects with strong interdependency, it is sometimes better to replicate the entire group of related objects at the same time. This type of semantic information is used in [9] to improve the efficiency of replication. It is also possible to formulate the replication problem from an economy point of view, i.e. to consider replication as a trade behavior. The servers are the trading parties and the goods for trade are the data objects. Each server can determine whether to buy or sell object replicas, or decide whether to maintain, keep or discard the replicas it has. Such decisions are made based on the data requirement of the local users as well as the maintenance cost to keep the replicas up to date. This approach is

taken by the replication subsystem of Mariposa [12].

To summarize, a dynamic replication scheme must determine, among other things: (1) a consistency control protocol; (2) a cost model for estimating the access and replica maintenance cost; (3) the replication structure; and (4) the replication control algorithm(s). The scheme we proposed employs the primary copy model and the ROWA (read-one-write-all) protocol for consistency control, and an improved cost model upon existing dynamic replication schemes for cost estimation. We do not place any limit on the level or placement of object replicas.

## 3. System Model

The active replication scheme was designed to work in a wired/wireless LAN environment. Servers connected by traditional fixed network are the sites for allocating and maintaining object replicas. Servers that are capable of providing wireless communication services are called mobile support stations (MSS). The radio coverage of a MMS is called a cell. Any computational device which is carried with a mobile user and is capable of wireless communication is called a mobile host. Each user has a profile to keep track of the user's current location and related information. A hand-off process is taken between two adjacent MSSs to ensure seamless transition when users move from cell to cell. This also implies that a MSS can detect any entrance and exit of users into and out of its cell.

To simplify our algorithm design, we adopt the popular primary copy model and ROWA protocol for replica consistency control. In this model, we select, among an object  $O$  and all its replicas (named  $O$ -scheme), one as the primary copy ( $P(O)$ ). The site that maintains  $P(O)$  must have all information about the replicas of  $O$ , including the number and locations. All sites that keep a replica of  $O$  must know where  $P(O)$  resides. All reads to  $O$  can be satisfied by the nearest copy of  $O$ ; while all writes to  $O$  must be sent to  $P(O)$  first and then propagate to all sites in the  $O$ -scheme for replica update. The cost of replica update is proportional to the size of the  $O$ -scheme. We note that more deliberate models and consistency protocols can also be used with our framework. The detail of adopting other models or protocols, however, is not within the scope of this paper.

## 4. The Active Replication Scheme

Our method improves upon existing dynamic replication schemes with special design toward mobile environment. We employ user profiles to record users'

read/write patterns so that our algorithms can be tailored to satisfy as closely as possible each individual user's information requirement. For predicting future access pattern, we offer the opportunity for a user to specify daily schedule as hints of the user's mobility pattern and data requirement. This is desirable in practice since mobile users do not move at random. They often come to a location at a predetermined time with a specific purpose in mind. Also the work a user is currently engaged in has a strong relationship with the data required. This is why a user's daily schedule can be served as valuable hints for predicting the future. However, a user may not follow schedule strictly. In such case, we resort to past statistics for making replication decision. In using the read/write histories, we propose the concept of *open objects* for better cost estimation. We selectively adopt the read/write statistics of only those objects that are currently in use or likely to be used in the near future. The following sections discuss detail of the active replication algorithms.

## 4.1 System Information

### User Profiles

User profiles are the places for declaring schedules and maintaining per-user access statistics. A schedule is a declaration of daily activities. Each schedule entry declares the time, location, and activity a user plans to do with optional data objects requirement. A mapping is performed to determine the default data requirement of a schedule entry by considering, for example, characteristics of the user, location, and activity. The default requirement is unioned with the explicitly requested data objects to form the information requirement. Figure 1 is an example schedule in a hospital environment. An asterisk signifies an emergency object which are unconditionally replicated. Schedules are explicit hints about users' mobility and access patterns. They have been used regularly to record our daily activities. Where we are and what we do have a strong relationship with what we need for doing our job. Our method simply takes advantages of existing practice for improving information services. A user who appears at the right place and time is called "a user on schedule". Otherwise, he/she is called "a user off schedule".

The read/write histories are a user's access statistics on data objects in the system. Since a user accesses only part of the data objects at a time and location, we capture this characteristic by defining the concept of *open objects* (in a similar sense as open files) to be the set of objects accessed since entering the current cell and, for a user on schedule, the set of data objects declared explicitly or implicitly in the current schedule

Schedule of Dr. Taylor		
Time	Location	Activity & Objects
00:00 - 01:50	OR1	Operation32
10:30 - 12:00	PD1	Outpatient: ER3906*
14:00 - 16:00	MD3	Outpatient: ER3906*
16:10 - 19:00	OR5	Operation36
21:00 - 21:50	PL1	Project24: PR2409
P.S.		
MD: Medical Dep.	DR: Delivery Room	
PD: Pediatrics Dep.	OR: Operation Room	
PL: Pathology	ER: Epidemic Report	
PR: Pediatric Report	*: Emergency Object	

Figure 1. An example schedule.

entry. Open objects represents current and near future data access range. The information is used in calculating the access cost when making replication decision.

### Replication Server

A replication server must maintain information about local users and the object replicas, handle read requests and replica update messages. The information about local users includes user profiles and each user's on/off schedule status. For each data object, a server maintains an access record called *local open read* which is the sum of all read histories of local users having this object opened. This is used for representing the read pattern of the current cell toward the object. Since the past read histories of users who no longer have this object opened are not counted, our algorithm is more precise and responsive to the changes in access pattern than algorithms that simply accumulate all the past read requests to the object.

For a server that maintains a primary copy of *O*, a record of the current *O*-scheme must be kept up to date. A global access record called *global open write* is used to reflect the current write pattern to object *O* in the entire system. This is the sum of the write histories of all users who have the object opened.

### Emergency Events and Objects

We allow the declaration of *emergency events* and *emergency objects* for must-have objects that require fast access. An emergency event is any situation that demands quick response. In many application domains, such events can be identified in advanced. An emergency event usually has a set of default information requirement that can also be identified beforehand. Whenever such event occurs, the system unconditionally replicates all emergency objects associated with

that event. A user can also declare emergency objects in his/her schedule. To provide emergency services, a server must maintain a *local\_emergency\_counter* for each object declared as emergent to count the number of emergency events or claimed users to that object. The object is replicated whenever the counter is greater than zero.

### System Events

Our algorithms are activated upon the occurrence of certain events. The system events that trigger the execution of the replication control code include the READ event, the WRITE event, the UPDATE event (a server receives a consistency update message), the ENTER/EXIT event (a user enters/leaves a cell), the TIME-CHECK event (the system performs a periodic check for replica management), and the EMERGENCY event (the occurrence of the predefined emergency event). On each event occurrence, a corresponding part of the replication control code is triggered.

### 4.2 Cost Model

Our cost model follows the basic idea of comparing cost saving of allocating a replica with that of replica maintenance cost (Section 2). The access cost is calculated based on network transmission cost. A local access does not incur any transmission cost while the cost of a remote access is counted as the network distance between the current site and the nearest site with a replica of the desired object. If the distance is  $d$ , then the cost per access that can be saved from allocating a replica in the current site is  $d$ . Similarly, the extra cost of maintaining the replica on each update is also  $d$  since the message is propagated from the nearest site in the  $O$ -scheme. If a replica of  $O$  is allocated in the current site, then the total saving per time unit is  $d \times local\_open\_read$ , while the extra replica maintenance cost is  $d \times global\_open\_write$ . Naturally, when  $d \times local\_open\_read \geq d \times global\_open\_write$  (or equivalently,  $local\_open\_read \geq global\_open\_write$ ) it is beneficial to have a replica in the current site. Otherwise it is not cost worthy to do so.

### 4.3 The Replication Algorithms

For ease of presentation, we define the following symbols:

- $O_1, \dots, O_k$  are the data objects in the system.
- $B_1, \dots, B_n$  are the base stations in the system.

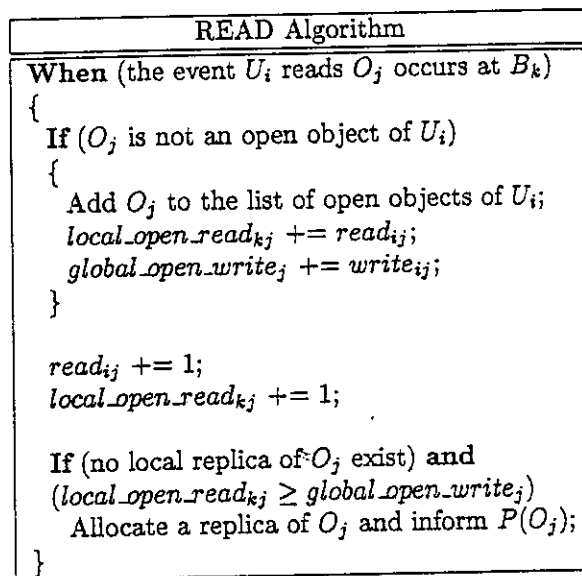


Figure 2. The READ Algorithm.

- $C_1, \dots, C_n$  are the corresponding cells.
- $U_1, \dots, U_m$  are the mobile users in the system.
- $P_1, \dots, P_m$  are the user profiles.
- $S_1, \dots, S_m$  are the user schedules.

We use  $local\_open\_read_{kj}$  to denote the local open read record of  $B_k$  on object  $O_j$ . Similarly,  $global\_open\_write_j$  is the global open write record of the object  $O_j$  maintained in  $P(O_j)$ . The local emergency counter of  $B_k$  on object  $O_j$  is represented by  $local\_emergency\_counter_{kj}$ . We also use  $read_{ij}$  and  $write_{ij}$  to denote the read and write histories of  $U_i$  on object  $O_j$ . Both of these records are maintained in  $P_i$ .

Each server runs an event detector to detect the occurrences of system events described in Section 4.1. Based on the event type, the corresponding replication control module is triggered and executed. There are eight modules for handling different types of situations. The algorithms used in each module are presented in Figure 2 through 9.

### 5. Simulation and Comparison

We have developed a mobile information system simulation environment which allows us to model and experiment on various network configurations, user mobility and access patterns. We compare the performance of five representative replication schemes: (1) No replication, (2) static replication, (3) dynamic data allocation [14], (4) adaptive data replication [15], and our

WRITE Algorithm
<pre> When (the event <math>U_i</math> writes <math>O_j</math> occurs at <math>B_k</math>) {   If (<math>O_j</math> is not an open object of <math>U_i</math>)   {     Add <math>O_j</math> to the list of open objects of <math>U_i</math>;     <math>local\_open\_read_{kj} += read_{ij}</math>;     <math>global\_open\_write_j += write_{ij}</math>;   }    <math>write_{ij} += 1</math>;   <math>global\_open\_write_j += 1</math>; } </pre>

Figure 3. The WRITE Algorithm.

EXIT Algorithm
<pre> When (the event that <math>U_i</math> exits <math>C_k</math> occurs) {   For all open objects <math>O_j</math> of <math>U_i</math>   {     Remove <math>O_j</math> from the open objects of <math>U_i</math>;     <math>local\_open\_read_{kj} -= read_{ij}</math>;     <math>global\_open\_write_j -= write_{ij}</math>;      IF (<math>O_j</math> is an emergency object requested)       <math>local\_emergency\_counter_{kj} -= 1</math>;      IF (<math>O_j</math> has a replica in the site) and       (<math>local\_emergency\_counter_{kj} = 0</math>) and       (<math>local\_open\_read_{kj} &lt; global\_open\_write_j</math>)       Discard the replica of <math>O_j</math> and inform <math>P(O_j)</math>;   } } </pre>

Figure 6. The EXIT Algorithm.

UPDATE Algorithm
<pre> When (an update to a replica of <math>O_j</math> occurs at <math>B_k</math>) {   IF (<math>local\_emergency\_counter_{kj} = 0</math>) and   (<math>local\_open\_read_{kj} &lt; global\_open\_write_j</math>)   Discard the replica of <math>O_j</math> and inform <math>P(O_j)</math>; } </pre>

Figure 4. The UPDATE Algorithm.

TIME CHECK Algorithm
<pre> When (a local time-check event occurs) {   For each local <math>U_i</math>   {     IF (the on/off schedule status has changed)     {       call EXIT Algorithm for <math>U_i</math>;       call ENTER Algorithm for <math>U_i</math>;     }   } } </pre>

Figure 7. The TIME CHECK Algorithm.

ENTER Algorithm
<pre> When (the event that <math>U_i</math> enters <math>C_k</math> occurs) {   IF (<math>U_i</math> is on schedule)   {     For all <math>O_j</math> required in <math>S_i</math>     {       Add <math>O_j</math> to the list of open objects of <math>U_i</math>;       <math>local\_open\_read_{kj} += read_{ij}</math>;       <math>global\_open\_write_j += write_{ij}</math>;        IF (<math>O_j</math> is an emergency object)         <math>local\_emergency\_counter_{kj} += 1</math>;        IF (no local replica of <math>O_j</math> exist) and         ((<math>local\_open\_read_{kj} \geq global\_open\_write_j</math>) or         (<math>local\_emergency\_counter_{kj} &gt; 0</math>))         Allocate a replica of <math>O_j</math> and inform <math>P(O_j)</math>;     }   } } </pre>

Figure 5. The ENTER Algorithm.

EMERGENCY ENTER Algorithm
<pre> When (an emergency event occurs at <math>B_k</math>) {   For all <math>O_j</math> required in the emergency event   {     <math>local\_emergency\_counter_{kj} += 1</math>;      IF (no local replica of <math>O_j</math> exist)       Allocate a replica of <math>O_j</math> and inform <math>P(O_j)</math>;   } } </pre>

Figure 8. The EMERGENCY ENTER Algorithm.



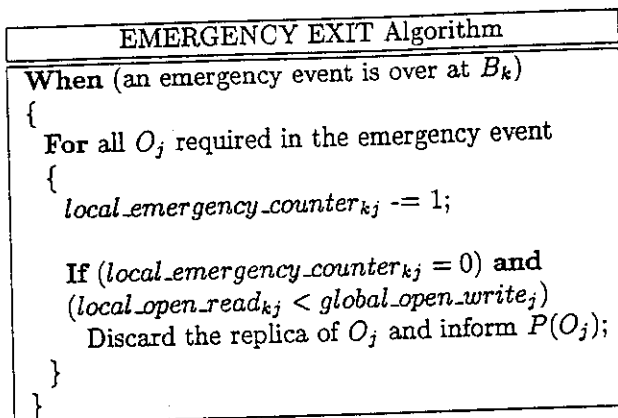


Figure 9. The EMERGENCY EXIT Algorithm.

(5) active replication algorithm. For each scheme, we measure the average access cost (both read and write), average response time (for read request), and average local availability of data. Both the access cost and response time are computed based on network distance. The local availability is the percentage of read access that can be satisfied locally. Because of the space limit, we only present the results on access cost and local availability.

We classify sites and data objects into classes. An *access class* is a set of objects that are routinely accessed by some fixed set of users. A *location class* is a set of sites that are constantly visited by some fixed set of users. Each user is assigned an access class and a location class to represent his/her access practice and mobility range. The variation of user behavior is controlled by the following parameters: (1) *movement locality* (the probability of moving within location class); (2) *access locality* (the probability of accessing within access class); (3) *schedule conformability* (the degree to which the user's actual behavior follows his/her schedule); (4) *write ratio* (the ratio of a user's write requests with respect to all requests). We have conducted a variety of simulation on different setting of the parameters. For each set of experiment, we varied one of the parameters while keeping others fixed so as to isolate the effect of the former.

## 5.1 Movement Locality

From Figure 10, it can be seen that the higher the locality the lower the average access cost. This is because when the users tend to move within a fixed area, most replicas are allocated around that area which results in lower access and replica maintenance cost. Higher movement locality also results in higher local availability (Figure 11). When the locality is low, however, our

active replication scheme is the only one that can still maintain a good level of data availability. This is due to the use of schedule and predictive replication.

## 5.2 Access Locality

It can be observed from Figure 12 that the higher the access locality the lower the average access cost. Stronger locality implies a more stable access pattern. This is advantageous to all replication algorithms, even for the case with no replication (since the initial distribution of data is also made according to the access class partition). From Figure 13 we can see that except for no replication, all algorithms achieve higher availability when the users reveal stronger access locality. For both set of experiments, our algorithm performs consistently better than all other algorithms.

## 5.3 Schedule Conformability

From Figure 14 we can see that schedule conformability does not have significant impact on all five algorithms. This is conceivable for algorithms which do not employ schedule. For our algorithm, a user conforms to his/her schedule simply means he/she shows up at the right place and time. It does not place any constraint on the objects accessed by the user. Remote access is still allowed. Furthermore, access cost includes write cost. Therefore schedule conformability does not have a strong impact on access cost. The situation is different, however, for local availability (Figure 15). For other algorithms, schedule conformability has no effect. For active replication, we observed higher local availability with higher conformability. This is a successful demonstration of the benefit of using schedule.

## 5.4 The Effect of Write Requests

From Figure 16, all algorithms incur higher access cost when the write ratio is higher. The impact is especially evident on static replication. When the write ratio approaches 100%, the average access cost of static replication becomes very high since all replicas incur only consistency maintenance overhead. Dynamic replication schemes, however, are capable of discarding replicas that are no longer beneficial. When the write ratio approaches 0% (i.e. close to read-only access pattern), the cost calculation of three dynamic replication schemes almost always results in positive decision and thus a nearly full replication situation. In such case, most read requests can be satisfied locally, resulting in very low access cost. Local availability is also affected significantly (Figure 17). The higher the ratio

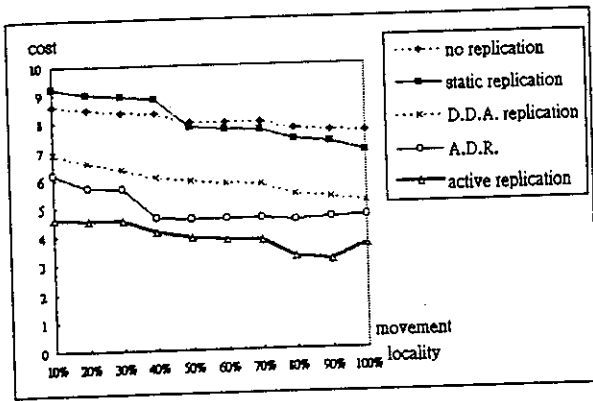


Figure 10. Average access cost of five replication schemes with varying movement locality.

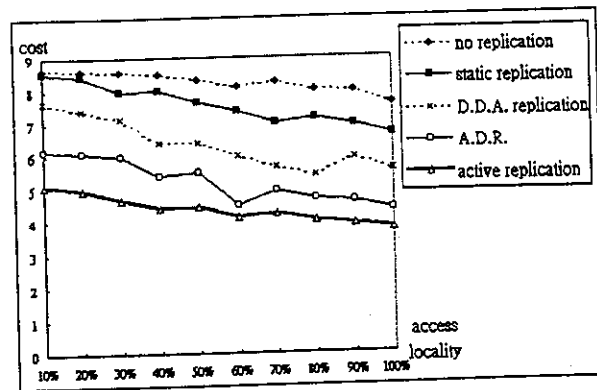


Figure 12. Average access cost of five replication schemes with varying access locality.

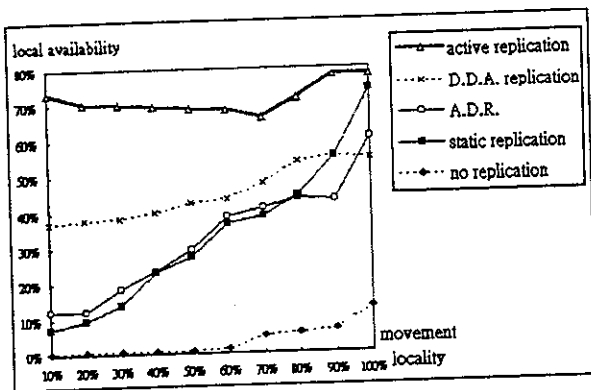


Figure 11. Average local availability of five replication schemes with varying movement locality.

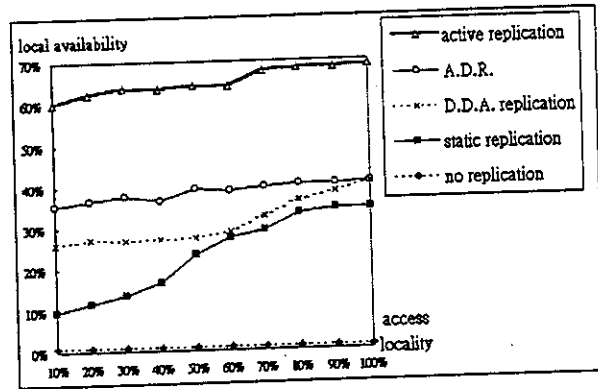


Figure 13. Average local availability of five replication schemes with varying access locality.

the lower the availability since the reduction of replicas increases the percentage of remote access.

## 6. Conclusions and Future Work

Traditional replication schemes are passive in nature, and rarely consider the characteristics of mobile environment. [We have proposed a dynamic replication scheme that actively provides replication services for mobile users.] With more precise and responsive cost model, as well as the help of schedule, [our scheme successfully demonstrates its ability to reduce access cost, improve response time, and achieve high local availability.] The implementation of our scheme calls for careful consideration of several issues such as the maintenance of access histories and user profiles, the selection of primary copy, the limits on the number of replicas, clock synchronization for periodic time check events, and the possible adoption of other cost models. (The perfor-

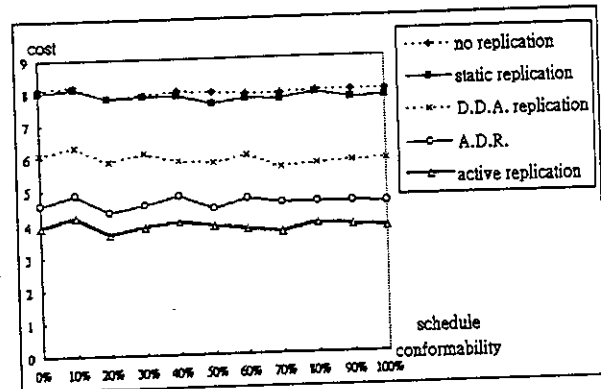


Figure 14. Average access cost of five replication schemes with varying schedule conformability.

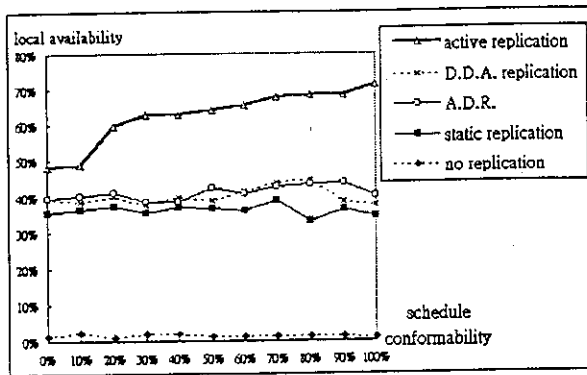


Figure 15. Average local availability of five replication schemes with varying schedule conformability.

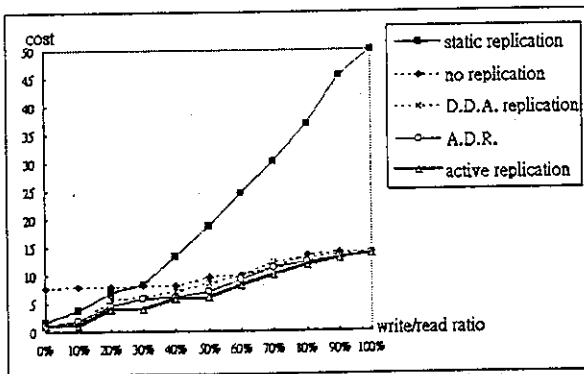


Figure 16. Average access cost of five replication schemes with varying write ratio.

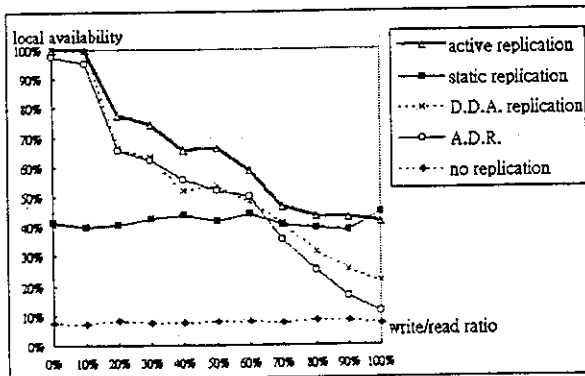


Figure 17. Average local availability of five replication schemes with varying write ratio.

mance of our algorithm can be further improved if the replication scheme is integrated with other mechanisms such as caching, prefetching, and data broadcasting.

## References

- [1] S. Acharya and S. B. Zdonik. An efficient scheme for dynamic data replication. Technical Report CS-93-43, Brown Univ CS Department, 1993.
- [2] B. R. Badrinath and T. Imielinski. Replication and mobility. In *Proc. 2nd Workshop on the Management of Replicated Data*, pages 9-12, Monterey, CA., 1992.
- [3] B. Ciciani, D. M. Dias, and P. S. Yu. Analysis of replication in distributed database systems. *IEEE Trans. on Knowledge and Data Engineering*, 2(2):247-261, 1990.
- [4] R. David, R. Peter, and P. Gerald. Replication requirements in mobile environments. Technical Report 970021, UCLA, 1997.
- [5] A. A. Helal, A. A. Heddaya, and B. B. Bhargava. *Replication Techniques in Distributed Systems*. Kluwer Academic Publishers, 1996. ISBN: 0-7923-9800-9.
- [6] Y. Huang, A. P. Sistla, and O. Wolfson. Data replication for mobile computers. In *SIGMOD Conference*, pages 13-24, 1994.
- [7] S.-Y. Hwang, K. K. S. Lee, and Y. H. Chin. Data replication in a distributed system: A performance study. In *DEXA*, pages 708-717, 1996.
- [8] M. C. Little and D. L. McCue. Computing replication placement in distributed system. In *Proc. IEEE 2nd Workshop on Replicated Data*, pages 58-61, Monterey, CA., 1992.
- [9] M. C. Little and S. K. Shrivastava. Using application specific knowledge for configuring object replicas. In *Proc. IEEE 3rd Intl. Conf. on Configurable Distributed Systems*, May 1996.
- [10] E. Pitoura. A replication schema to support weak connectivity in mobile information systems. In *DEXA*, pages 510-520, 1996.
- [11] N. Shivakumar, J. Jannink, and J. Widom. Per-user profile replication in mobile environment: Algorithms, analysis, and simulation results. *MONET*, 2(2):129-140, 1997.
- [12] J. Sidell, P. M. Aoki, A. Sab, C. Staelin, M. Stonebraker, and A. Yu. Data replication in Mariposa. Research Report S2K-95-62, Department of Electrical Engineering and Computer Science, University of California, Berkeley, 1995.
- [13] O. Wolfson and S. Jajodia. Distributed algorithms for dynamic replication of data. In *PODS*, pages 149-163, 1992.
- [14] O. Wolfson and S. Jajodia. An algorithm for dynamic data allocation in distributed systems. *Information Processing Letters*, 53(2):113-119, 1995.
- [15] O. Wolfson, S. Jajodia, and Y. Huang. An adaptive data replication algorithm. *ACM Trans on Database Systems*, 22(2):255-314, June 1997.