A Hypermedia Based System for Digital Information Archiving and Retrieval in A/E/C Project Management.

By

Kalpajit De

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Civil Engineering

APPROVED :

Dr. Michael C. Vorster, Chairman

Dr. Yvan J. Beliveau

Dr. Jesus M. de La Garza

Dr. James D. Lutz

December 1995

Blacksburg, Virginia

A HYPERMEDIA BASED SYSTEM FOR DIGITAL INFORMATION ARCHIVING AND RETRIEVAL IN A/E/C PROJECT MANAGEMENT.

by

Kalpajit De

Dr. Michael C. Vorster, Chairman

Via Department of Civil Engineering

(ABSTRACT)

The management of the high volume and different types of information that characterize its projects, has traditionally posed a constant challenge to the architecture, engineering and construction industry. The distribution and communication of such information to project participants, who are often geographically dispersed, is an important necessity and forms a significant part of this challenge.

Hypermedia computing on the Internet has emerged as an important and efficient means for the distribution and communication of information.

This thesis proposes hypermedia computing on the Internet as a solution to address the problem of information distribution in the architecture, engineering and construction industry. The problem of information management is examined, and relevant features of hypermedia and the Internet discussed. The viability of the technology to address the specific problem of information communication and distribution is demonstrated by building a hypermedia based information system as proof of concept. The thesis concludes with a discussion of future developments in this technology and the scope for further work in hypermedia based information distribution systems.

Acknowledgments

I wish to thank Dr. Mike Vorster for his input and advise in preparing this thesis, and most of all for his friendship, encouragement, and support during my graduate studies at Virginia Tech. I would also like to thank the members of my committee, Dr. Yvan Believeau, Dr. Chema de La Garza, and Dr. Jim Lutz for their reviewal and suggestions.

The contribution of the sample project data by Anderson & Associates, Inc., Blacksburg, Virginia, is gratefully acknowledged. In particular I would like to thank Sean Neff for his time and assistance.

To the gentle folk in the Deans Suite at Norris Hall, where I served as a graduate assistant for the past year and a half, thank you for your friendship. In particular, Linda Scott, Sandy Ritter, Faythe Rittenhouse and Linda Perkins, thank you for all the smiles and kind words, and making a place of work a very nice place to be.

I wish to thank my wife Paroma, for tirelessly proof reading this entire document, and for making each day as it comes, just a little more special and worth its while, like only she can.

I dedicate this thesis to Deepa and Kalyan De, parents and architects extraordinaire.

Table of Contents

List of Figures

1

Introduction

The latter decades of the twentieth century have been described as the information age (Dizard, 1982). With rapid growth in scientific development, knowledge has been expanding in many areas of human endeavor. The information explosion caused by this accelerated expansion in knowledge has brought with it a need for effective mechanisms that help manage the vast amounts of data that have become available. In his study, Porat (1977) empirically proved that 53% of the work force in the United States were engaged in "information needs". This number has expanded today.

The Architecture Engineering and Construction industry has traditionally been faced with its own set of needs for managing the large volumes of information that typically characterize its projects. The research presented in this thesis examines a specific problem in information management in this industry, and demonstrates the applicability of a chosen technology to address the same.

1.1 The Background

The introduction of computer technology has put immense data generation capabilities in the hands of individuals and organizations (Bates, 1989). Consequently, traditional paper based methods of data management, and modes of distributing information by physical transportation of printed paper are becoming displaced. In the technologically advanced society of the present merely upgrading such obsolete systems will be unable to address the needs of modern organizations and individuals, for high-speed and focused information delivery. Information systems of the present are required to perform two important tasks; (i) delivering a vast array of data in real time, and (ii) delivering such information in a format that facilitates its optimal utilization.

Architecture, Engineering and Construction forms one of the largest industry in the United States (Perkowski, 1988). It has been reported to account for 8% of the US GNP (US DOC, 1985). This industry has certain characteristics that distinguish its problems from its counterparts in the manufacturing sector. First, closely coupled with the characteristic of its considerable size, is the degree of fragmentation that exists in the construction industry(Choi and Ibbs, 1990). In fact, in comparison to its European and Asian counterparts, the US construction industry is far more fragmented. A census of US construction industry has revealed that there are over 1,400,000 establishments in the country, the vast majority of which have very few employees (Howard et al. 1989).

Fragmentation exists both at the individual phases of construction, as well as across the various project phases. There is a high degree of vertical and horizontal fragmentation. Vertical fragmentation is found between project phases - such as design and construction and horizontal fragmentation between the participants at a given phase - such as the designers of the electrical and mechanical subsystems (figure 1.1). This creates significant problems in communication and in the efficient management of a given project.

As conceptually illustrated in figure 1.1 the fragmentation is both vertical, between players in different phases of the project as well as horizontal, between players collaborating at a given phase. Often this severely impedes the distribution and flow of essential information, resulting in poor communication between the main players.

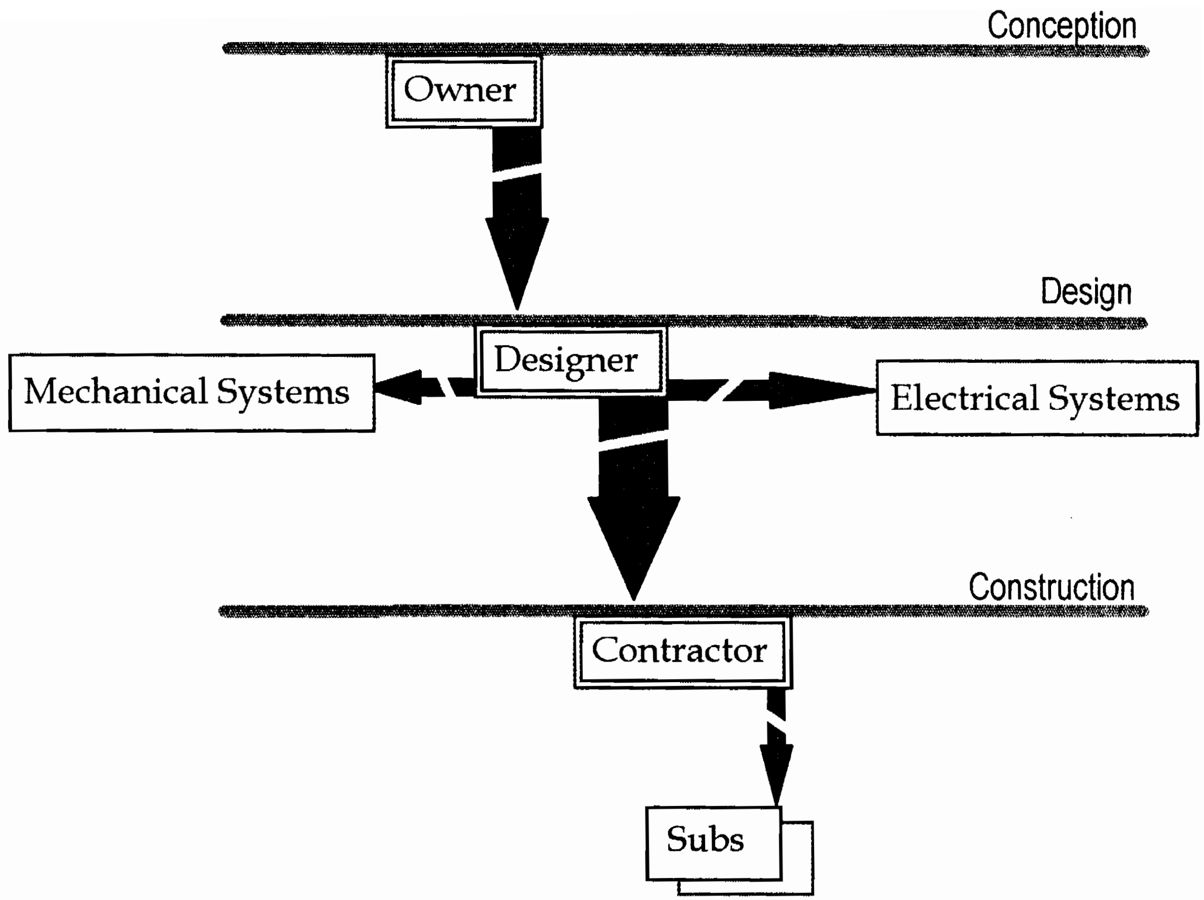

Figure 1.1 Horizontal and vertical fragmentation

Added to this, is the task of handling a plethora of project related data in numerical, graphical and textual format that characterizes the typical construction project. Conventional methods of data management have been insufficient in providing information which can be optimally used. Essential information has become concealed in a mass of accumulated drawings and documentation. The essence of this problem is well summarized by the frequently heard phrase, "data rich but information poor".

These problems have in turn, led to a high degree of fragmented decision making throughout the life of the project (Choi and Ibbs, 1990). Project participants, namely, the owner, designer and the contractor are often unable to visualize the project scenario in its entirety. They are either overwhelmed with accumulated technical and administrative information, or rendered myopic because too little information is made available to them. In the working environment of the construction industry, correctly made spot decisions are crucial. Such spot decisions can mean the difference between acceleration and delay, completion and rework, profit and loss, and indeed success and failure. Therefore, decisions taken on the basis of incomplete or outdated data, can lead to spiraling project costs at the least, and ultimately even to serious courtroom litigation.

This backdrop raises two issues which are of prime importance for success on the typical project. The first of these is, that there must be an effective vehicle for the conveyance and distribution of project related information. The second is that the quality of information conveyed should be in a format that facilitates its distribution, and remains appreciably simple to understand.

This thesis is concerned with the issue of information management in the A/E/C industry as discussed above. It should be mentioned at the outset that information management encompasses a large number of related issues. This includes defining what constitutes "information" in the A/E/C environment, and how that superset maybe broken down and categorized into a number of subsets. Furthermore, it is also essential at this point to understand that the methods of dealing with the different categories of information are different. Such differences depend largely on the nature and format of the data that is to be dealt with in that category.

In this thesis the research effort will focus on the particular problem of *distribution* and communication of certain types of A/E/C project related information, over a wide geographic area and among different participants in the project. The specific types of information taken into consideration for this purpose are discussed in some detail in chapter 2 of this document. The following sections state the objectives of this thesis, and discuss the scope and limitations of the work undertaken in this thesis.

1.2 Objectives

This thesis concentrates on the issue of distributing relevant information on an A/E/C project, to facilitate communication among concerned participants. The issue of distribution is viewed as a specific aspect of information management.

The objective of the thesis briefly stated, is to propose and demonstrate the utility of an upcoming technology in multimedia computing available on the Internet, in mitigating the problem of information *distribution* in the A/E/C industry, especially over a wide geographic area. More specifically, the thesis will focus on establishing the applicability of hypermedia computing on the Internet, as a possible technological solution to the problems discussed in distributing project related information.

A hypermedia based information system will be developed and discussed to demonstrate the advantageous features of hypermedia computing on the Internet and how it may be applied towards the problem of information distribution. The hypermedia software that will be utilized is available via the Internet. The objective also includes identifying features of the Internet that can be serve to enhance the applicability of this technology to the problem being dealt with.

1.3 Scope and Limitations

The research effort documented here seeks to investigate problems associated with the distribution of information as discussed above. The aim of the thesis is to contribute towards the understanding of how hypermedia computing on the Internet can be used to address these problems, and establish the viability of this technology by building a proof of concept that demonstrates the features discussed.

To begin with, an attempt is made to define the terms data and information, and addressing the question regarding which types of data constitute information in section 1.5 and section 1.6. The problem of information management is then outlined, and a discussion provided on the importance of the distribution of information in an A/E/C environment. The thesis then proceeds to identify a technology that may serve to alleviate the problem through the use of a hypermedia based computing tool. The technology that is the subject of the research effort is hypermedia computing on the Internet, and more specifically, a hypermedia based storage and retrieval tool with a graphical user interface to the Internet, its means of data interchange, and its multimedia capabilities relevant to the above problem. An attempt is made to identify the scope and applicability of the technology to the problem, on the basis of the existing features of the chosen software. A brief discussion is provided on the advantages proffered by the multimedia features of this technology.

In order to test arguments presented in favor of the chosen technology, a hypermedia module has been authored based primarily on NCSA Mosaic, to demonstrate its feasibility and usefulness in an A/E/C environment. Other Internet tools providing specific multimedia support to Mosaic have also been considered for this purpose. The module developed here is meant to serve as a proof of concept for the thesis, by demonstrating how project information can be archived and retrieved over a wide geographic area. For this purpose, information sampled from an actual A/E/C project has been used to illustrate the functionality of the module

The research avoids the direct use of any commercially available software package, but does include discussion on the means to link the distributed information with some such selected packages. The scope of research is limited to applying the capabilities of the currently available versions of hypermedia browsers such as NCSA Mosaic, to the problem mentioned.

Where specifically mentioned, the NCSA Mosaic browser refers to the shareware version available by anonymous file transfer protocol over the Internet. The term hypermedia will refer primarily to the hypermedia features of the NCSA Mosaic browser and World Wide Web protocols. It maybe noted here that there are at present commercially available versions of Mosaic, or very similar hypermedia software that are possibly more robust, and offer additional features. These have not been covered in this thesis because this research sought to investigate the viability of computing resources that are widely available, universally popular, and thereby the unofficial industry standard. The issue of security relating to communication of information over the Internet, and restricting access to information via hypermedia browsers is discussed on a theoretical basis at the end of chapter 4. The discussion there attempts to provide an overview of relevant security concerns, and how these may be addressed.

The concluding part of the thesis outlines the future directions of the Internet and hypermedia computing on the Internet, and the implications thereof. A brief discussion is provided on possibilities for further research in this area.

1.4 Outline

The presentation of the research for this thesis has been structured into five sequential parts as shown in the flow chart in figure 1.2. As illustrated in the figure, the approach taken has been to identify a specific problem, discuss a technology that may be applied towards addressing the problem, present a solution in the form of a proof of concept, and conclude with a discussion on the future directions of the technology and scope for further research.

The opening sections of chapter 1 have provided a background for the topic of this thesis, stated the objectives of this research, and outlined the scope and limitations of the thesis. The forthcoming sections of chapter 1 present a discussion of what constitutes "information" in an A/E/C project, and identify and define the problem of information management in the industry.

Chapter 2 presents a discussion on information systems in the A/E/C industry. This includes a brief mention of practices and procedures that are commonly found in A/E/C firms that do not extensively utilize high end technology. The chapter attempts to outline the general considerations information systems development for the A/E/C industry. A discussion is then presented on the categorization of A/E/C data and information, and the need for distribution. The chapter closes with a note on multimedia and its impact as a tool for this industry.

The primary purpose behind the content in Chapter 2 is to gain further understanding of information management in the A/E/C industry, and thereby obtain an overview of which aspects of the problem a hypermedia based tool may address.

Chapter 3 will introduce the technology that is proposed here as a remedy to the problem. To start with, a discussion on the Internet, its workings, and relevant issues are mentioned. The concept of hypermedia including relevant definitions and related technical aspects are then outlined.

While the objective in Chapter 3 is to introduce the technology, Chapter 4 attempts to explain how this technology may be applied to the specific contexts as outlined in the first chapter. The features of the hypermedia browser that allow for the storage and retrieval of information, and how this may facilitate communication among project participants is explained here. This is achieved by presenting HARM (Hypermedia Archiving and Retrieval Module), the proof of concept of this thesis, which is a hypermedia module authored based on NCSA Mosaic. The objective of this module is to demonstrate some of the features of hypermedia in the context of facilitating information management in an A/E/C project.

Having defined the problem being dealt with, discussed a technology that can be applied towards mitigating the problem, and provided a demonstration of the same via the proof of concept, Chapter 5 will conclude the thesis with an outline on the future directions of the Internet, and scope for future work relevant to this thesis.

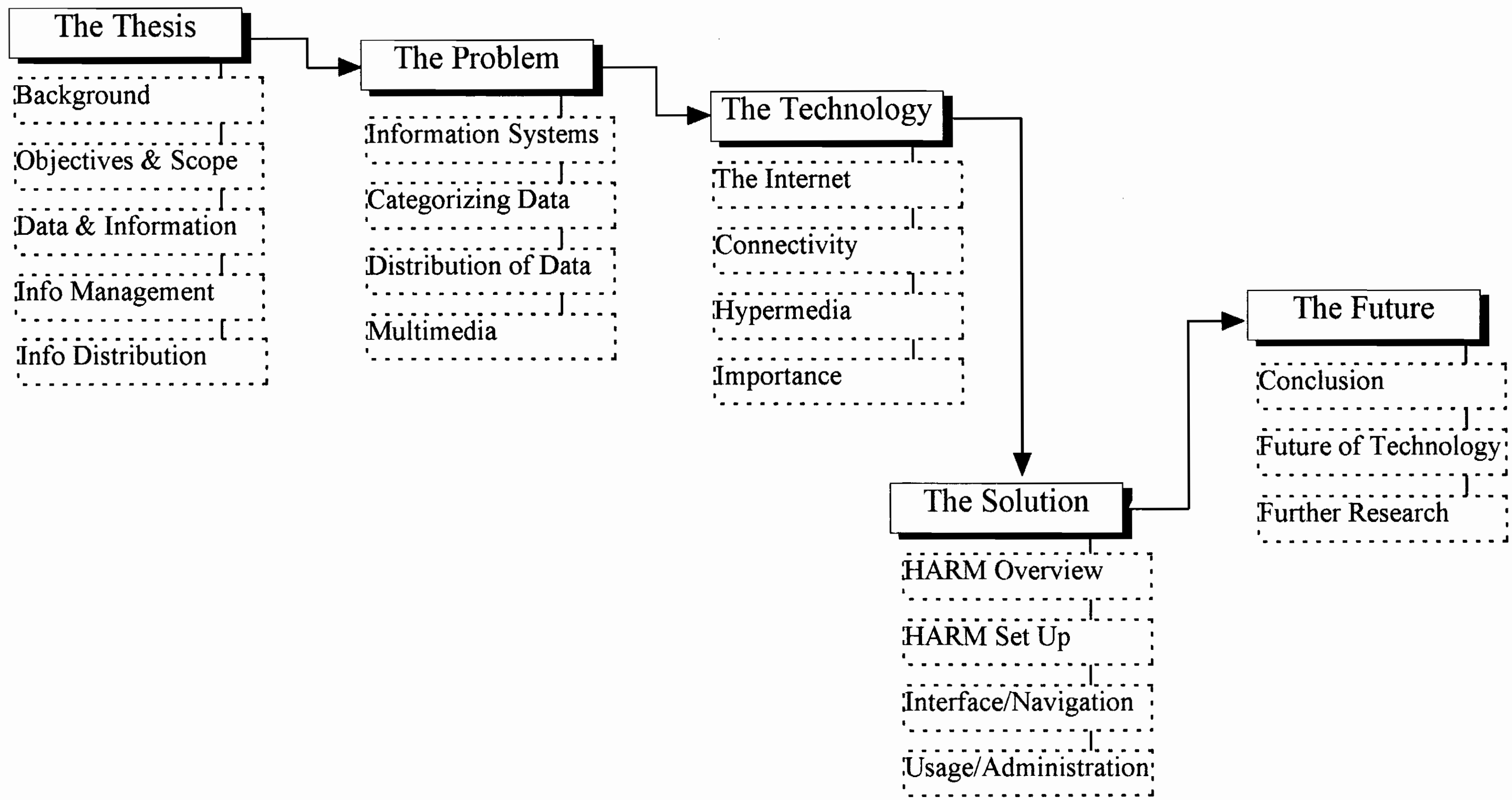

Figure 1.2 The structure of the thesis

1.5 Data and Information

The distribution of information being the crux of this thesis, the terms information and data are frequently used. In order to maintain a focus on the context in which these terms will be used, it is necessary to develop a clearer understanding of what these terms mean, and highlight the distinction between the two.

Information is a generic term that encompasses a many substantive meanings having to do with education, news, facts, computerized databases, empirical statistics and the like that serve to provide certain knowledge on a given issue (Chaffee, 1991). While knowledge can thus, be considered as a byproduct of information, the question then arises from where and how does information originate. The answer to this question is that information originates from data. Individual elements of data can be looked upon as facts, and collection of such facts on a given issue form a data set. In isolation, a single element of data is usually of limited or negligible significance. It is a combination of facts in the form of one or more data sets that provide the basis for obtaining information.

This idea introduces the distinction between data and information, for while data represents facts, information is the meaning that people assign to these facts (Davis and McCormack, 1979). Data thus needs to be routed through some process that combines the facts it represents in some meaningful way in order for it to make the transition to the state where it can be used as information.

In many instances, the "process" referred to above, simply involves sorting or reorganization of the available facts in a format that provides new meaning or greater understanding to the combined data set. Once the data has been put through the processing stage, it is now available for use as information. Typically such information is utilized for some form of decision making, which in turn may precipitate some activity. The execution of the activity, would in turn, give rise to fresh data which might or might not be passed through the above mentioned stages again. These events can be pictured as an information cycle as shown in figure 1.3 below.

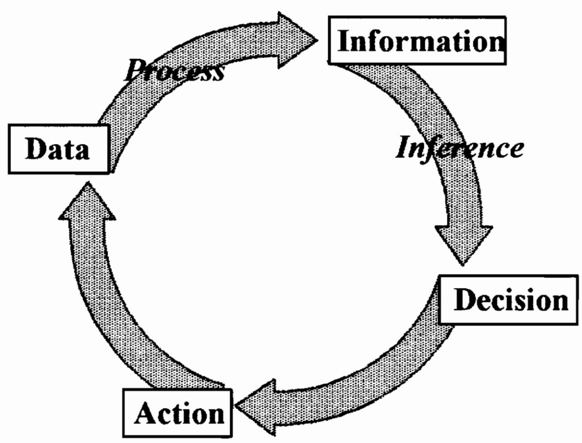

Figure 1.3 The information feedback cycle[1]

The process of data being transformed to information is depicted by the hatched arrow between the two phases. The remaining hatched arrows represent other processes that might be involved between the respective phases.

---

[1]Modified from *The Information Age*, Davis and McCormack, Addison-Wesley, 1979, p24.

Having provided an idea of how data and information can be distinctly conceptualized in the context of the thesis, it is possible now to revisit the two terms and attempt to define the latter. As defined by the Webster's[2] dictionary information is the "communication of news, knowledge (or) facts." To extend this definition for the current context, information can be said to consist of data that has been retrieved and put to use for informative or inference process, argument, or as a basis for forecasting or decision making (Murdick et al., 1984).

The distinction here, between information and data, should be kept in mind. While a large amount of data maybe available on a given issue, all of it may not be put to use. Thus available data becomes information only when processed or put to use to aid in inference, argument, decision making or some like purpose. Simply put, information is a subset of data, and while all information consists of some data, all data need not necessarily constitute information. To summarize the essence of the above discussion, data represents facts while information is the meaning that is assigned to those facts after they have been processed in a desired manner. The information obtained thus becomes an aid in decision making.

---

[2] *Webster's Dictionary*, NTH Edition, 1994.

1.6 Data and Information in the A/E/C Industry

The previous section attempted to provide a conceptual understanding of what the terms information and data encompass in a generic sense. In the context of this thesis the definition provided above could be somewhat extended to reflect the needs of the working environment of the A/E/C industry.

Information comprises of news, knowledge or facts that aid in decision making. This is achieved by giving participants in a process a comprehensive understanding of the global picture of which that process is a part. Naturally, the participants in this context, refer to at the least to the three major participants involved in a typical A/E/C project -- the owner, designer, and contractor.

It is now essential to take a closer look at the typical kinds of data involving A/E/C projects, and how they contribute to the body of useful information. In the environment of the construction industry, the typical project is characterized by a vast amount of data that tends to increase or build up over the life cycle of the project (Hemmet, 1993). The data itself exists in diverse forms, and the different forms vary widely in their purpose, functional use, and even amount of use.

For instance, both a section of the contract document detailing procedures for processing change order work, and a set of still photographs taken on site that show certain constructability constraints, qualify as important data. Both sets of these data may require retrieval for reference analysis and decision making. However, while the former remains unchanged over the duration of the project, the latter is most likely to change as the project progresses and constructability constraints change. Similarly project drawings detailing the design of a particular component of the facility, and documents that detail the specifications of materials for that component both qualify as forms of data on the project.

Data on a construction project could be in textual, graphical, pictorial, video and even audio formats (Naiem et al. 1993). The actual format or medium of the data will, as mentioned before, depend largely on its purpose and function. It is also worth noting that much of the data generated may be limited to single use only. Referring back to the above instance, certain design drawings maybe rendered obsolete due to subsequent change orders, still pictures of a partially constructed component maybe outdated after construction of that component has been completed, and material specifications may differ from one component to another. From the above discussion three significant characteristics of data related to an A/E/C project emerge.

- The different functions on a project such as contract administration and design alterations generally require different kinds of data.

- The data required for a given function or a certain format maybe of a static or dynamic nature. In other words, it may remain unchanged over the duration of the project or may require updating or replacing as the project proceeds.

- The data can exist in different formats such as text, graphics, pictures etc.

As pointed out earlier, not all data qualifies as information. It is only when different available elements or sets of data are combined or cross referenced in a meaningful manner to contribute to decision making that they form information. In the current context information will consist of any data that is put to use by any of the three major participants in the project -- the owner, the designer and the contractor, to perform their expected roles over the life-cycle of the project. More importantly, it will also comprise of any data that can be put to use to aid the process of communication between these major participants, so as to provide them with a clear understanding of the progress being made, problems being encountered, and the status of the project at a given time.

1.7 The Problem of Information Management

The introduction to the thesis provided a perspective on information management and cited it as being a crucial issue for industries and organizations wishing to preserve a competitive edge in contemporary times. This thesis takes the approach of viewing the management of information as a challenge. Before addressing specific issues it would be helpful to summarize the key facets of the problem in this section.

This thesis is focused on the problem of distribution and communication of information. As the introduction specifies, distribution and communication represent just one of the many issues that comprise the problem of information management. Therefore, the discussion that follows, rather than being all inclusive attempts to identify and provide an overview of the current status of the problem of distribution and communication of information in the A/E/C industry. It may be kept in mind that while this discussion deals with issues that relate to the A/E/C industry, many of the same problems are common to other industrial sectors as well.

There are two primary issues to information management:

(i) organizational

(ii) technological.

The organization of the A/E/C industry is characterized by certain features that distinguish it from other industries. The effective management of information is obviously crucial to any given industry; but, in the A/E/C environment, the leading aspects of this problem arise from the unique way in which this industry is structured and operated. Hence, in order to put this problem in proper perspective a brief outline of these features would be helpful.

A construction project involves several participants who interact with one another over its life cycle. The major participants are:

(a) the owner or client desiring to have a given facility constructed,

(b) the designer of the facility responsible for putting together the design in accordance with the client's needs or specifications, and

(c) the contractor responsible for producing the physical construction product.

In addition to these three basic participants there maybe one or more financiers providing funding for the project, and almost invariably several subcontractors performing specialized work under the supervision of the main contractor. The successful completion of a project on schedule and within budget demands coordinating the roles and responsibilities of these participants. In particular, to ensure that the end product meets desired specifications the owner has an essentially active role to play throughout the life cycle of the project.

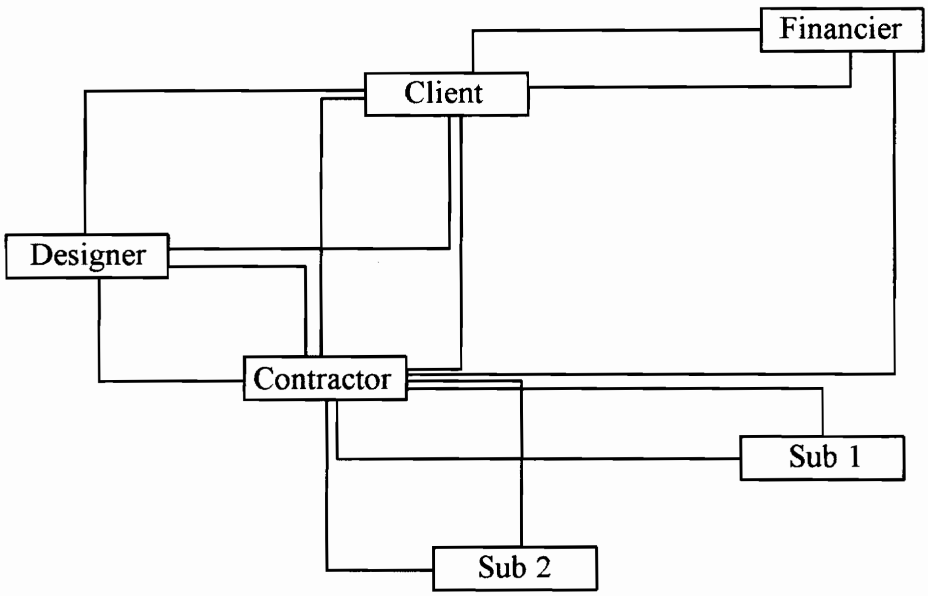

Figure 1.4 Multiple lines of communication

Figure 1.4 depicts the essential lines of communication that need to be established between the participants mentioned above. While the figure shows the basic scenario, the actual number of participants will depend on the size and complexity of the project. Each participant needs to maintain lines of communication to and from each of the other participants it interacts with. Thus with two participants in the picture, such as buyer and seller in the retailing industry for instance, only two main lines of communication would be required. In contrast the lines of communication required in the scenario shown in figure 1.4 are at least 13 in number. An increasing number of participants, therefore, have the impact of increasing the organizational complexity of the project. This in turn, increases the problem of effective information flow between the participants.

It must be remembered that A/E/C firms are essentially service businesses. Although the performance of their services results in tangible products such as buildings, commercial engineering, or industrial structures, the principal output is a service geared toward meeting the specified needs of the client (Bessom, 1975). Each project is thus a customized product and no one project is quite like the next. Even for a single A/E/C firm, be it involved in design or actual construction, projects are likely to differ considerably in size, nature, and scope.

Requirements in terms of performance and specifications also change depending on the nature of the client or the owner of the project. For instance in awarding a new contract for the design and construction of a desired facility, a government agency would invite competitive bids whereas a private owner seeking construction services may opt for a negotiated contract that stipulates entirely different conditions (McCaffer and Baldwin, 1984).

These differences deprive the A/E/C industry of the advantage of being able to reuse a large part of the data of a given project in the next project. This also necessitates the establishment of fresh lines of communication for each new project, further increasing the difficulty of developing a fixed pattern or methodology of information exchange. Further aggravating this problem is traditional fragmentation, a long standing characteristic of this industry.

The other major issue involved in information management as stated before is technological, having to do with computerization of the industry. Introduction of computer technology in the industry has created numerous means to accelerate individual tasks in the construction process. Computer technology has afforded construction and design firms the means to automate time intensive tasks such as structural analysis and design, production of design and shop drawings, estimating and scheduling, and data storage and manipulation (Parfitt, 1993; Rocha and Morad, 1993).

The benefits that have been realized in a number of areas, such as savings in time and reduction in design errors, have been nothing short of dramatic. However, widespread computerization has also introduced new problems that are far from being insignificant. The ability to perform time consuming, data intensive activities with great speed and efficiency using computers has contributed to a significant increase in the volume of data generated over the life of the project. Computerization has also ushered in a new genre of clients who are better acquainted with professional procedures, have greater demands in terms of performance expectations, and expect to be kept well informed of progress in their project (Perkowski, 1988). Naturally this creates a greater need for A/E/C firms to be able to distribute relevant information in an effective manner.

Computerization has also, so far, tended to reinforce rather than alleviate organizational fragmentation. The explanation of this seeming paradox is rooted in two basic and interrelated causes, namely, (i)system incompatibility, and (ii)a lack of industry-wide standards (Choi and Ibbs, 1990). As a result even construction firms that have implemented similar level of computerization in their work processes may have done so on different hardware platforms that support differing software between which data exchange is either difficult or not feasible. It is not uncommon to find several platforms such as PCs and UNIX workstations in the same firm for different applications, hindering internal data exchange within that firm. Even where a construction or design firm has alleviated the problem of internal data communication by means of standardized in-house software, the lack of industry-wide standards almost invariably hinders the direct exchange of such data outside the firm.

These problems, commonly grouped under the phrase computer fragmentation (Howard et al. 1989 ; Choi and Ibbs, 1990), are posing a challenge to A/E/C firms striving to remain competitive at the turn of this century. For several years now the growing need for integration among computer systems has been recognized (Perkowski 1988). Consequently, ways and means of achieving an environment of computer integrated construction has been the subject of considerable research in recent times. To date such research efforts (Howard et al. 1989; Miyatake et al. 1993; Parfitt et. al. 1993, Teicholz et al. 1994) have offered a variety of excellent solutions that tie together isolated applications within a given firm. A comprehensive solution to the problem continues to be evasive, however, as communicating data outside the boundaries of a firm over a wide area and across different platforms, is still fraught with obstacles in light of the fact that industry wide standards and universal platforms are yet to evolve.

Research on developing market trends in the industry (Krippaehn et al. 1992, Yates 1991) indicate that US based A/E/C firms will increasingly face a *geographically dispersed* client base. In many cases such clients will be owners located in a foreign country possibly half way across the world. Such firms are therefore now being confronted with a increasing need to coordinate the distribution of information among various parties who may be located at a considerable distance. Commercial vendors who offer global networking services have provided only a partial solution since none have a proprietary hold on the industry. There is no assurance that all participating firms on a given project will be availing of the same network service provider. Connections via multiple network service providers is of course a possibility but this once again considerably restricts flexibility on the kind of data that can be usefully shared. In addition such services are still very expensive.

Hence, while the introduction of computers to the design process has changed and even improved the means of generating data, it has not fundamentally changed the methods of sharing information.

1.8 Distribution

Section 1.7 attempted to outline the problem of information management from the organizational and technological viewpoint, and identify where the issues of distribution and communication of information appear in this backdrop. In the context of this thesis it would be helpful to review what makes the issue of distribution one of prime importance in an A/E/C environment.

The timely conveyance of necessary information that keeps all concerned participants continuously informed of the overall scenario has been cited as a crucial factor for the success of a construction project (Sanvido et al. 1992). Effectively conveyed information facilitates the process of interaction between the parties involved and promotes a clear understanding between them on vital issues such as design intent, quality control, project duration and limitations of the contract. The need for this cannot be over-stressed, for the bane of any construction project is a conflict of interests surfacing between concerned parties. Most often, this situation arises due to a breakdown of communication that leads to a misinterpretation of their respective roles and responsibilities as defined by the contract. Organizing available data in a manner that enables its use as useful information, and aids the process of communication between the participants on a construction project has been a constant challenge for the A/E/C industry.

To recapitulate the discussion in section 1.5, data by itself does not provide information. Meaningfully combined or cross referenced data represents potentially useful information, but only if it is made available to the persons or parties in need of the same. Making information available to the right people in real time empowers them to make informed decisions (Palmer and Varnet, 1990). Indeed, by itself, this defines one of the end objectives in the management of information.

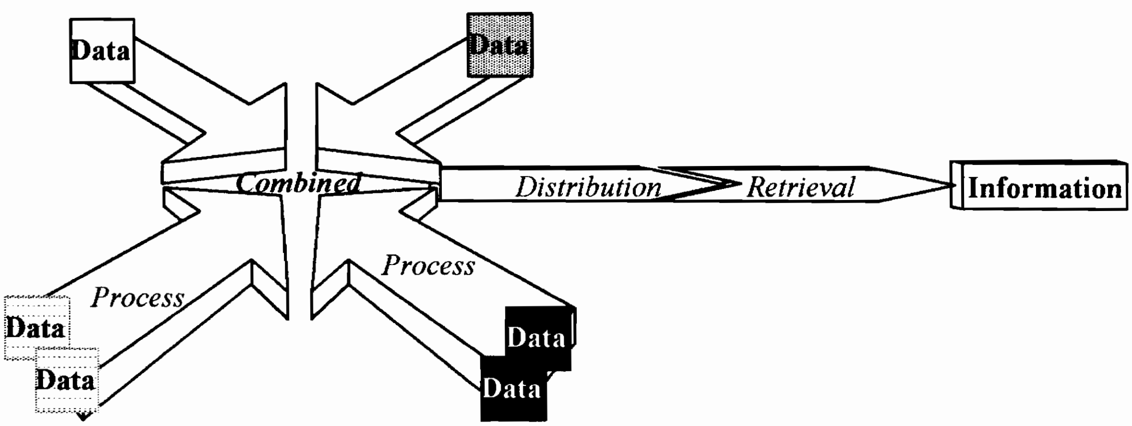

Figure 1.5 The distribution of information for decision making

Information management itself, as pointed out before, is a generic term that encompasses a number of related issues, each of which pose challenges of their own. Such issues include manipulation, processing, transformation, and distribution of information. Of these, the issue that emerges as relevant to the above discussion and continues to be a long standing problem in the A/E/C industry, is that concerning the *distribution* of project related information. The crux of the problem is to be able to make the project related data widely available in a format that is retrievable in real time for instantaneous reference and decision making. Indeed, only under such conditions does the 'data' hold the realistic possibility of transcending to the level of 'information' as defined earlier in this section 1.5. Figure 1.5 attempts to illustrate the process conceptually.

This chapter of the thesis has provided the background for the discussion that follows. With an understanding of what constitutes information in the A/E/C industry, it is now essential to devise a means to categorize the same, and identify which types of information will most frequently require distribution among project participants. In developing a tool for the distribution of such information, it would also be useful to obtain an overview of information systems development in the A/E/C industry, and the impact that multimedia technology has had on the same. These concerns are addressed in the following chapter.

2

Information Systems in the A/E/C Industry

Chapter 2 of the thesis introduces and develops a discussion on some of the main issues that need to be taken into consideration in the development of information systems in general. In particular, a focus is laid on issues that are relevant to information systems that will serve the A/E/C industry.

The discussion on information systems development attempts to bring out two important issues:

(a) being able to categorize A/E/C data and information, and,

(b) identify what data and information are relevant for the problem of distribution being dealt with here.

The chapter concludes with a brief note on the impact of multimedia technology on information systems, and its possible applicability as a tool for the A/E/C industry.

2.1 Information Systems Development

In all kinds of organizations information and data constantly flow from one point to another. Such information flow is between individuals and groups both within the organization and outside it. The definition and distinction of the terms information and data as outlined in chapter 1 make it evident that the goal of any information system must necessarily be making available the right information at the right time to end users in need of that information. The following sections provide an overall idea of what different information systems represent and seek to achieve. In the context of this thesis some of the major issues that need to be taken into account in developing an information system are also discussed.

2.1.1 Systems and Technology

An information system can be looked upon as any established procedure, mechanism or technology for the transfer of information or data from one point to another (Baxendale, 1994). Individuals and groups involved at different levels of decision making require systems that help them assimilate and reorganize disparate arrays of data from various sources to obtain meaningful information. Information systems in organizations may range from being rudimentary to sophisticated but all seek to meet the overall goal of assimilation and reorganization of data into information.

Vital information is recorded either on paper or in a digitized form at some juncture over the course of the point to point transfer process. Traditionally, information management systems were geared towards only the recording and representation of information. The concept of information itself as an element of value was absent. Most often such systems were carriers of information rather than being information based (Palmer and Varnet, 1990). Traditional carrier-based information systems that have evolved include libraries, records management, microform, and paper based archiving.

The advent of digital and computer technology has radically altered this scenario. Computer technology has given rise to more advanced systems for office automation, digital processing and data archiving in computer data bases. Such systems have popularized the term information technology, which refers to the processing, manipulation and transfer of data in an electronic or digital format. The design of some of these advanced systems have also continued within the carrier based paradigm, each system developing a particular pattern that fits the carrier. Naturally, with respect to achieving the overall goal of conveying the right information to the right person at the right time, each system has addressed the issue within the limitations of their respective information carriers. Some further discussion on this is provided in the next section.

Digital technology has increased the scope for the manipulation and processing of data for specialized purposes even before such data is transmitted as information. As pointed out in section 1.4 of Chapter 1, there are certain organizational and technological issues that should be taken in to consideration in developing and information system for the A/E/C industry. In order to maintain clarity a few of these issues are briefly described below (Gardner, 1993; Hemmet, 1993; Baxendale, 1994)

- *Storage* — Can be regarded as the archiving of large volumes of data in digital format in a manner that facilitates easy retrieval of all or part of the data for subsequent use. For instance, maintaining a complete record of inventory, purchase orders and the like with the aid of a database application program.

- *Manipulation* — Can be regarded as subjecting data to certain alterations or iterations to produce a desired set of values. An example would be iterative procedures carried out in the dimensioning of a structural member subject that must be designed to withstand certain loading conditions.

- *Processing* — Can be regarded as piping data retrieved from an existing database through certain filters to produce a leaner focused data set in a desired format, that is, data specific to some given purpose. An example would be generating a report of all costs associated with a particular portion of the project in a specified time period.

- *Transformation* — Can be regarded as utilizing a set of raw data to generate a completely new data set that can provide essential information. For instance data relating to costs, quantities and calendar dates and deadlines being utilized by a scheduling program to produce a construction schedule.

- *Distribution* Can be regarded as storage, retrieval and transfer of selected data that in most cases has passed through one or more of the above stages, to the various participants involved in the project. An example would be a designer's office providing a contractor with data such as drawings and specifications relating to a change order required by the owner.

It should be noted, however, that the issues discussed above do not encompass all aspects but constitute only the primary features involved in information systems. The descriptions provided here are not meant to serve as definitions but rather to illustrate the distinction between these issues, and clarify which of these is relevant to this thesis.

Most information systems combine storage with one or more of these above mentioned functions to manage information in accordance with the needs of the organization or environment it is designed for. Information technology thus (a) seeks to replace time consuming error prone manual operations with computerized versions of the same, and (b) paper based transactions with digital communications between computers and their application packages. In doing so an information system may address one or more of the following areas or functions (Somogyi, 1993) ;

- The system may partially or fully automate administrative type business activities which are prone to error when performed manually.

- The system may provide the facility of organizing, storing and retrieving a large quantity of data far more effectively than could be achieved using manual methods.

- The system may enable the real time flow of information over geographically dispersed locations

- The system may support business processes with data that is of value to the satisfactory functioning of the business by providing access to the different sources of data.

The component technologies that comprise a particular information system may vary widely from those that make up another. For instance, a system may be based on software developed in-house, or may consist of a number of vendor supplied application programs that have been interfaced or integrated together in some way, each of which perform a specific task. The flexibility of the system naturally depends upon the number of functions that it can handle. As such, this is influenced to a large extent by the expense that is budgeted for the system. It should be kept in mind that provided a given system is competent in addressing the specific needs of information handling it was designed for, it has been successful in fulfilling its prime objective (Turk, 1992). While added functionality in such cases would be a bonus, the level of sophistication of the technology employed in an information system and the number of above functions it can perform, is not necessarily a measure of its effectiveness.

2.1.2 Limitations of Information Systems

Regardless of the type of data being dealt with or the level of sophistication of the technology involved, almost all information systems come with certain inescapable limitations. Among these, is the fact that, information needs of individuals or groups serving different functions in an organization vary considerably (Lawler and Rhode, 1976). Problem solving groups may require precise quantitative data, consulting groups may call for the directive type of information such as standards, guidelines, laws and regulations, while individuals in decision making roles may need a broader array of synthesized and aggregated information (Palmer and Varnet, 1990).

Traditional methods commonly cause inevitable delays in conveying information from one point to another. Often such methods result in unnecessary duplication of data and efforts. Materials tend to get misplaced or misdirected resulting in the failure of the system to meet its stated goal. Apart from the inherent difficulties of tracking of information in transit, is the fact that those producing the information are often not fully aware of those who may need it. Consequently the required connections to the right destinations are error prone.

Compounding these problems is of course the inability of conventional methods to meet the often vital requirement of transmitting the data in real time. In many cases the value or utility of the data that could have served as information is thereby diminished or lost, as by the time it reaches its intended destination, a more current version of the same is available. In an A/E/C environment progress reports, correspondence relating to a construction or design schedule serves as an example of data that must be transmitted in real time to concerned parties to be of significant value.

Advances in information technology during the past decade have offered a variety of solutions to mitigate some of the above stated problems. The terms listed in the previous section provided an idea of some of the features that are found in advanced information systems. These are essential to comprehensively meet the general goals of an information system. An existing traditional system would be unlikely to be able to incorporate all of these features. The main reason behind this is that most computer systems are designed to address a specific aspect of the problem of managing the data, and additional features are usually added only as and when extended needs develop (Howard et al. 1989).

In some instances application packages serve a specific function well but do not interface easily with other applications producing related information. For example, a scheduling software package that produces an accurate construction schedule with supporting charts and graphics to aid in understanding the transformed data, may not lend itself to being interfaced with a CAD package that was used to create the design drawings for the project.

A significant part of this problem also concerns incompatible computing platforms such as UNIX based workstations and DOS based PC's. Typically, an application program that runs on a given platform and its respective outputted data, would not be compatible or accessible from a different platform. Thus, there is a tendency for different forms of information to get dispersed in incompatible carriers. This further increases the difficulty of bringing together the differing types to produce a complete set of integrated information. In other words most systems come prepackaged with certain features built in, rather than being tailored to specific individual needs.

2.1.3 Organizational Issues

In trying to meet the needs for the distribution of information in an organization, it is important to take into account its organizational or administrative structure. An overview of the organizational structure provides insight into the manner in which the lines of communication between various groups and personnel within the organization are maintained. For the given organization it is then possible to decide whether the information system will seek to enhance the existing lines of communication or open alternative routes for the same. Of the number of such administrative structures utilized in various organizations, the following are relevant in the context of the A/E/C industry (Palmer and Varnet, 1990).

- *Team Structure* — Commonly comprises of group of specialists and managers who interact vigorously, and share joint responsibility for running the system.

- *Democratic Structure* — Commonly comprises a central management team at the center of a group of specialists, who in turn have a group of operating personnel on their periphery.

- *Free-Form Structure* — This is characterized by a largely flexible interaction between management and operating personnel, to facilitate adaptability to changing circumstances.

- *Multiform Structure* — As implied by the name, this form of an organizational set up is usually a mixture of the above mentioned forms. Commonly a small controlling center may have lines of communication and control that coordinate the interactions between two or more types of structures.

From the discussion on the set up in the A/E/C environment which involves several major interacting participants, it is evident that the multiform structure presents the best fit. This further reinforces the need for multiple lines of communication to be maintained between participant entities on a given project as discussed earlier in Chapter 1 section 1.3 and 1.4, and serves to highlight the importance of being able to distribute relevant data among them.

A second organizational perspective that is worth keeping in mind is that an information system basically seeks to influence the behavior of its end users in a positive fashion. An information system influences behavior by assisting, guiding and thereby empowering individuals to make informed decisions that are consistent with the overall goals of the organization (Lawler and Rhode, 1976).

In the context of a typical A/E/C project a wide variety of individuals are called upon to make decisions at both high and low levels daily. Ensuring the distribution of relevant information to these individuals on time fashion must, therefore, be essential in the development of an information system that seeks to impact their behavior and consequently the given project in a positive manner.

2.1.4 Strategic Issues

The strategic issues that influence the development of an information system depend largely on the particular aspect of information management that the system will primarily be required to facilitate. There are several areas in which a system with a focus on the aspect of distribution should be able to provide a strategic advantage (Somogyi, 1993). The following are a summary of those that are relevant to an A/E/C firm.

- *Administration* The system should be able to provide readily accessible information that aids in the assessment of past performance, monitors current activities and plans for future periods.

- *Communication* The system should facilitate regular conveyance of data with the outside world as well as in between the component groups that are involved in the project.

- *Coordination* The system should facilitate a meaningful connection between different activities on the project, and should provide assistance in achieving a connection between them. On an A/E/C project, the chain of activities tend to be data rich in the sense that successive activities require or generate information. The system should provide the ability to carry information parallel with the chain of activities.

- *Services* The system should be able to enhance the quality of services provided by the firm that employs the system. This it should do by enabling the firm to deliver a product that is not only associated with a great deal of data, but is also information rich.

Apart from such specific areas, an information system should be able to link the information technology being employed with the business objectives of the organization that employs it. In other words, it should ensure that the technology supports the business and not the other way round (Stewart, 1993).

The "people issue" of employing new technology must also be given due consideration. As discussed earlier, data transcends to the level of information only when it is used for decision making. The technology should be such that encourages its use by individuals in the industry, for though the system must provide the data it is the end users who must perform the ultimate decision making function. Information systems that provide sophisticated features but are difficult to use fall short of expectations from a strategic viewpoint. The use of such systems become limited to a smaller group of individuals and organizations, rather than being adopted by the industry at large for the enhancement of business processes.

2.2 Categorizing A/E/C Data and Information

The discussion presented in Chapter 1, section 1.3, provided an overview of the diverse kinds of data that are generated on the typical A/E/C project. Naturally not all of such data are required to be distributed between the different parties involved in the project. There is a vast amount of data of different kinds that tend to accumulate. In order to be able to identify which of these require distribution to effectively serve as information, it is essential to try and categorize different kinds of A/E/C data. Data available on a given project could then be itemized and placed into its respective category. This would facilitate the orderly storage of different kinds of data and make the process of retrieval of a given type easier.

There are several ways in which available data could be subdivided and sorted depending on its source, purpose, format, function and whether computer processible or not. Indeed, a comprehensive classification of all kinds of data is a subject in itself and would require extensive treatment beyond the scope of this discussion. However, in the current context, to achieve the purpose outlined above, the data could be categorized taking into account the characteristics identified in the previous section. To begin with, such data could be grouped on the basis of function. On an A/E/C project the following are the distinct functional areas (Gardner, 1993).

- *Operational* Comprising work related to engineering design and drafting. Data in this category would include plans, drawings and specifications that would be put to use for analysis and design of the given facility.

- *Managerial* Comprising work related to the management of a given project. Data in this category would include estimated quantities, production standards, construction or design schedules, resource requirements and the like. These would be used for manpower and resource allocation, project monitoring and budget control.

- *Administrative* Comprising work related to daily administrative functions on the project. Data in this category would be used for accounting, invoicing, payroll and other day-to-day functions.

The above basis for categorization does not take into account the static or dynamic nature of a given type of data as discussed in the previous section. A categorization to distinguish between these would therefore be pertinent and is outlined next.

- *Static* Comprising of data that is most likely to remain unchanged over the life of the project. The general conditions of the construction contract if not the contract as a whole, and specifications for materials to be used are two possible examples of data that would be placed in this category.

- *Dynamic* Comprising of data that is most likely or certain to keep changing over the life of the project. Progress reports on the amount of construction put in place, or data on prevailing site conditions as discussed in the previous section are examples of data that would be placed in this category.

- *Dated* Comprising of data that is liable to, but not necessarily, change over the duration of the project. Data such as design drawings that may need updating or alteration due to a change order as discussed in the previous section, or the construction schedule that may need updating due to unforeseen delays are examples of data that can be placed in the dated category. These types of data or parts thereof may not change over the life cycle of the project, but the nature of the typical A/E/C project suggests that they may. Thus such data cannot be placed under the first two categories defined above but need to viewed in category that is in-between these. The name 'dated' is given to this category to imply that the data in this category is valid only till a specified date, and may require revision or review thereafter.

A final method of distinction would be appropriate on the basis of whether the data is in a readily usable format or whether it is more akin to raw data that must be piped through some process or transformed in some way as discussed in section 2. to be used as information. A distinction between available data based the following two terms take this aspect in to consideration.

- *Preliminary* Comprising of data that requires some further manipulation, processing or transformation in order to be considered as potential information.

- *Terminal* Comprising of data which after passing through some required process or in its native format is ready for use without further alteration being required.

The dead and live weights that a load bearing member of a structure will be subjected to is an example of data that could be qualified as being preliminary. Such data would be processed and manipulated for structural analysis and design before being transformed into design data such as the dimensioning of the member. The designed dimensions of the member is an example of data that could then be qualified as being terminal. It maybe noted that the latter two terms outlined above provide a means of further qualifying data that falls into the categories defined earlier in this section, rather than being a basis for categorization themselves.

The summary of the discussion presented in this section can now be viewed in the form of a three by three matrix as shown in figure 2.1 below. The advantage of the matrix is that it provides a template that can be used for categorization of any available data in a unified manner. It would now be possible to place a given kind or form of data into its appropriate slot in the template. As shown above by the hatched areas the template provides six such slots for the categorization of any available data. Data placed in each of these slots can then be further qualified by giving due consideration to whether it is in preliminary or terminal form as discussed above.

	Operational	**Managerial**	**Administrative**
Static			
Dynamic			
Dated			

Figure 2.1 A template for the categorization of available data.

The ideas presented in this section provide a foundation for identifying which types of data need to be taken onto account in developing a system for the distribution of information. As mentioned earlier, the system proposed by this thesis will attempt to take advantage of certain multimedia features to aid in the task of distributing and communicating information. However before to obtaining an understanding of how the multimedia capabilities of the technology to be employed will be a relevant, it is essential to obtain a background perspective of multimedia computing in the A/E/C industry. Section 2.3 addresses this issue.

2.3 Distribution of Data

The opening sections (2.1 and 2.2) of this chapter presented a definition and description of information systems and technology in the context of this thesis. The discussion also attempted to obtain an understanding of the main functions performed by such systems, one of these being the distribution of information. Having thereafter provided a framework for the categorization of typical A/E/C data, it is possible at this juncture, to identify specific kinds of data that are appropriate for and require distribution. It should be noted however, that the central theme of this thesis is to propose a technology for developing an information system that can be used for the *distribution* of relevant data.

The objective of this section is, therefore, not to provide a complete list of all the different kinds of data on an A/E/C project that may require distribution, but rather to identify a representative few of the same in order to lay the foundation for defining what the specific requirements of such a system would be. To begin with, some of the specific forms of A/E/C project data that were mentioned as examples while categorizing data in section 2.3, have been placed in their respective slot in the three by three matrix template as shown in figure 2.2.

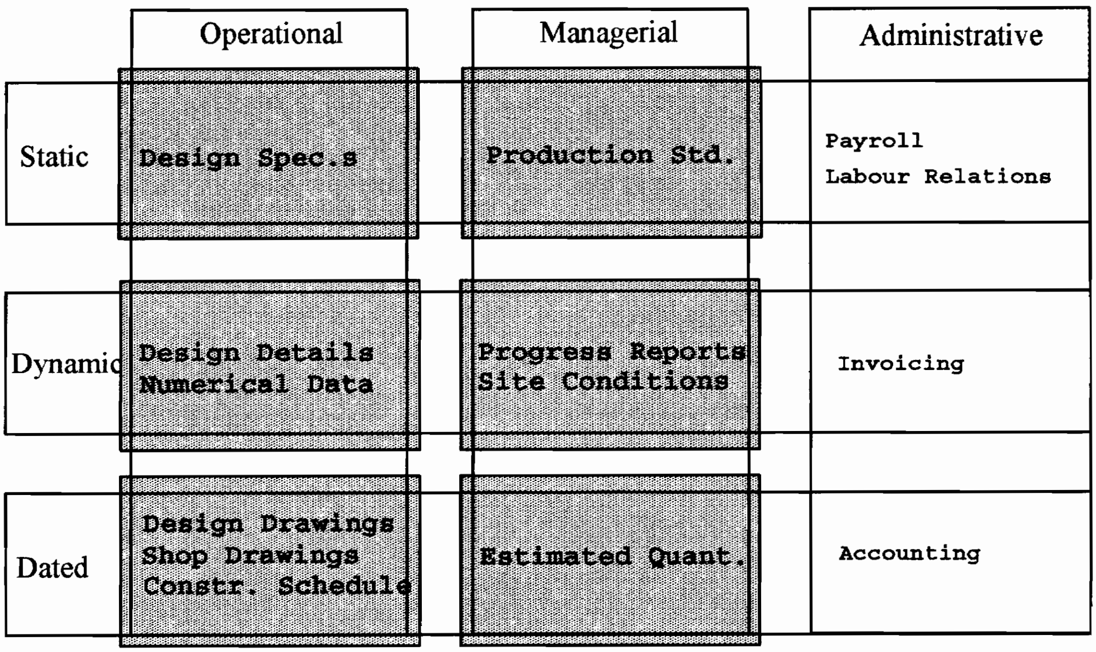

Figure 2.2 Data requiring distribution.

As discussed in Chapter 1 section 1.1, a construction project is characterized by a number of participants who are organizationally and sometimes even geographically dispersed. The objective behind the distribution of data on a project is to make it available over the widest possible area, such that it can be accessed by all parties concerned, as and when the need arises.

It follows that the data which requires distribution in such manner is of the kind that is likely to be *referenced repeatedly* if not frequently by different participants involved with the given project. Furthermore, in most cases, it is desirable that different forms of distributed data be *interlinked* in some meaningful manner, such that the user may make a meaningful association between them to aid in the process of inference and decision making. This would imply that a given form of distributed data, in most cases, should be able to complement one or more of other forms of data that are distributed, when the two are associated during retrieval by the user of the system.

The contract specifications and the construction schedule are two forms of data that can be cited to illustrate this point. Over the duration of the project unforeseen circumstances may require the rescheduling of certain activities by the contractor. The contractor would then need to refer closely to the conditions of the contract that outline the policies for such circumstances, and reschedule the affected activities accordingly. After the rescheduling is completed the owner of the project would be likely to have an interest in being kept abreast of the changes effected and whether or not these were effected in accordance with the policies laid forth in the relevant section of the contract specifications. Given the strong possibility that the owner and contractor of this hypothetical project may be geographically located in different parts of the country, the advantages of an information system that allows the distribution of this information to the above parties in real time, is quite evident. Clearly, the contract and the schedule are two forms of data that are not only required for frequent reference by both parties, but also provide complementary information when associated with one another.

Considering the Operational and Managerial categories from Chapter 2 figure 2.2, illustrative situations can be drawn up similar to the one above for most forms of data that fit into these slots. With respect to the time dependent characteristics of these slots on the template, in other words, whether a given functional type of data is Static, Dynamic or Dated, the primary factor that needs to be taken into account is how often a given type of data requires updating or revision in order to ensure that the information being provided is accurate. Naturally, Dynamic data types will require maximum close monitoring and most frequent updating, while Static types the least.

Most operational data types that are dynamic may not require distribution as their usage maybe limited to a very focused group in the project. For instance, numerical data pertaining to engineering analysis and design is likely to be useful only to the design team. On the other hand selected drawings that provide an insight into the design may require frequent referral by different parties, qualifying as data that provides information of increased value if it can be effectively distributed. Another example is Managerial data that is also Dynamic such as progress reports on the amount of construction put in place. Such data may be of considerable interest to most parties in the project.

The six slots on the template in Chapter 2 figure 2.2, all identify types of data that may require distribution. The relative importance of these would naturally vary to some extent, and due emphasis could be laid on the more prominent slots on a project by project basis. It maybe noted that most kinds of data under the functional category Administrative would not be suitable for distribution. Data pertaining to the Administrative category such as accounting ledgers, materials invoices are usually maintained separately by each of the different players for their own purposes. Such data apart from being confidential, provide little information on the project per se.

The following section (2.4) outlines how the multimedia capabilities of an information system is important by presenting an overview of multimedia computing itself.

2.4 The Impact of Multimedia

Multimedia is a new facet in computer technology that rapidly developed over the past few years. Multimedia has had a major impact on modern computer driven environments (Keyes, 1994). While this technology continues to develop with the promise of newer and more advanced features becoming available in the near future, the capabilities that are actually available today make it an aspect of computing that is almost indispensable in the development of an information system for the contemporary A/E/C environment ("Intelligent interface ...", 1993).

This section provides a generalized definition of multimedia computing and attempts to point out relevant features that could serve usefully in disseminating information in an A/E/C environment. The discussion here serves to justify the focus of this thesis on the importance of the multimedia features of the technology chosen for the proposed system.

2.4.1 Multimedia Computing

As Costea and Gillam (1993) point out a simple definition of multimedia would be the seamless capture and integration of a variety of digital media such as audio, video, graphics and text for the purpose of enhancing communication, understanding and creativity. Multimedia computing can be viewed as an aspect of information technology that seeks to utilize the multiple forms of media to enhance the effectiveness of conveying information, to the user of a given system which incorporates this technology.

There are primarily two classes of multimedia (Kindborg & Pettersson, 1991):

(a) the interactive and,

(b) the non-interactive.

Interactive multimedia allows the user to interact with the information being presented. One of the basic forms of interaction is that the user is able to choose the path of navigation or the sequence in which the material is viewed. The degree of interactivity varies depending on the technology. A majority of such systems or software is based on a scripting language that enables the author of a multimedia module to build a flexible navigational structure through the information while being able to orchestrate how the material is put together and interconnected (Costea and Gillam, 1993).

Non-interactive multimedia, on the other hand, simply presents the information in a sequence pre-defined by the creator of the system. This class of multimedia is used more for creating instructional sessions or demonstrations that take advantage of a variety of media but present the material in a fixed format.

The use of interactive multimedia has been growing in a number of engineering industries (Hodges and Sasnett, 1993). Product design processes have been aided by interactive multimedia technology. Multimedia tools that can interface with other applications have been developed. These tools serve as a front end and allow easy retrieval of different parts of the design. Tools to create interactive training systems for personnel and create product demonstrations have also been developed. Another area in which it has been used effectively is interactive communication tools for video teleconferencing. A still developing area includes combining artificial intelligence techniques with interactive multimedia to create a 'smart' environment for the retrieval of information.

2.4.2 Multimedia as a Tool for the A/E/C Industry

The advantages of being able to *visualize* data with the aid of graphics or pictures has almost universally been agreed upon. Research in this area has demonstrated the benefits of digitally archiving photographs such as those that illustrate site conditions, equipment, as well as other forms of multimedia, and linking the same to a database of related project information such as schedule or cost related data (Greishaber, 1993; "Filing of construction photos...", 1993). One of the greatest advantages proffered by the use of multimedia is the ability to visualize different forms of data that have been linked together, often in one seamless interface. In varying degrees, certain multimedia tools can now bring together disparate data forms and present them to the user in a format that is easier to understand.

The characteristics of the A/E/C industry with respect the large quantities of data and their disparate forms would appear to make multimedia a much sought after technology. In architecture, engineering and construction, multimedia computing offers the chance to improve individual or collective creativity in representing disparate forms of data, thereby enhancing the problem-solving skills of groups as well as individuals.

At a basic level there are two areas where multimedia computing can be employed in the A/E/C industry. The first of these is in creating marketing tools for the services of a given A/E/C firm in the form of a multimedia presentation of the organization's capabilities, expertise, past achievements and projects. The multimedia presentation could then be packaged in some form of portable electronic storage media such as a CD ROM or data disk and distributed among potential clients to serve as an extended form of advertising. Several progressive A/E/C firms have reportedly begun to avail of this form of advertising in their marketing efforts (Gerwick & Woolery, 1983; Robinson, 1991).

The other area in which multimedia could be beneficially employed is in the creation of tools that present information on an ongoing project to personnel involved with specific aspects of that project . Different forms of media such as graphics, audio, and video could be utilized to supplement textual information to provide an overview of different parts of the project. The broad objective would be to enhance the understanding of various personnel about the global issues involved in the particular project. At a more advanced level, multimedia tools could be used to supplement current methods of information dissemination on A/E/C projects. As discussed earlier, for an environment that is largely data rich and information poor, the use of multimedia tools to present different kinds of data in an interlinked and seamless manner could provide immense advantages in the actual use of such data as information. The information system presented in chapter 4 of this thesis attempts to take advantage of this aspect of multimedia computing.

A/E/C being an industry largely bound to traditional means and methods, the concept of multimedia computing has been greeted with more caution and skepticism than enthusiasm. To an extent this viewpoint has arisen from the traditional mind set that the scope of work undertaken in the A/E/C industry precludes the use of new forms high end technology. In part, this outlook has been reinforced by a lack of clarity in being able to define exactly how multimedia computing can be used to benefit the industry.

For a majority of A/E/C organizations an anticipation of high costs in employing multimedia computing tools on an A/E/C project has also been a deterrent in embracing the technology. However, apart from the advantages that multimedia may proffer in an A/E/C environment as discussed above, there are several reasons that make this form of technology a worthwhile investment for A/E/C organizations wishing to enhance the manner in which they disseminate their information. An important consideration, is the fact, that although multimedia is still a developing part of the computer industry with much room for improvement, the level of technology available today offers a variety of interesting features that can be effectively utilized for the purposes outlined (Keyes, 1994; Michel, 1991). Moreover, with software and hardware prices continuing to fall, multimedia technology is becoming more affordable to even small sized A/E/C firms operating within a tight budget.

In summary, multimedia computing offers the advantages of being able to link and present different forms of data from a uniform interface, the ability to expedite navigation through the information, and the means to create an interactive environment between users and their data, in the development of information systems. A superior method of information dissemination is a key factor that augments the competitiveness of an organization, and the incorporation of multimedia technology in the dissemination of information is rapidly gaining commercial acceptance in many industries (Carson, 1994; Hodges and Sasnett, 1993; Michel, 1991).. Thus the advantages that multimedia can provide in developing an information system is virtually indisputable.

The discussion so far provides a background that underlines the importance of those features of the system, and why they are integral to its effectiveness. Some further discussion of on the kinds of A/E/C data that provide additional value on a project if made available to different participants in real time is provided in chapter 4 of this document, where the proposed information system is explained. Specific types of data that the system serves to distribute are mentioned there. The thesis has so far attempted to underline the importance of being able to distribute certain types of A/E/C data, and to broadly identify which kinds of the latter require distribution. The forthcoming part of the thesis provides a brief overview of the technology required for the system that will perform the function of distribution, namely the Internet.

3

The Internet and Hypermedia

Chapter 3 provides an overview of relevant features of the technology, that is, use of multimedia computing in distributing information that is proposed by this thesis. There are two aspects to the technology, namely the Internet and hypertext computing. Prior to describing the hypertext based system on the Internet which is the information handling tool proposed in this thesis, the two facets of the technology are discussed separately. This provides the background necessary for the discussion that follows in Chapter 4, which outlines the design and conceptual set-up of the proposed information system.

As noted in Chapter 1 section 1.2, the objective of this thesis is to identify and demonstrate the applicability of hypermedia computing on the Internet towards solving the problem of distributing project related information on a construction project. This is not intended to imply, however, that this is the most replete, technically advanced or ideal solution to the problem of information management on a construction project. Rather, the focus is to demonstrate that hypermedia computing on the Internet is both a viable and attractive solution to addressing some of the key issues in the problem, and that there are certain distinct advantages to be gained from the employment of this technology in particular. Keeping this in mind, additional discussion contrasting this technology with other alternatives, computer based or otherwise has not been included here.

In order to be able to beneficially employ hypermedia computing on the Internet as a tool for expediting information distribution, it is necessary to gain some understanding of the Internet, the basic features of its working, and resources available on it. Sections 3.1 and 3.2 present a discussion on the Internet.

3.1 The Internet

The Internet is the backbone of the technology proposed in this thesis. An overview of what comprises the Internet would, therefore, be a useful to the discussion.

The Internet can be broadly defined as a vast interlinked web of computer networks spread out over the world, that transmit and receive data using the IP protocol (Hahn and Stout, 1994). However, this definition does not hint at the extent to which the Internet has come to influence the use of networked computing in contemporary times. It is necessary, therefore, to take the discussion further into how the Internet began, some of the underlying technology that it involves, and where it stands today.

3.1.1 Historical Background

The beginnings of the Internet can be traced back to the year 1969, This year saw the beginning of an experimental network involving the use of four computers called the "ARPANET" (Gilster, 1993). The ARPANET was a government initiated experiment in packet-switched networking by the Advanced Research Projects Agency (previously known as the Defense Advanced Research Projects Agency), Department of Defense, which linked research scientists with remote computers. This allowed them to share hardware and software resources such as disk space and databases (US DOCS, 1994). The concept of packet-switched networking is explained briefly in section 3.1.2.

Later, the ARPANET was split into two networks, the ARPANET and the MILNET, the latter being an unclassified military network. While sharing of information between the networks continued, access was limited to institutions involved with defense research. The Ethernet was built in the 1970s. The Ethernet allowed the connection of several dozen computers at one site in a building or on a university campus. The Ethernet was an example of what is referred to now as local area networks (LANs). The evolution of such LANs connected to the ARPANET saw a considerable growth in the size of the latter. Along with these developments it became evident that one common wide area network (WAN) would be unable to handle the demands placed upon it. As a result many more networks were built. These developments spurred similar growth of network computing among different organizations, the government, the private sector and educational organizations in the U.S. This growth also spread to other countries like the U.K, Canada and France (US DOCS, 1994).

This worldwide dispersion gave rise to the idea of connecting multiple networks of LANs and WANs without any one network having to be in control. A new set of protocols was devised to allow a wide range of networks to be connected, allowing a computer on one network to communicate with another computer on any other network. Some further discussion on protocols utilized by the Internet system is reviewed in section 3.1.2

Cooperative, decentralized networks such as the USENET (User's Network), and UUCP a worldwide Unix communications network came into being towards the end of that decade, serving universities, and later, commercial organizations (Hahn, 1993). Other networks to serve research and academic institutions such as the BITNET and the Computer Science Network (CSNET), began providing nationwide network capabilities. Though these initially were not a part of the Internet, special connections were made to facilitate the exchange of information between the communities.

In 1986 the National Science Foundation Network (NSFNET) was developed, linking researchers across the country with five supercomputer centers. The NSFNET continued to expand, encompassing state wide academic networks which, in turn, were linked to large university and research community networks, eventually replacing the ARPANET. The ARPANET was dismantled in 1990 and the CSFNET in 1991. Since then the Internet, an open-ended cross connection of LANs and WANs, each separately administered, but all adhering to the same basic protocols has grown exponentially and spread worldwide (US DOCS, 1994).

3.1.2 Communication Protocols

A protocol is a set of conventions that determine how data will be transmitted and received by different programs (Gilster, 1993; Lynch & Rose, 1993)). Such protocols also determine how the network moves messages and handle errors. The use of a standard set of protocols allows the exchange of data between dissimilar hardware platforms. This is a major feature available on the Internet.

In 1974 Robert Kahn, a major figure in the development of the former ARPANET and currently president of the Corporation for National Research Initiatives (CNRI), and computer scientist Vinton G. Cerf, currently president of the Internet Society and vice-president of CNRI, developed the Transmission Control Protocol/Internet Protocol (TCP/IP), a system of protocols that made wide area networking possible (US DOCS, 1994). The TCP/IP contributed to the success of the Internet. In fact the importance of TCP/IP to the Internet can hardly be overstated.

TCP/IP was developed using public funds, and is considered an "open", non-proprietary suite of protocols. This means that any number of companies can develop the hardware and software necessary for a network connection. As a result, there are implementations for the TCP/IP suite for virtually every type of computer system (Lynch, & Rose, 1993; Thomas, 1995).

The illustration in the figure below (Figure 3.1) attempts to represent how TCP/IP enables interconnection and data transactions to take place between a diverse array of computing platforms. The Internet can, thus be, visualized as a meta-network with a highly decentralized structure that interconnects a wide array of computing equipment from large-scale regional networks connected to high speed backbones, office networks, to stand alones like PCs, Macintosh computers, Unix workstations and so on. This allows for rapid exchange of diverse information. The sheer variety and amount of information that can be exchanged over the Internet remain unparalleled when compared to traditional modes of information transfer.

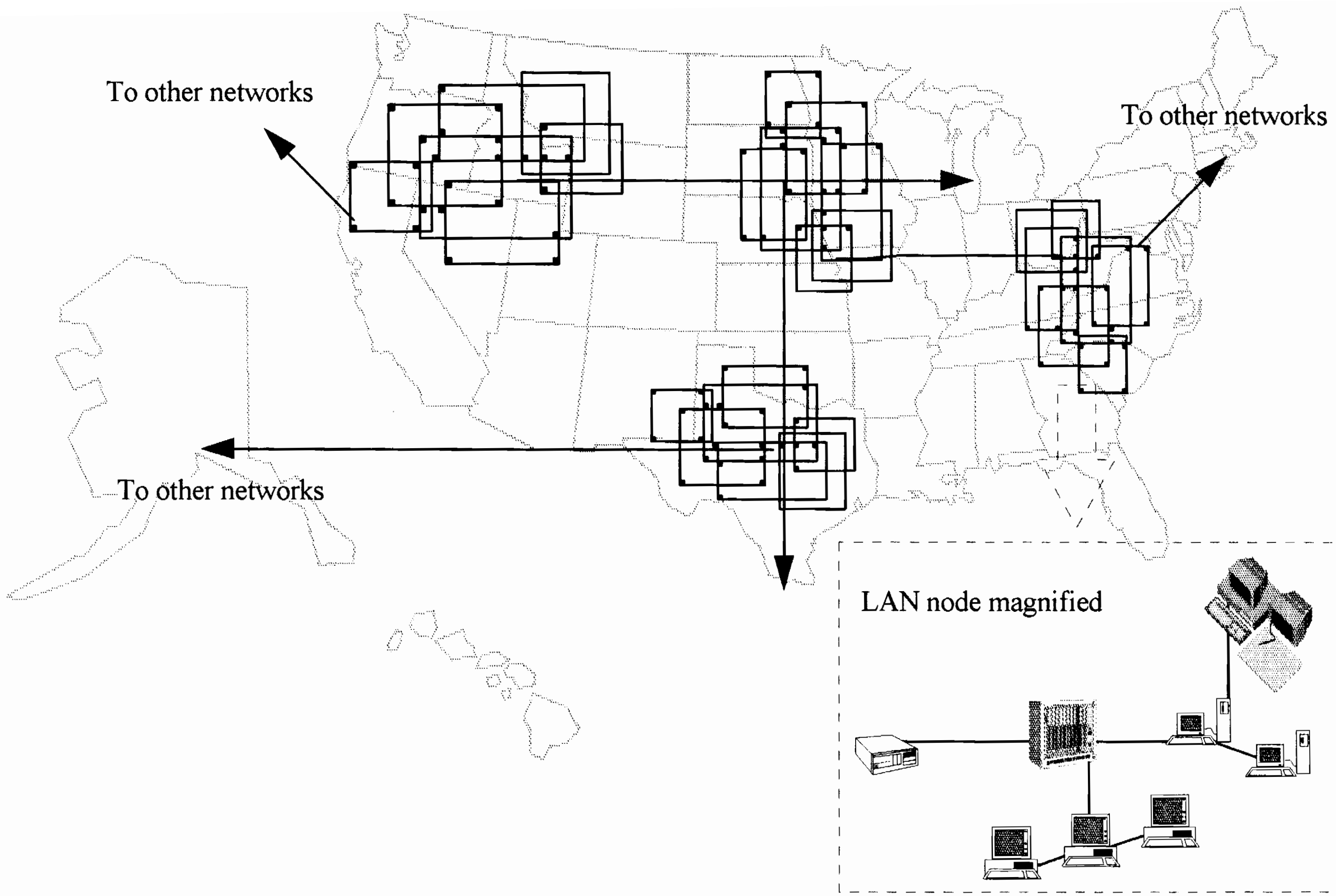

Figure 3.1 : A conceptual view of the distributed nature of the Internet

While TCP/IP is the prime set of protocols on which the Internet is based, other standards such as the Open Systems Interconnection (OSI), a protocol suite developed by The International Organization for Standardization (ISO) are also being integrated. Systems that make use of other protocols like the BITNET are connected to the Internet through special gateways that allow their users access to the resources and information on the Internet (LaQuey, 1993). On the other hand UUCP network users can gain access to the Internet from dial-up telephone lines. Figure 3.2 which is a map prepared by Lawrence H. Landweber and the Internet Society, shows the extent of the Internet's international connectivity (Gilster, 1993).

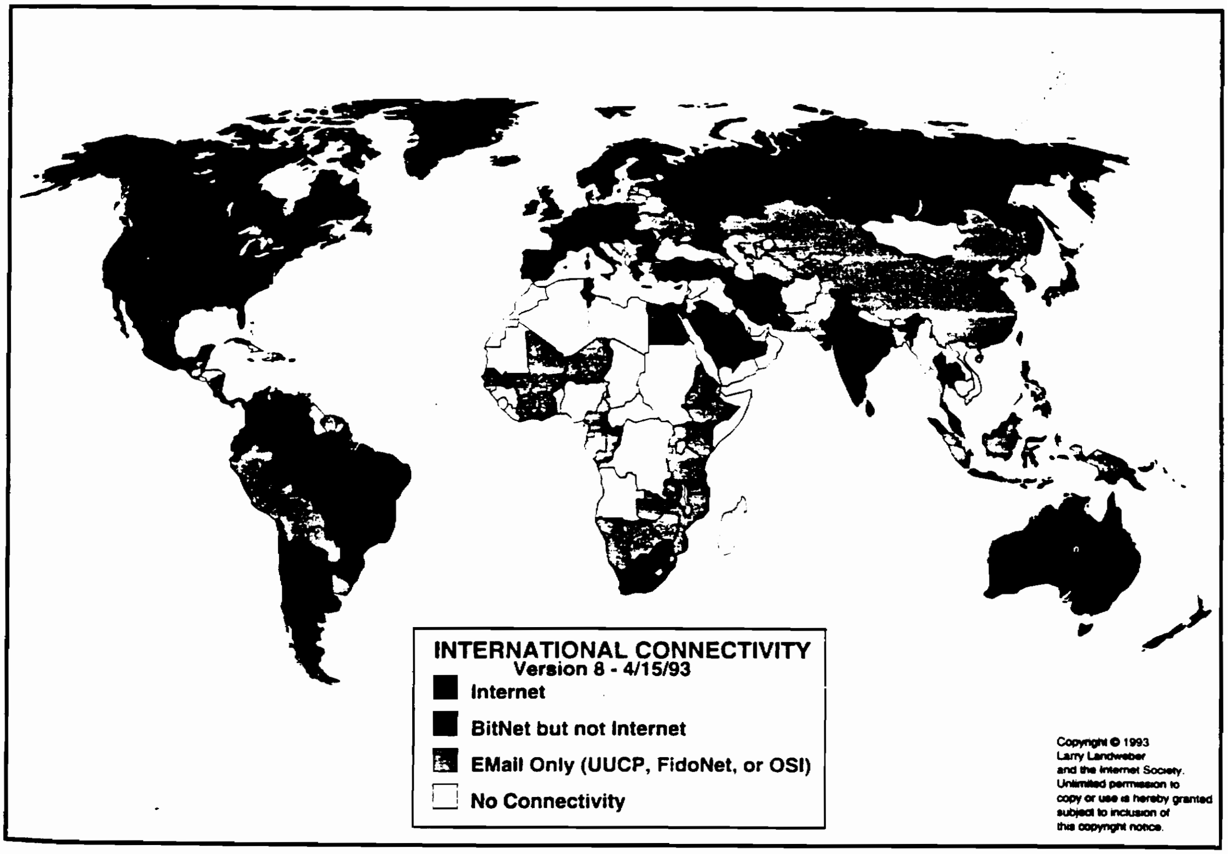

Figure 3.2 : International network connectivity

3.1.3 Packet Switching

As mentioned in section 3.1.1, the Internet began as an experiment in packet-switched networking, and continues to work on this concept. Within a TCP/IP system data that is to be transmitted is broken into small chunks called packets. Each packet usually consisting of several hundred bytes, contains some header ID tags with routing information and a sequence number (Lane and Summerhill, 1993; LaQuey, 1993). Splitting the data into packets provides certain important advantages. It enables the data communication lines to be used simultaneously by several users. This can, therefore, transmit all kinds of packets at the same time. Data is also not required to be sent directly between two computers. Instead, it can be routed from computer to computer until it reaches the destination machine identified by its routing information.

The individual packets that make up a particular data message do not have to take the same route to their common destination. Rather, the network routes each packet from one computer to another (called packet switches) using the best connection that is available at that instant. It would be an oversimplification to view the Internet's connectivity as a network of wires or fiber optic cabling alone. Digital data is moved through special hardware devices called routers which connect networks and employ sophisticated algorithms to determine optimal paths for network traffic (Gilster, 1993; Lynch and Rose, 1993). The packaged data is transmitted through a variety of media which include telephone lines, standard dial-up or leased lines, fiber optic cable connections, satellite networks and even microwave radio transmissions.

The transmission of data in such packets has the added benefit that should an error occur in the transmission of a packet, only that packet need be resent, instead of the network having to resend the next message. The utilization of multiple paths ensures that if a particular path is unavailable, the network automatically finds another through which to transmit the data. The destination system then uses the sequencing numbers of each of the packets to reconstruct the data message. In the event of an error occurring with a particular packet, the destination system transmits a message to the sender requesting for retransmittal of that particular packet. Computers at either end of a packet network connection may operate at different speeds as the network acts as a buffer to adjust for the difference (Lane and Summerhill, 1993; Meltsner, 1995).

Evidently the above system contributes to a high degree of reliability in transmitting data across the network, and transmissions across a physical distance of thousands of miles, the time take only seconds (Hahn, 1993, p174; Lane and Summerhill, 1993).

3.1.4 The Basic Set Up

Internet network connections do not have a specific model. There is however an informal hierarchy amongst the networks that comprise it. This is depicted in a simplified manner in figure 3.3 below.

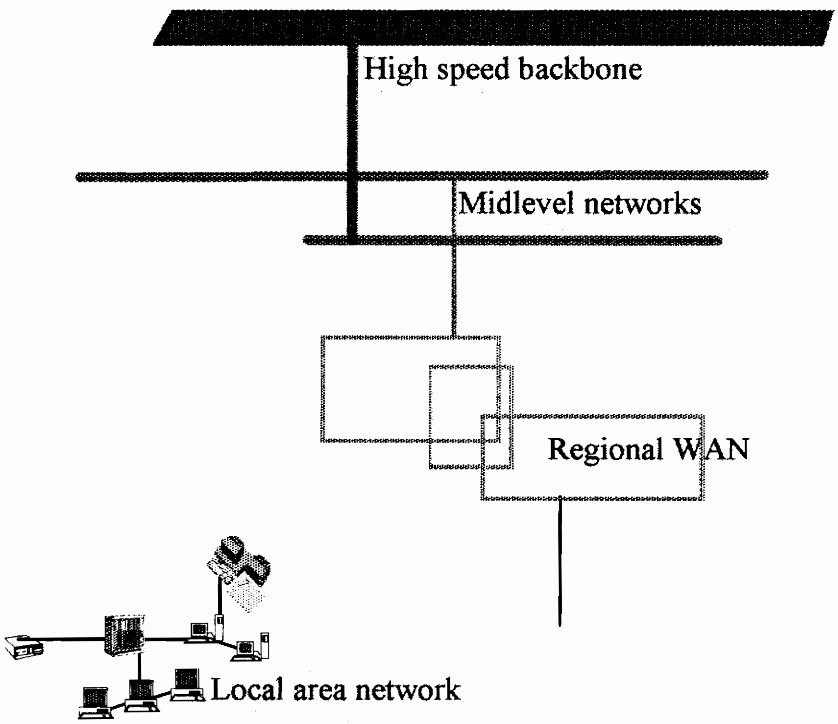

Figure 3.3 : Basic set up of networks on the Internet[1]

As shown in the figure above, the backbones, which are high-speed central networks accept traffic to and from the mid-level networks, which in turn do likewise to and from their respective member networks. Member networks distribute data to regional or

[1] Modified from LaQuey, 1993.

institutional networks which distribute data to individual computers on their net. It is important to understand however that the above set up does not place restrictions on how and at what point an Internet connection can be made. More information on requirements for connecting to the Internet has been covered in section 3.1.8.

A very significant point that encourages the extensive use of the Internet is that its numerous interconnections between these networks that number in thousands, is entirely seamless to the end user. Known as the Internet-working or Internet-concept, it effectively conceals the details of the numerous linkages via routers and packets and so on, to the user who is, therefore, able to access and utilize the resources available in a transparent fashion("The Internet ..", 1995; Lynch & Rose, 1993).

3.1.5 Administration and Coordination

An important point to be noted is that each of these networks comprising the Internet, are responsible for administering and regulating the traffic that flows within them. Hence, if two computers on the Internet that are within the same LAN need to exchange data, there is no need for that traffic to flow out of the net and traverse the backbone before reaching its destination. Routing, in this case, is performed from within the LAN. A broad operating principle in the Internet community is that each network is responsible for its own funding, administrative procedures, and, therefore, remains completely autonomous. This has been another contributing factor to the phenomenal growth in size of the Internet (LaQuey, 1993, p28; Gilster, 1993; Hahn and Stout, 1994).

Although the Internet is decentralized in nature as discussed above, its activities are coordinated to some extent by a number of agencies. These agencies include the Internet Society, Internet Architecture Board, The Internet Engineering Task Force (IETF), Internet Research Task Force (IRTF) and Internet Assigned Numbers Authority (IANA), and the Computer Emergency Response Team (CERT), among others. These bodies have varying roles including promoting the growth of the Internet into a global research and information infrastructure, coordinating the research and development of TCP/IP protocols and its standards and developing technology that may be applicable in the future (US DOCS, 1994). The role of the CERT is mentioned in Chapter 4, section 4.6.1 of the thesis, when security on the Internet is discussed.

3.1.6 Resources and Services

Resources on the Internet refer to the information, software, hardware and other useful entities that can be accessed by its users. Internet resources are as vast in quantity as they are diverse in quality, and defy any attempt at quantification. Information available ranges from that for entertainment to the highly scientific. Access is available to a host of hardware and peripheral devices, as well as a great deal of useful software.

Internet services can be broadly categorized as being the mail service, the file transfer service, the remote login service, and the USENET news service. It is well beyond the scope of the current discussion to detail each or any of the above features in depth. However, mail service, the file transfer service, the remote login service are relevant and integral to the proposed system described in Chapter 4 of the thesis, and a short description of these is outlined below (Thomas, 1995; Katterman, 1995; Finn, 1995).

Mail Mail or electronic mail (*e-mail*) refers to the transmission and receipt of electronic messages or files between computers on a network. The mail service allows to and fro exchange of messages between any two properly configured computers on the Internet. Between computers that have direct connections to the Internet message transit times vary from being instantaneous to taking a few seconds.

File Transfer File transfer refers to the ability to transfer files from one networked computer to another using the *ftp* program. The acronym ftp comes from File Transfer Protocol, the protocol underlying the program.

Remote Login Refers to the ability to log in to a remote computer as a user and to avail of the resources on that computer on an interactive basis. This is achieved via the *telnet* program.

3.1.7 Use

By its very nature the Internet appears to defy accurate quantification, and estimates concerning its size and spread tend to differ. However, there can be no doubt that the growth in use of the Internet has been phenomenal. This fact is central to the idea proposed by this thesis, which attempts to suggest a technology whose viability is attested to by its widespread and growing use. A brief note on how the Internet has grown in terms of size and use is provided below.

In 1982 merely a 100 odd computers constituted the ARPANET, and a few score others formed the NSF-sponsored CSNET which also availed of the Telnet public data network. By the end of 1990 there were 313,000 computers registered on the Internet, by 1992 the number had climbed to 727, 000 (Lattor, 1992). As of March 1994 there were 2.2 million computers on the Internet, linking approximately 27,000 networks in over 70 countries around the world. It has been estimated that approximately another 30,000 networks avail of the technology though they are not considered to be connected to the global Internet. (Lattor, 1992; US DOCS, 1994).

The Internet system is tied into a majority of public and many private messaging services that extends the number of users able to, at least, exchange e-mail to approximately 20 million. Users include businesses, professionals, researchers as well as private individuals. It has been estimated that the current rate of growth per month is about 15% (Gilster, 1993 and US DOCS, 1994). An idea of the monthly growth can be visualized in figure 3.4 through 3.6 (Dillon et al. 1993).

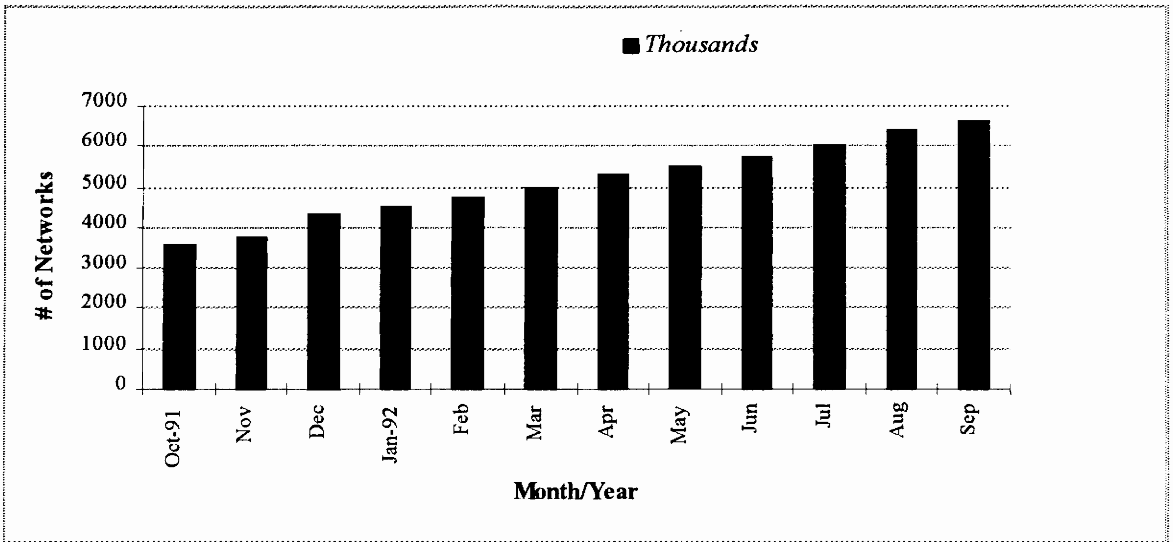

Figure 3.4 Growth in the number of networks on the Internet between October 1991 to September 1992

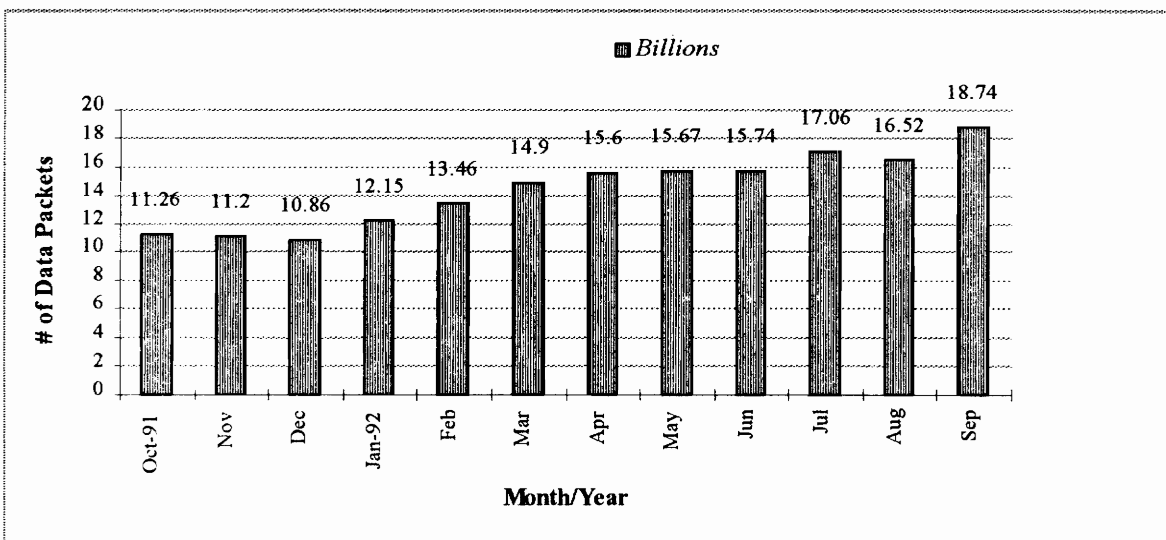

Figure 3.5 Growth in network traffic in terms of data packets on the Internet between October 1991 to September 1992

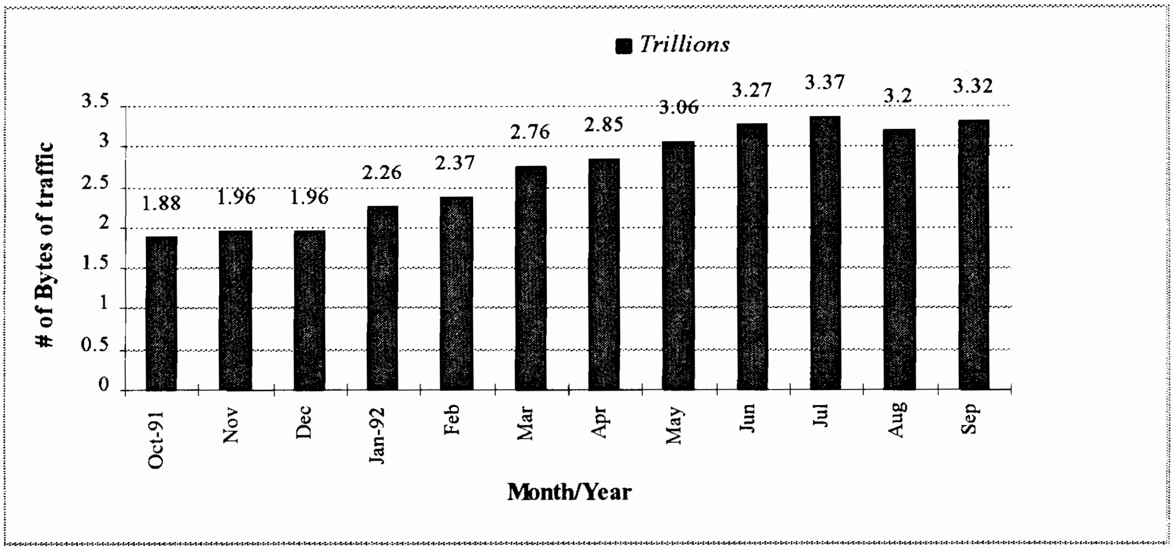

Figure 3.6 : Growth in network traffic in terms of bytes on the Internet between October 1991 to September 1992

The system has been doubling in size annually in terms of users, networks, hosts. Clearly the Internet is growing faster than any other form of telecommunications system, including the telephone network. Highly pertinent to the context of this thesis, is the fact that, over half of the networks registered are associated with commercial users. While it is evident that this rate of growth cannot continue indefinitely, it has been estimated that the user population will exceed 100 million by 1998 (US DOCS, 1994).

3.2 Connectivity Issues

At the most basic level for individual computer users or small businesses, getting an Internet connection is not a difficult process. The only requirements in such cases would be the personal computer or workstation itself, a modem, communications software and a phone line connection.

Modems are simply devices that convert the digital signal from the computer into an analog sound wave that can be transmitted across telephone lines. Currently, many hardware vendors offer personal computers that have built in 14,400 to 28,000 bps modems, and communications software. The software sets up the three way communication data between the computer, the modem and the remote computer that is being accessed. The important requirement is that the communications package include the full range of file transfer protocols (such as Xmodem, Ymodem, Zmodem, Kermit etc. which differ from the Internet's file transfer protocols).

For high speed modems, the software must support terminal emulation. Terminal emulation enables the users computer to act as a virtual terminal for the remote computer. The widest used terminal emulation is the VT 100 terminal emulation, named after the VT 100 terminal once produced by the Digital Equipment Corporation.

The requirements for connecting the LAN or WAN of a commercial organization to the Internet is a more involved procedure. This entails obtaining the connection through a network provider such as the Commercial Internet Exchange (CIX), which include CERFnet, PSI, US Sprint, and UUNET. The organization would also have to take care of certain administrative details such as registering for its own Internet Protocol network number and a unique domain name that identifies the organization's network with respect to the Internet. These identifiers can be obtained from the US Internet Registrar, Defense Data Network (DDN), Network Information Center (NIC), operated by Government Systems, Inc. (LaQuey and Ryer, 1993). The computers on the network with Internet access will have to run the TCP/IP suite of protocols, and some specialized equipment such as routers, and CSU/DSU will have to be maintained by the organization. The CSU/DSU will handle the connections between any leased digital data lines dedicated to telecommunications, and the routers whose function has already been explained.

Connection to the Internet can also be achieved through Serial Line Internet Protocol (SLIP), or Point-to-Point (PPP) protocols. This is a 'dial-up on demand' type of connection that is relatively less involved and less expensive, though with a corresponding drop in performance from the point of view of speed of data transmission.

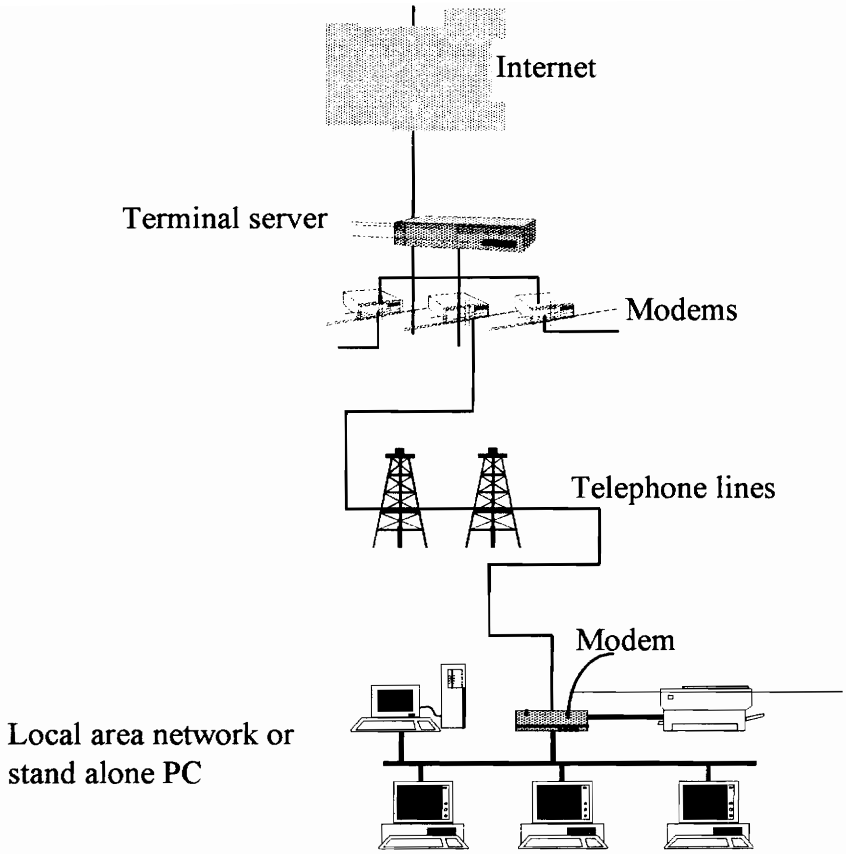

Figure 3.7 : Conceptual set up of a SLIP/PPP connection to the Internet[2].

Basically SLIP and PPP enable advanced types of client connections, that enables a direct connection via telephone lines making the computer a peer computer on the Internet. This is unlike the standard client software access, which provides only access through an network provider, and does not permit direct interactive exchange of data between the Internet and the computer (Meltsner, 1995).

[2] Modified from LaQuey, 1993.

With respect to rates of data transfer, the Internet is the fastest global network in existence. The NSFNET has among the fastest overall speeds of data transmission at around 45 megabits per second (Mbs), using what are known as T3 lines. The T3 backbone became operational in December 1992, and transmits data at approximately the equivalent of 1400 pages of single spaced typescript per second. Another kind of line called the T1 transmits data at a rate of 1.5 Mbs (Lane & Summerhill, 1993).. In general, the speed of data transmission is a function of the technology that has been employed. Direct connections are of course faster than connections via dial-up telephone lines.

3.3 Hypertext and Hypermedia

The second part of the technology that is central to the solution proposed by this thesis is the utilization of hypermedia computing. Although the information system presented in Chapter 4 will be based specifically on hypermedia computing on the Internet, it is essential to provide some background of hypertext and hypermedia technology in general. This section outlines the basic issues of hypermedia such as mechanisms and terminology, and points out the distinction between the concepts of information *searching* and *browsing*. It also discusses some of the relevant considerations in processing information using hypermedia technology.

3.3.1 Terminology and Mechanism

"Hypertext can be defined as a system to manage a collection of information that can be accessed non sequentially. It consists of a network of nodes and logical links between nodes (Lucarella, 1990). Glushko (1991), emphasized the evolutionary relationship of hypertext to printed documents. He defined hypertext as "a computer repository of information organized for access in ways that are analogous to nonlinear access conventions that are used in printed information". A more comprehensive description is provided by Smith and Weiss (1988), who state that hypertext is an approach to information management in which data is stored in a network of nodes connected by links. Nodes can contain links, text, graphics, audio, video as well as other forms of data. The nodes, and in some systems the network itself, is viewed by means of an interactive browser, and manipulated through a structure editor.

The nodes and links that can be set up in hypertext provide a very flexible structure in the presentation of the information. Since the information stored can be in a variety of forms such as images, graphics, audio and video, as well as textual data, the resulting structure is referred to as hypermedia. Since the system presented by this thesis proposes the active representation of information in multimedia format where appropriate, the terms hypertext and hypermedia are used interchangeably to mean the same thing. It is useful, however, to note the distinction between multimedia per se, and hypermedia. The term multimedia which was defined earlier in Chapter 2, section 2.4.1 of this thesis emphasizes the technology base, whereas hypermedia emphasizes the applications that such technology helps perform. Thus while multimedia presentations need not involve hypermedia, hypermedia usually involves the application of multimedia technology (Glushko, 1991).

In describing the mechanism, the central idea of hypertext is the manner in which information is represented to the viewer, or as more appropriate to the current context, the user. In the printed form such as a book, information is presented in a fairly rigidly structured format. The reader is expected to make progress through the matter in what can be visualized as a two dimensional manner. Typically this would suggest that the reader progresses from (a) the top to the bottom of a printed page, and (b) sequentially from the first to the second chapter and so on.

In contrast, a hypertext information system provides the user with a three dimensional environment. Instead of being led through a body of information sequentially and in a rigid hierarchical format, the user may select his or her preferred path of progression by following the built in nodes and links in the information. Having provided an idea of the mechanism around which hypertext revolves, it would be expedient to summarize the terminology that is commonly used in describing hypertext systems (Croft, 1990; Jonassen, 1989; Nielsen, 1990).

Documents Documents are analogous to the pages in a printed book, and are the elements of information in a hypertext system. Usually this term refers to a digital text file that may or may not have embedded images, graphics, multimedia and links to other documents. Often such documents are text files that have been converted to hypertext using a hypertext markup language.

Markup Tags Hypertext systems are built out of documents that have been marked up with special hypertext tags. Mark-up tags are defined by a mark-up language such as the Standard Generalized Markup Language (SGML) or the Hypertext Markup Language (HTML). HTML is a subset of SGML and the standard markup language for hypertext authoring on the Internet. The tags in such a language mark all the text structures in a document to provide the desired formatting such as bold face, italics, paragraphs etc.

Link Anchors Link anchors define the elements between which a link is established. As such link anchors can be thought of as special tags that possess certain attributes. The function of a link is usually explicitly defined by a script which determines the element or document to which a jump will be executed. The visual entity that behaves as a link maybe textual, or an image, and in hypermedia systems are often referred to as hotlinks.

Nodes A node usually refers to an entire document which may contain several hotlinks to other nodal documents.

Jumps A jump can be looked upon as the event of going from one document to another as a result of selecting a hotlink. The number of jumps that can be executed in a given system usually depend on the manner in which that set of documents has been authored. The ability to provide jumps in a hypertext system is central to the idea of three dimensional presentation of information, as these provide the user with several choices of navigating through the information in whichever order he or she may desire.

In general, a set of hypermedia documents are characterized by their content, structure, style and access (Waterworth, 1990). Content refers to the actual information contained in the document, while structure refers to the manner in which the component documents of the system have been linked together and presented. Style refers to the nature of interactivity built into the system and the manner in which the information has been displayed. Access determines which users are permitted to read and interact with the information presented by the system.

3.3.2 Information Browsing

In proposing the use of hypermedia for designing an information system there is a basic but important distinction to be made between information retrieval systems based on a query mechanism and those based on a hypertext mechanism. The difference between these two kinds of systems can be depicted by the simple flow chart shown below (Croft, 1990).

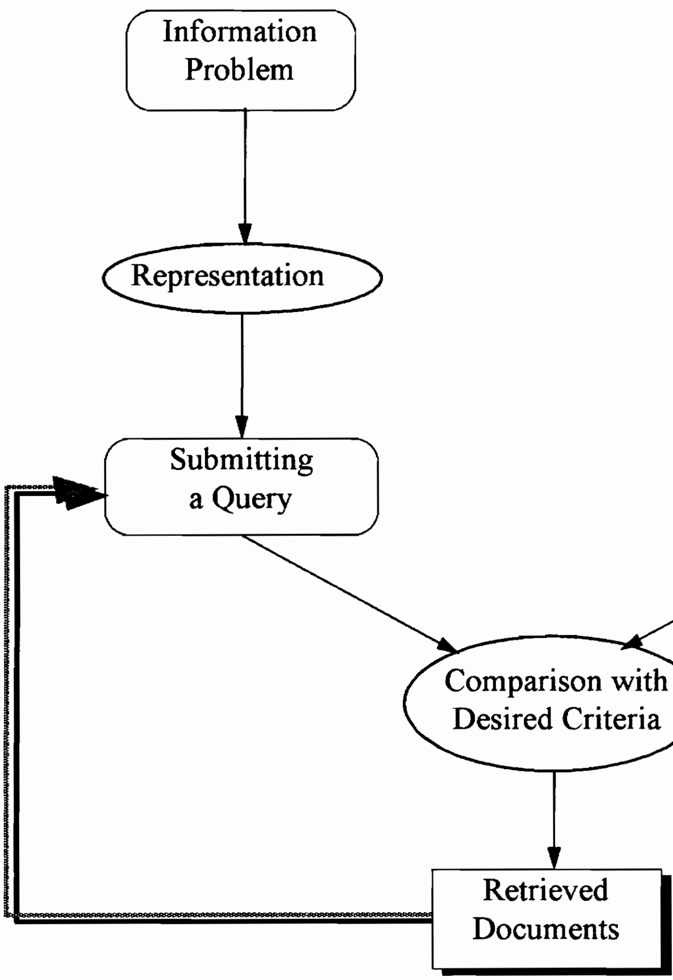

Figure 3.8 : Query and browsing in information systems

As the above figure (figure 3.8) shows the main distinction between the two mechanisms is that hypertext systems lay stress on multimedia data and representation of networked information which facilitates browsing instead of queries. This is one of the characteristics of hypertext systems that facilitate the browsing through a great deal of information, that make the adoption of this technology advantageous for the problems in information management identified in earlier in Chapter 2 of the thesis.

3.4 The Importance of Hypertext

The concept of hypertext is not new, though it is only in the current decade that computer technology has advanced sufficiently to enable its usage as an information processing and handling tool. As stated earlier, the concept appears to have evolved from the manner in which printed information was processed. Proponents of this technology have argued that hypertext imitates the human mind and it speeds and greatly facilitates the process of information retrieval (Zimmerman et al. 1993).

Another important aspect of this technology is the creation of certain standards in its structure and authoring that have made it platform independent. The primary standard for hypertext today is SGML (International Organization for Standardization, 1986). SGML was made an international standard in 1986 for representing the logical structure of documents irrespective of the specific kind of hardware or software used to create them. An SGML document contains standard tags or mark up, making it significantly easier to exchange and process electronic documents as the documents have fewer assumptions regarding their mode of processing or final appearance. Text tagged in this fashion, therefore, is more suited to on-line viewing than that formatted for printed output.

In addition to this, are the significant advantages proffered from being able to customize the navigation within the information, and being able to design a custom interface. The latter is discussed in greater depth in Chapter 4 of the thesis.. The advantages of being able to reduce the amount of actual paperwork that requires processing has already been discussed in Chapter 1 of the thesis. Further discussion on the advantages of adopting the specific system proposed by this thesis are pointed out in the next part of the thesis.

This chapter has thus outlined the relevant features of the technology chosen to address the problem of information distribution discussed in chapter 2. Chapter 4 that follows, now presents HARM the proof of concept module for this thesis. The description of the module serve to demonstrate how the technology discussed in this chapter can be utilized, in what manner it maybe set up, and the main features that can be taken advantage of.

4

Hypermedia Archiving and Retrieval Module

The preceding chapter of the thesis provided a background and description of the Internet and hypermedia. As discussed earlier the intent of this thesis is to propose hypermedia computing on the Internet as a solution to the problem of information distribution and management, and demonstrate it by building a proof of concept that incorporates some of the features already discussed.

In Chapter 4 of the thesis, the Hypermedia Archiving and Retrieval Module for Architecture, Engineering and Construction (HARM A/E/C) is presented as this proof of concept. This part of the thesis begins with an overview of the module. It also attempts to point out the various possibilities and issues involved in utilizing the chosen technology, by discussing it in the context of HARM.

4.1 Overview

The proof of concept, HARM, is a module of interactive hypermedia documents, providing point and click access to a variety of construction project information through a uniform interface. The purpose behind constructing this module was to demonstrate how some of the features of hypermedia computing on the Internet can be used to address the problem of accessing a diverse range of information, over a geographically dispersed area.

4.1.1 Objective of HARM A/E/C

The objective of HARM A/E/C is to provide the ability to easily access and retrieve project information in real time from virtually anywhere in the world, by means of a computer connected to the Internet.

This is illustrated by figure 4.1 and 4.2 in the following pages. As an example, the figure 4.1 shows the geographic distribution of projects for a hypothetical construction organization in the United States. Figure 4.2 depicts how the same distribution can be interlinked through the Internet. . The terminals connected by the dotted lines in the figure 4.2 are a conceptual depiction of the Internet, and how it can be used to convey information over a great distance. For instance, the figure shows how information on a project located on the west coast of the United States, can be viewed from a computer located in the eastern part of the country by means of a hypermedia browser.

The use of this technology would provide the ability to retrieve and view this information from any of the terminals connected by the dotted lines shown. In other words, any computer that is connected to the Internet and has a hypermedia browser such as Mosaic or Netscape running on it can be used to access this information.

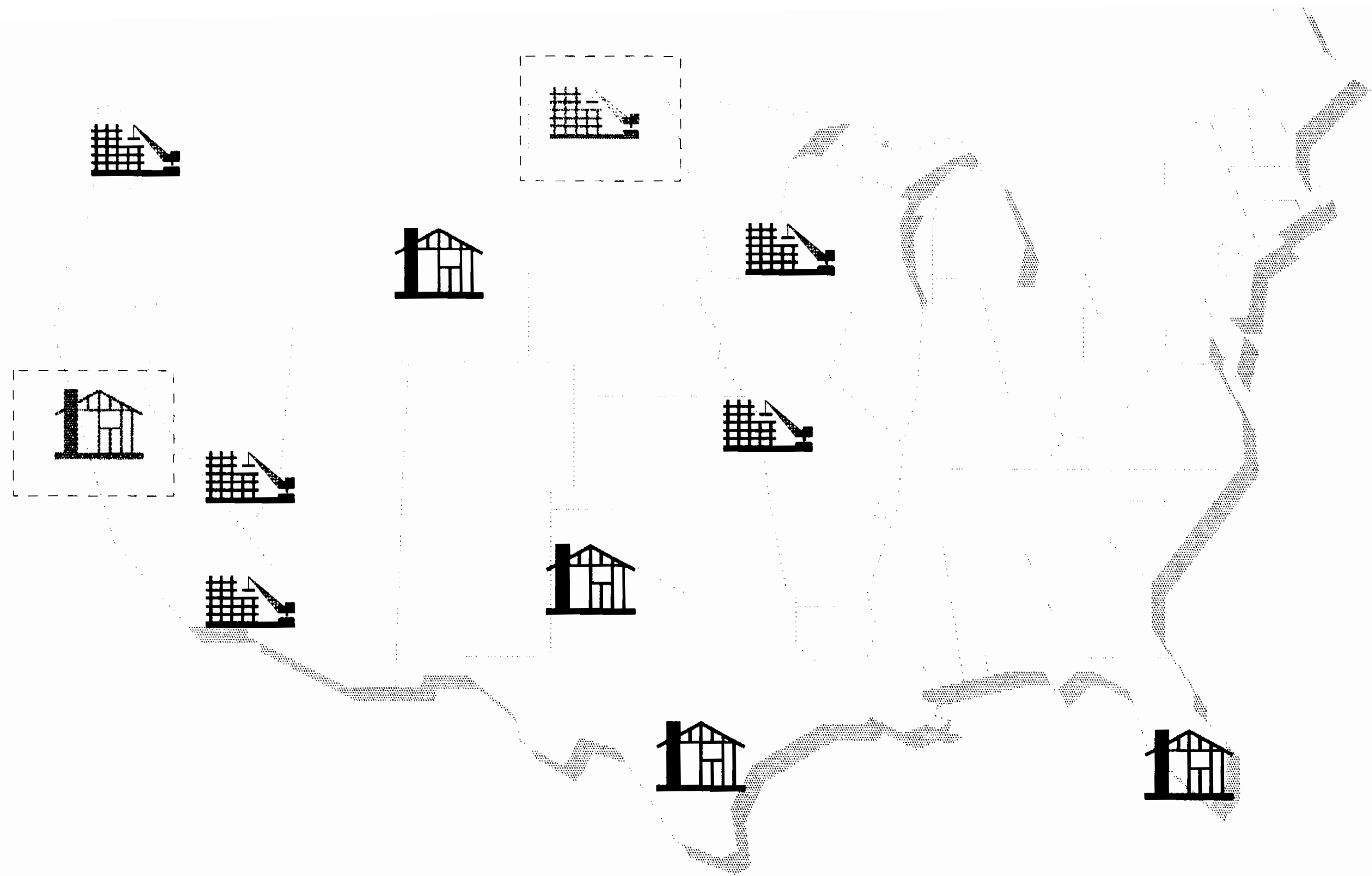

Figure 4.1 : Map showing geographic dispersion of projects in the U.S of a hypothetical construction organization
Note the projects marked with the dotted rectangle in figure 4.2

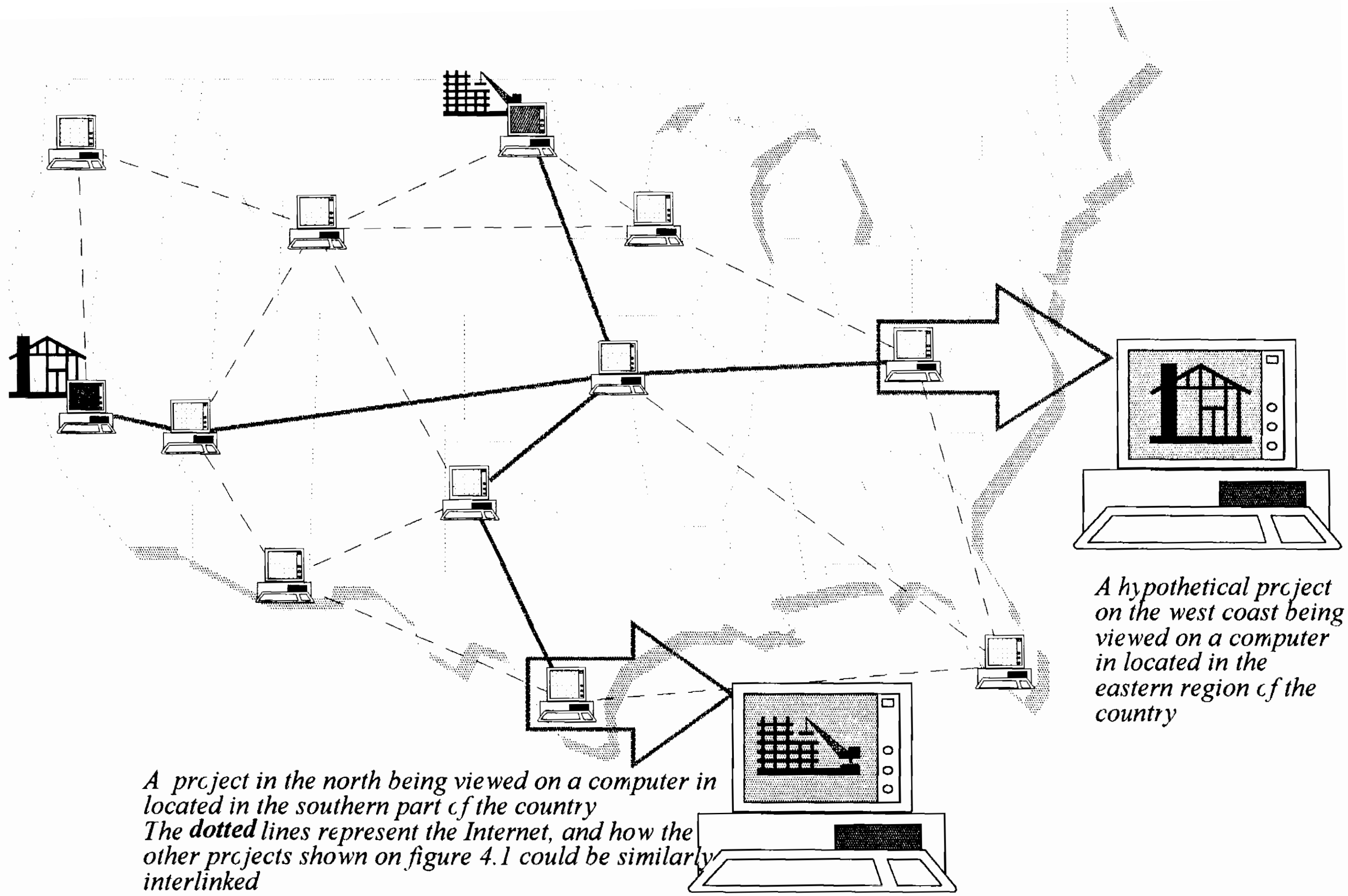

Figure 4.2 : The idea depicted here is that information on a given project can be accessed from any location greatly reducing the problems of project participants being geographically dispersed

4.1.2 Information Accessible through HARM

The Chapter 2, section 2.2, of this thesis provided a discussion on the various types of information and data that characterize a typical construction project. This chapter (Chapter 4) discusses how HARM provides the ability to access information on multiple projects. Based on the discussion presented in chapter 2 of the thesis regarding the categorization of information, figure 4.3 shows the first few levels of breakdown of how information access is organized within this module.

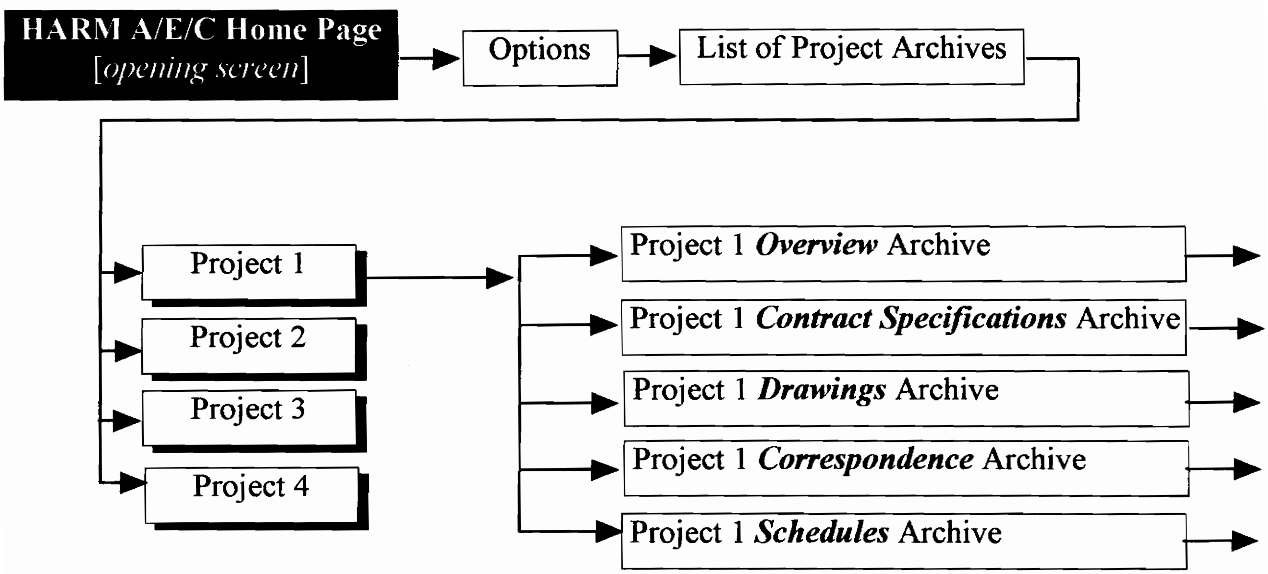

Figure 4.3 : Organization of information within HARM A/E/C

As shown in the above figure, specific information on a particular project is grouped under the number assigned to that project. Under each project, detailed information is available from the following archives.

Project Overview Archive This archive contains information that provides a broad overview of the project. Such information may include an introduction to what the nature and scope of the project is, and the parties involved in the project. If a more technical level is desired, this archive may house information on the architect's design intent, and how these issues are being addressed to by the contractor.

Contract Specifications Archive This archive houses the documents that comprise contract specifications for the project. The entire set of specifications may be made available on-line in this archive for quick and easy interactive retrieval, as opposed to having to search through volumes of paper documentation.

Project Drawings Archive The drawings archive may consist of design drawings, shop drawings, and even architectural drawings. Drawings can be stored and retrieved in their native file format, such as a CAD file in the data exchange format (DXF), or as scanned images.

Project Schedules Archive The schedule for preparing architectural and structural plans and drawings, and performance of quality reviews and checks is provided in this archive.

Project Correspondence Archive This archive contains selected correspondence between parties involved in the project. This adds value by providing information on communiqués and commitments that have been put on record in the course of the project. This does not, of course, include such correspondence that is deemed to contain information that is sensitive in nature.

Other Information HARM also provides point and click access to other information resources available on the Internet that are of relevance or interest to the A/E/C industry. These pointers or hyperlinks are to be found under the 'Other Resources' category of the module and attempt to encourage users of the module to exploit the vast wealth of information that is available through the Internet.

It maybe noted that since the actual on-line version of HARM A/E/C has been developed for the purposes of demonstration only, it contains actual information on only one project. However, the set up of the module has been designed to handle an indefinite number of projects as further discussion will make evident.

4.1.3 Functionality

The hypermedia technology on which HARM is based provides point and click access to all of the information accessible from the module. As discussed later in section 4.4 of this chapter the module has been constructed to permit the user to navigate through a great deal of diverse information quickly and with ease. By pointing and clicking on a given hyperlinked phrase or button, the user is able to ("Intelligent document...", 1992; "Intelligent interface...", 1993).

- Interactively access more specific information on a particular subject
- View scanned images or documents
- Communicate by means of electronic mail with other users connected to the Internet
- Launch certain application programs running on the users' local machine.

- Launch an automated file transfer session, to retrieve data or access a remote database.

In addition, most Internet browsers including Mosaic and Netscape have certain standard features that add to the convenience of using this technology[1]. The user can save a retrieved file, regardless of format, to a storage device such as the hard drive of his or her local machine. Users cannot, however, make any changes to the information they are viewing unless they have the authorization to access the HARM information server. The following sections discuss certain specific features of HARM and the set up of the module.

4.2 The Set Up

This section describes the set up proposed for HARM-A/E/C. The discussion on the set up comprises two basic issues. The first relates to how information has been structured internally within the module, and the intent behind such a structure. The second issue concerned with external set up, outlines how the module may be deployed on the Internet by a construction organization, and the basic technical requirements to do so.

---

[1] Source: NCSA Mosaic On-line documentation at URL http://www.ncsa.uiuc.edu/SDG/Software/WinMosaic/HomePage.html, and Netscape on-line documentation at URL http://home.mcom.com

4.2.1 The Internal Set Up

The internal set up of the module refers to how and where the information served by HARM is stored. To explain how this is done, figure 4.4 graphically depicts the directory breakdown structure of the module.

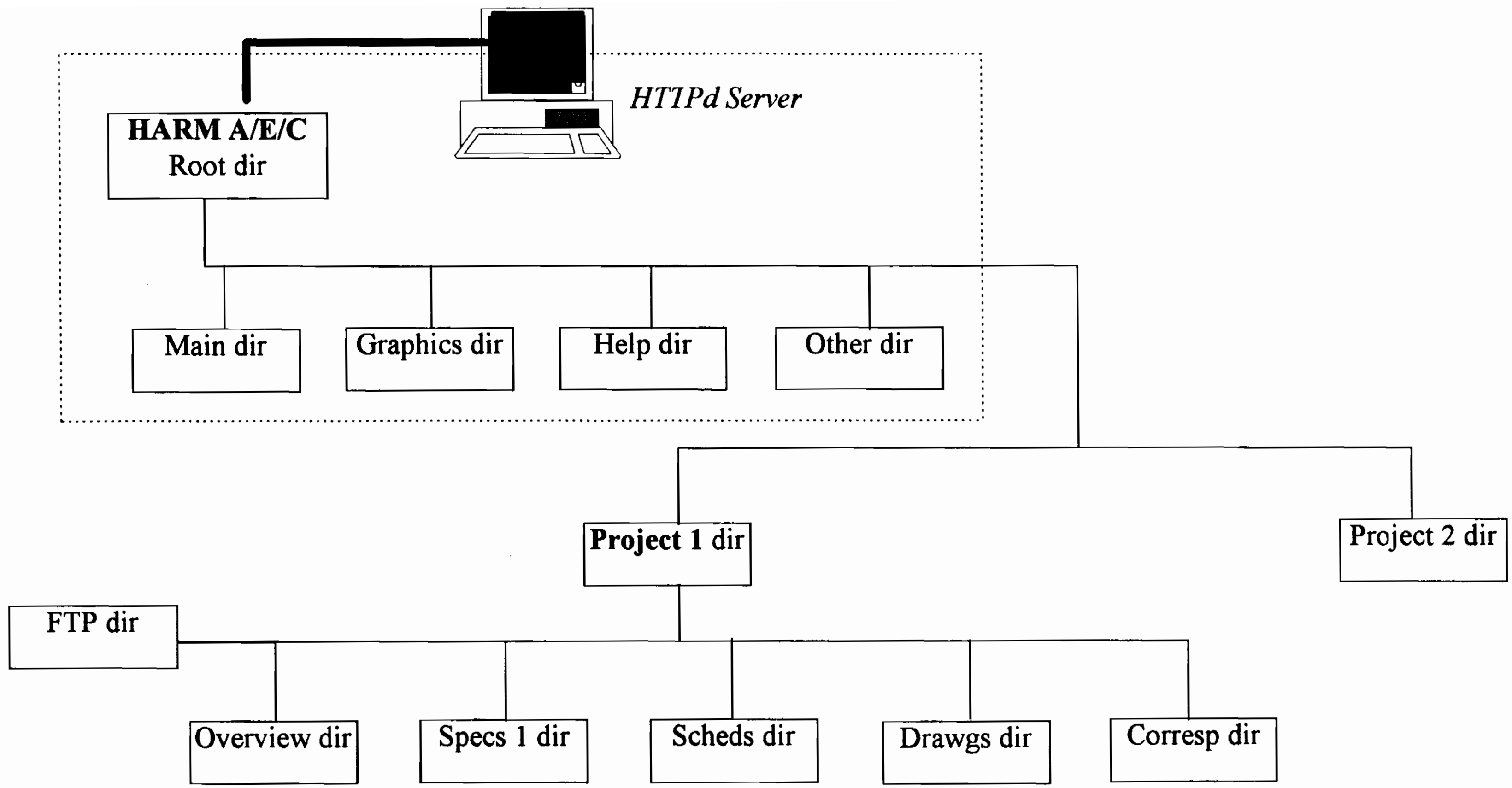

Figure 4.4 : The HARM directory breakdown structure.

The directory structure for the module has been set up to facilitate the administration and maintenance of the system, once the system is deployed. In figure 4.4, the section enclosed by the dotted box consists of directories that are integral to the module itself. These comprise files that make up the interface, including graphics files such as buttons and icons, on-line help files, and the section of the module that provides access to other resources on the Internet. This section of the directory tree is independent of any specific project and will require relatively infrequent maintenance and upgrade. This section of the directory tree must be located on the machine set up as the main HTTPd World Wide Web server for HARM.

Keeping in mind that HARM may be required for use on several projects, each project is allotted a separate directory on the main server as evident from figure 4.4. In turn, each of the information archives under a given project is allotted a separate directory. Each of the archive directories contain the actual data files that can be accessed through the HARM interface.

The intent behind this structure is to organize the information in a modular fashion, each project being independent of any other, and each archive under a given project independent of other archives. This makes it possible to track, update or alter information under a specific archive for a particular project quickly and easily without affecting any other parts of the module.

4.2.2 The External Set Up

The external set up outlined here discusses the components required and issues to be considered for setting up HARM to serve information as discussed in this thesis. The minimum requirement for the set up is a personal computer or workstation with a direct, SLIP, or PPP based connection to the Internet, and software suited for configuring the machine as a World Wide Web server. In addition, a Web browser program for accessing the module and the Web is required. Since there is a considerable degree of flexibility in the components that can be used, the specifications mentioned below refer to the existing set up built for the proof of concept version of HARM. The discussion which follows outlines any changes that may be recommended for an actual industry scenario.

Configuration The set up of the HARM is based on the client server model as outlined in the previous chapter. The entire module is housed on a single machine that has been configured as an NCSA Mosaic HTTPd World Wide Web server. This module is accessible from any computer with direct or indirect access to the Internet, and running a Web browser as mentioned above.

Hardware The machine employed as the HTTPd server for HARM is a Unix based Silicon Graphics Iris Indigo XS24 Workstation with the X Window system. The hardware is located in the Construction Engineering and Management Laboratory on the Virginia Tech campus.

Sc ftware The server software for HARM is the NCSA Mosaic HTTPd suite version 1.3 for the World Wide Web. The client software running on the same machine is NCSA Mosaic 2.0 for the X Window system.

In addition to the above, the image maps used as a part of the navigational tool in HARM are based at a second Web server, also located on the Virginia Tech campus. As discussed in section 4.3.1, this serves to illustrate the fact that information accessible from the HARM interface can be located on several geographically dispersed machines without having any effect on the look and feel or usage of the module.

Mosaic client software, like that of most popular browser programs, is available for all three major computing platforms, making access to all hypermedia documents on the World Wide Web platform independent. However, the HTTPd software for setting up a server is currently available for only the Unix platform. Hence, setting up an IBM PC clone or Macintosh as a Web server would require an alternative software suite, though this would not make any impact on the set up issues discussed earlier.

4.3 The HARM Interface

The interface of HARM refers to the look and feel of the module. The intuitive feel of the interface represents one of the primary advantages gained the utilization of hypermedia based information access on the Internet (Kindborg, & Pettersson, 1991). There are two aspects of the HARM interface that merit some discussion. The first has to do with characteristics that are intrinsic to the use of this technology. This gives HARM a *seamless* interface and allows the user to access and retrieve information in a completely *transparent* manner. The second aspect pertains largely to the way the module was designed and authored and refers to the *uniformity* and *multimedia* characteristics of the interface, that further enhance its easy use. These are explained in the following sections.

4.3.1 Features Related to the Technology

Seamless As mentioned earlier, each screen presents several hyperlinks to other information, in the form of highlighted text or buttons. The locations of different information sources to which the hyperlinks on any screen point are likely to be widely dispersed and drawn from different computers. However, this is not made apparent to the user to whom all the information appears to be available from a single screen. This is described as the seamless characteristic of the interface (Croft, 1990).

This is illustrated in figure 4.5 which points out the actual sources of the information being pointed to by the hyperlinks on this page. To the user, however, this is neither apparent, nor does it affect the ease with which he can select and follow any of the hyperlinks presented on this page.

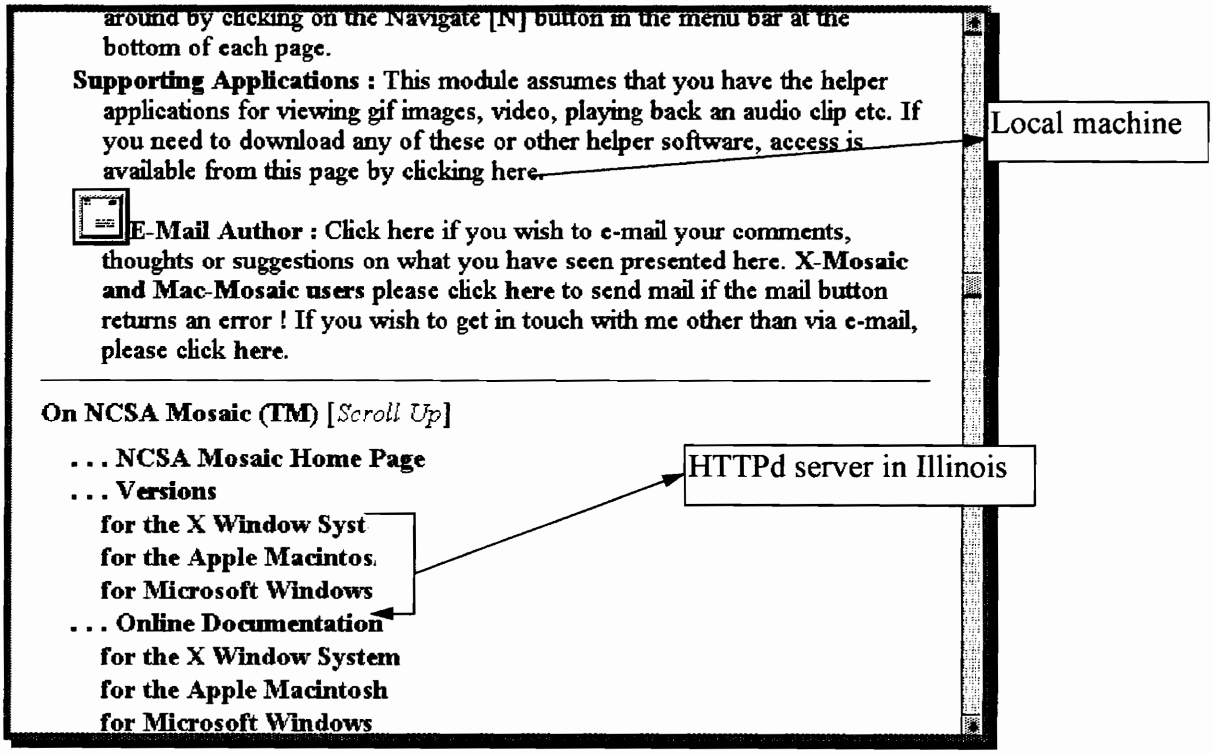

Figure 4.5 : A HARM screen illustrating the seamless characteristic (The boxes with arrows show the source to which the hyperlink points).

Transparent The manner in which the interface is linked to the information remains completely hidden to the user, who need only be concerned with choosing which link he wishes to follow (Jonassen, 1989). More precisely, the process activated by clicking on a hyperlink which causes some information on a remote computer to be retrieved and displayed on the users local machine takes place in the background. This aspect of hypermedia computing on the Internet makes it easy to use. This makes it feasible for users with virtually no prior training in the use of computers to take advantage of the kind of technology that HARM represents.

4.3.2 Features Related to the Design

Uniformity Uniformity means that the basic features of the interface which the user sees remains constant throughout the entire module (Nielsen, 1990). Instead of having to try and understand the format in which the information is presented in each different hypermedia document, the user finds the same basic options for navigating and interacting with the information regardless of where he may be in the module. In other words, the look and feel of the interface remains the same throughout the module.

There are two aspects that contribute to the uniformity of the interface. These are the standard features of the browser, and those of the module itself. Figure 4.6 shows the interface of NCSA Mosaic, the preferred browser for this module. The figure highlights a few of the buttons on the interface that contribute to some of the aspects of the functionality discussed earlier in section 4.1.3.

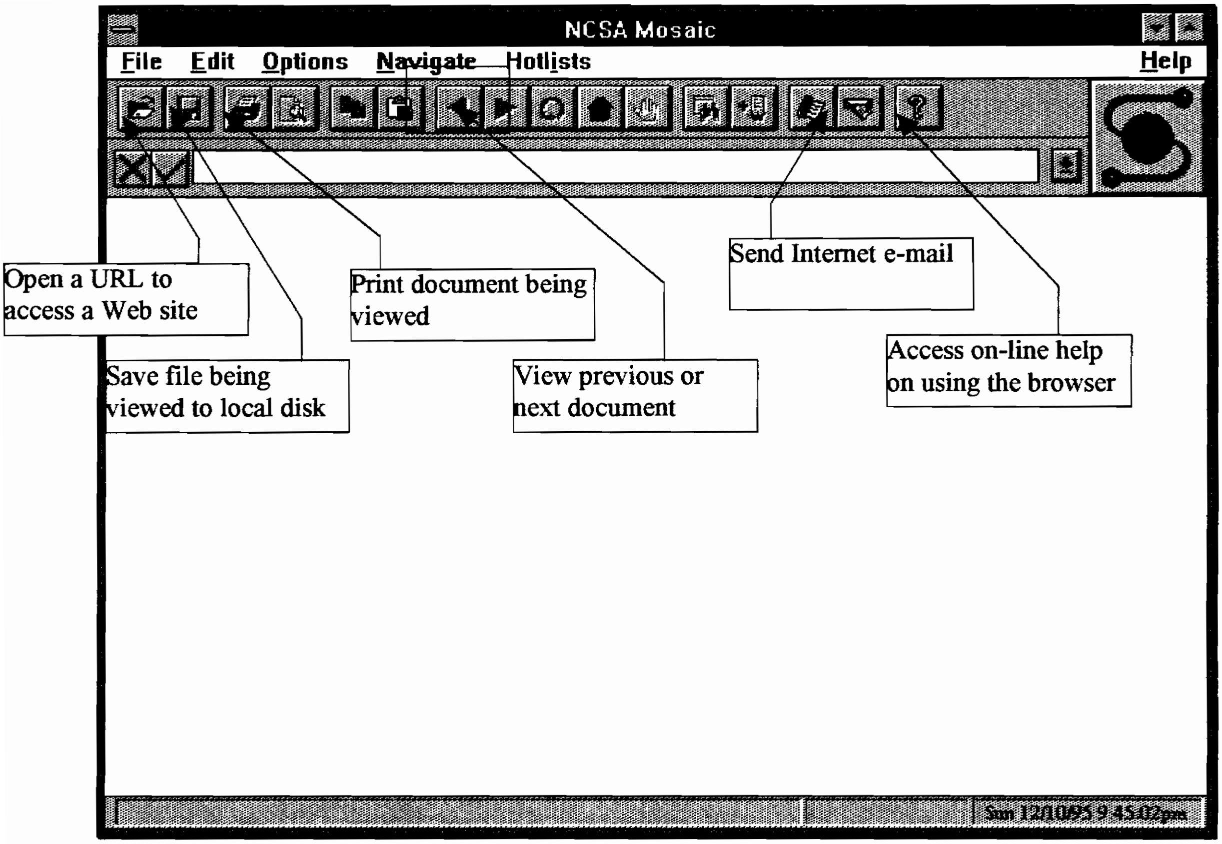

Figure 4.6 : The NCSA Mosaic browser interface.

In addition to the graphical user interface (GUI) of the hypermedia browser Mosaic, the format in which information is presented in HARM has been kept constant throughout the module. The brief discussion that follows on the navigational features built into the design, will further demonstrate the advantages to be gained from maintaining this uniformity in the interface.

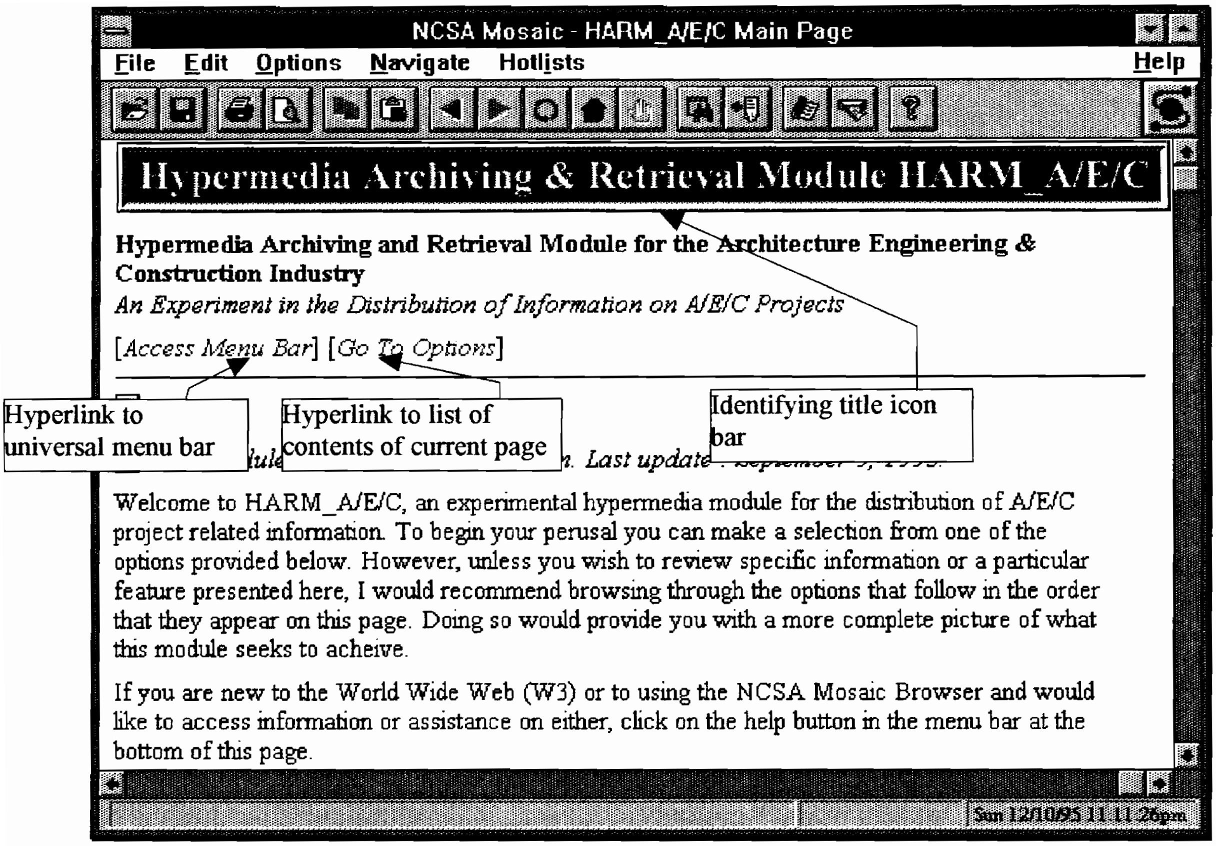

Figure 4.7 : A HARM A/E/C screen

Multimedia HARM represents an attempt to incorporate all the aspects of presenting information in a multimedia format discussed in earlier chapters. Apart from being able to access a variety of textual data, the module facilitates the use of multiple file formats, images and graphics. It should also be noted, that it is the multimedia functionality, an intrinsic part of hypermedia computing, that enables its distinguishing features such the use of graphic icons. These distinguishing features contribute to the look and feel of the module as well as the authoring of its built in navigational substructure described in the next section.

4.4 Navigation Through the Information

The discussion on hypermedia in Chapter 3 section 3.3, had mentioned the importance of being able to navigate through the multiple layers of information that are characteristic of the hypermedia environment (Kindborg & Pettersson, 1991). Research in this area indicates that users tend to experience a frequent loss of orientation due to the absence of a hierarchical structure of the format in which information is presented (Zimmerman, 1993; Malin, 1994). Since one of the driving factors behind the use of this technology is to facilitate ease in accessing information, this issue was treated with importance while authoring the module.

In the module, the solution to this problem can be found in its built in navigational tool. The flow chart shown in figure 4.8 illustrates the progression of a simple user session with HARM. The basic purpose of the navigational tool is to allow the user to access any part of the module, at any time during a session regardless of his current location. The latter is important. As illustrated in figure 4.8, the user is able to execute the first few desired jumps through direct hyperlinks in the document being viewed. However if he wishes, for instance, to access a document in the Specifications archive from his current location in the Overview archive, there may not be a suitable hyperlink within the document being viewed to execute the jump directly.

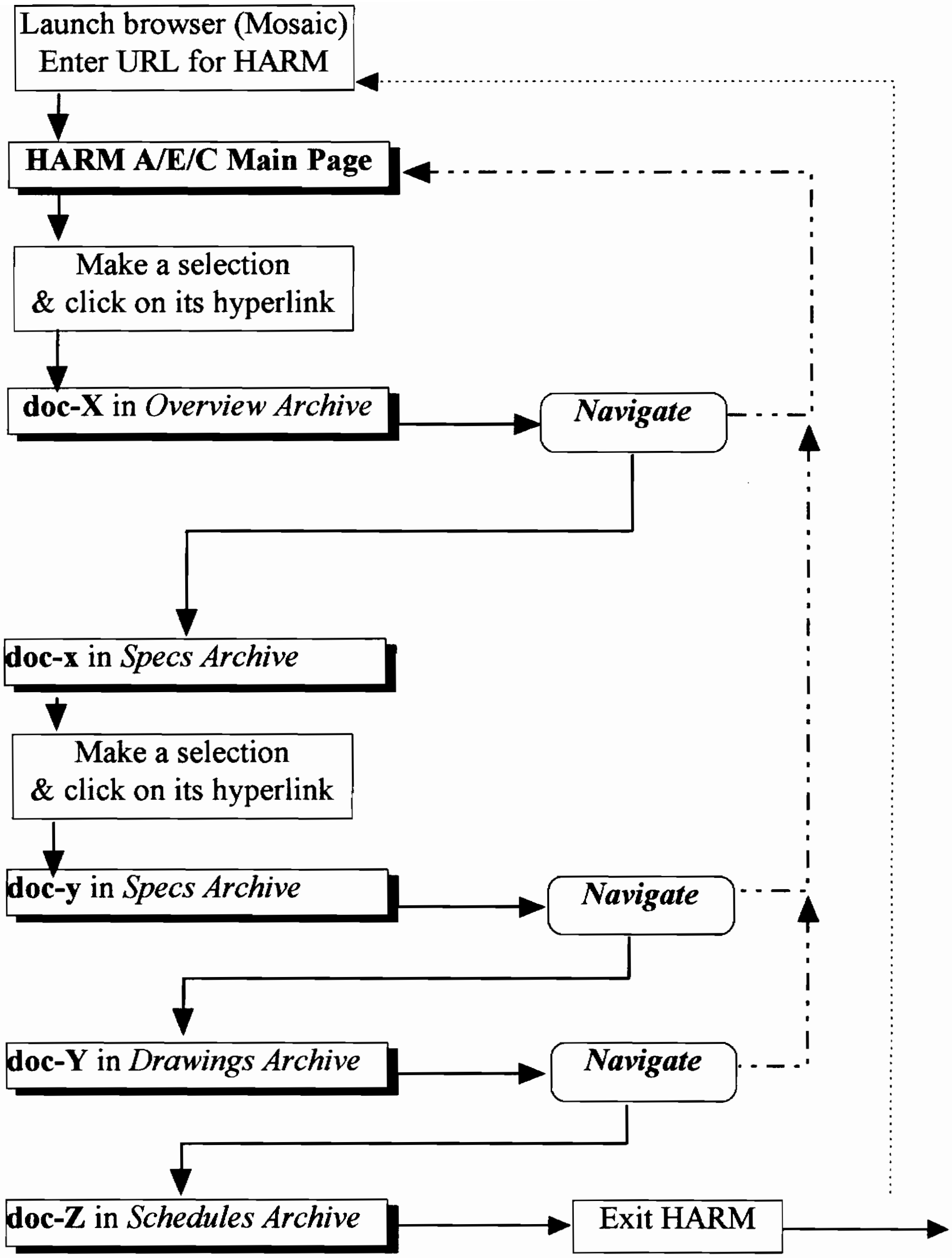

Figure 4.8: The utility of the navigational tool within HARM

Most browsers like Mosaic provide a pair of buttons (see figure 4.6) to allow the user to jump backwards or forwards. This jump, however, can only be executed to documents that have already been viewed. In the example shown in figure 4.8, this is, therefore, of no help in accessing document-x (doc-x in the flow chart).

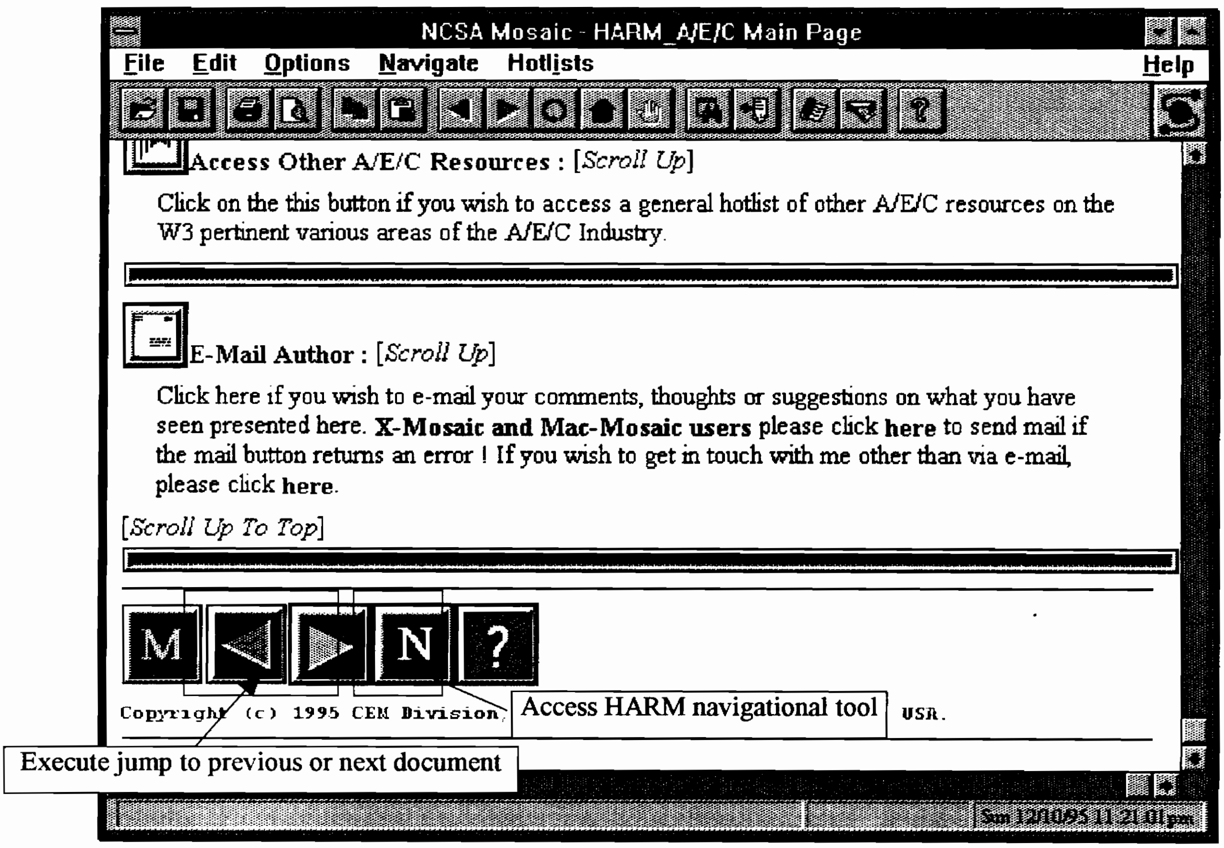

Figure 4.9 : The universal menu bar with the navigational buttons

Using the navigational tool in HARM the user can easily execute a desired jump to a given archive in just two steps as shown in figure 4.9. Within HARM, the navigational tool can be accessed by clicking on the navigate button of the universal menu bar shown in figure 4.9. The menu bar also provides a pair of buttons for moving forward and backward within the submodule of HARM being viewed.

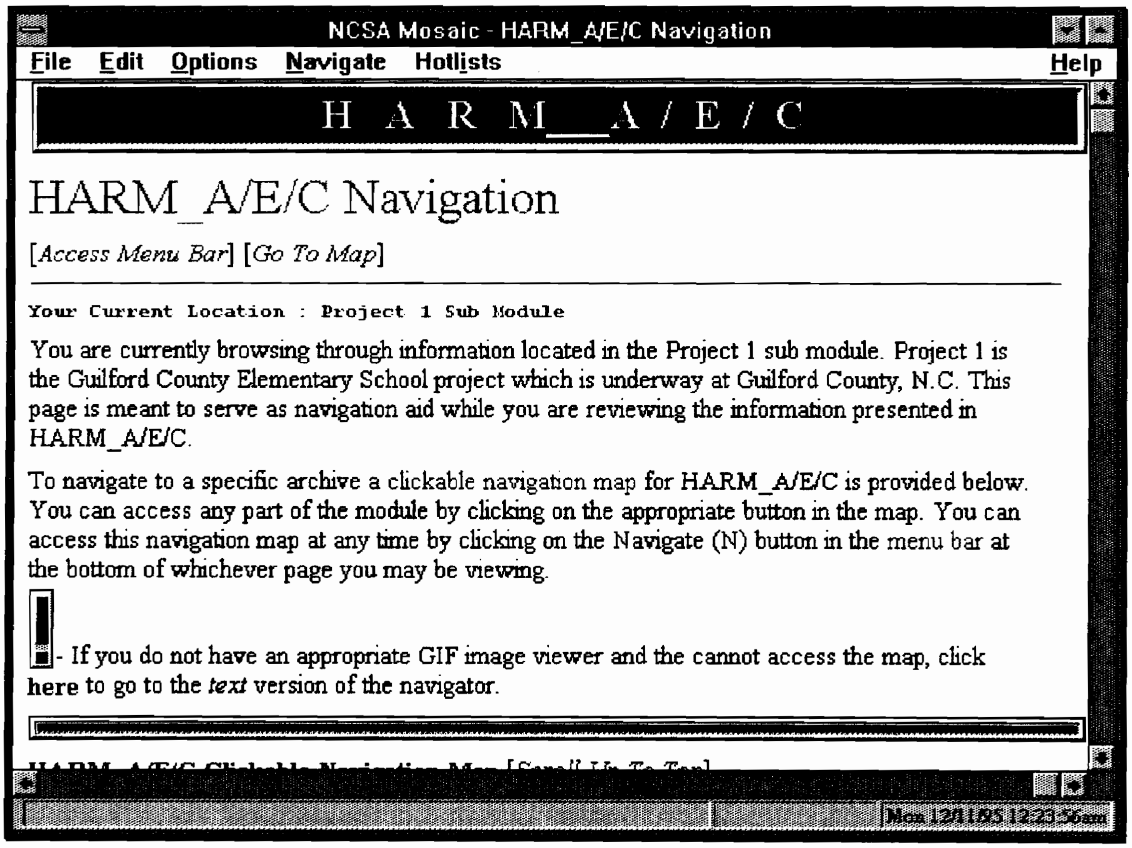

Figure 4.10 : The HARM navigation tool

Figure 4.10 above shows the navigational tool that is accessed by clicking on the navigate button shown in figure 4.9. The tool is a basically a hypermedia document that provides a clickable image of how information is organized in the module. This is called a navigation map. As seen above, the document also informs the user of his or her current location in the module. Clicking on the appropriate hyperlinked text allows the user to access the map shown in figure 4.11. Though most hypermedia browsers do support clickable image addressing, to allow for the contingency that the browser being used may not have this feature a textual version of the navigator is also provided here.

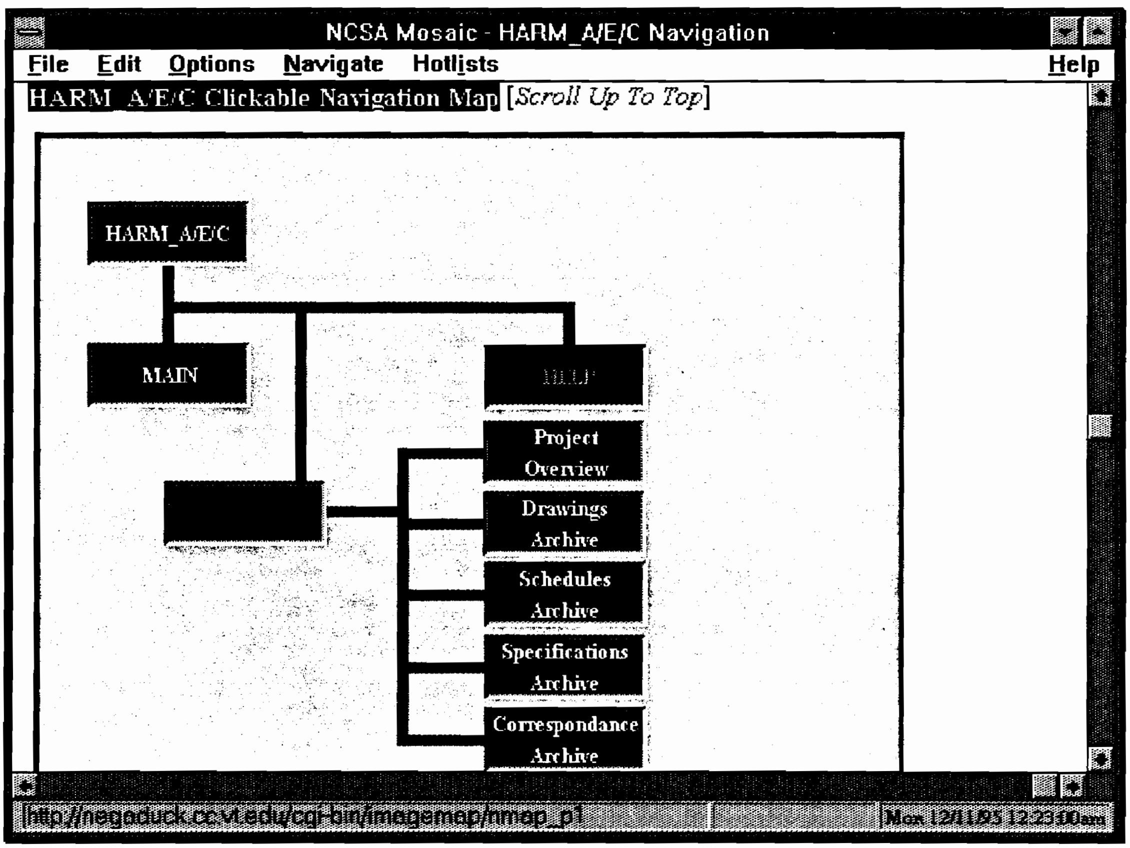

Figure 4.11: The clickable navigation map

The figure above shows the clickable navigation map. The user can access any desired archive or other location shown in the map as buttons. A description of the information available under each of these archives was outlined in section 4.1.2. As illustrated in the example presented in figure 4.7, the navigation tool allows the user to return to the opening screen of HARM, also referred to as the main page.

The navigational feature of the module totally obviates the need for remembering the sequence in which information was accessed, or even from what location in the module the current information is being viewed. This frees the user to take full advantage of the three dimensional nature of accessing information as discussed in the previous chapter, without any loss of orientation.

4.5 Use and Administration

In the first chapter of this thesis it was explained that the design of the hypermedia module would be directed primarily towards facilitating the distribution of information to organizations concerned with developing the physical construction product, in other words the contractor. It follows, therefore, that the set up and administration of the system will be carried out by the prime contractor. It is important to bear in mind that the system will work in a reliable fashion only if the organizations deploying the system understand their administrative responsibility, and decide at the outset how much of the information they wish to share with other parties on the project.

4.5.1 Use of the System

The use of system is a separate issue from its administration and refers to who will utilize the system after it has been set up, and which features they are likely to use most frequently. In this regard, the system should be useful to all parties and persons involved on a given project, who will make use of HARM A/E/C as an information service.

The features available for use have been outlined earlier in this chapter. At the most basic level the convenience of point and click access to various types of information will serve to keep users abreast of developments or aspects of the project with which they may be concerned. A greater level of use will involve taking full advantage of the Internet tools that are available for use from within the module itself.

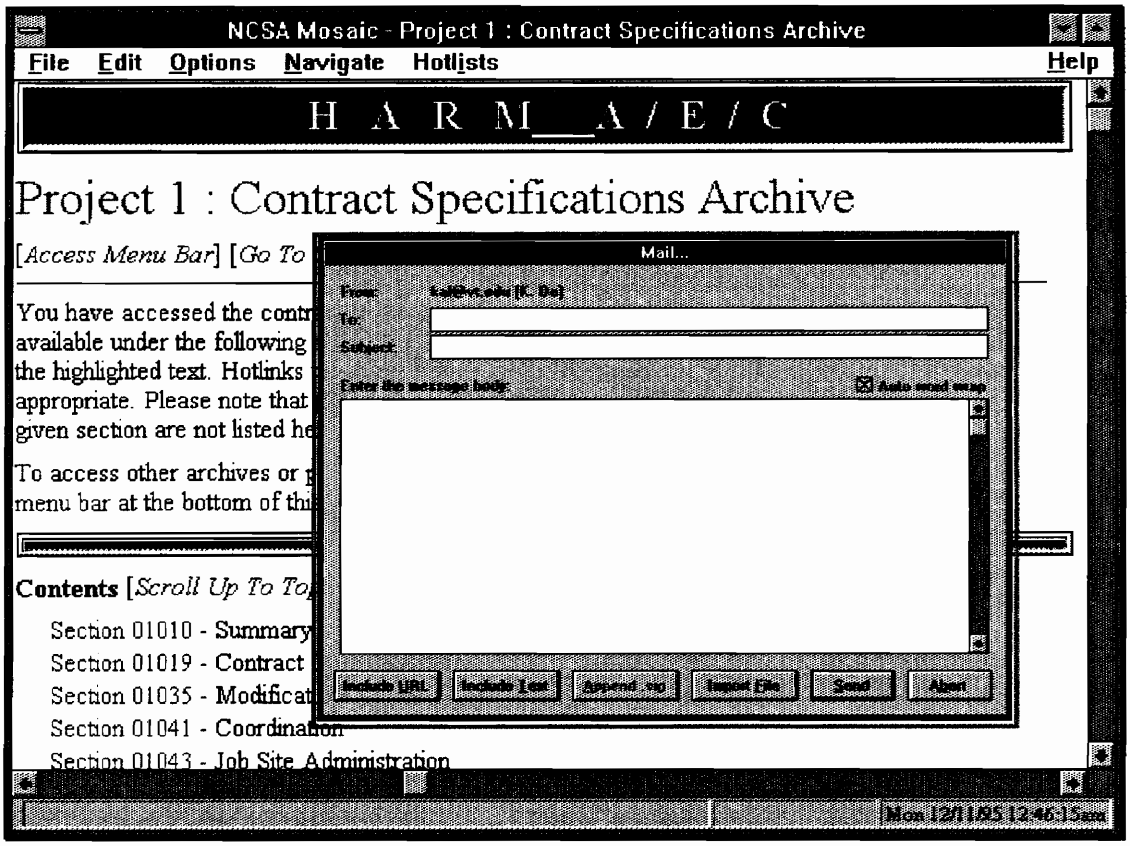

Figure 4.12 : Launching an e-mail session from within HARM

To illustrate the point, figure 4.12 shows an electronic mail session being launched from within the module by clicking on the button for sending mail. Similar functionality exists for performing file transfers to and from remote sites connected to the Internet. Thus, while reviewing information, users can communicate with one another, attach the text of the document being viewed, send in or download updated data files, all without having to exit from the module.

4.5.2 Administration of the System

The administration of the system refers to how the system will be maintained and updated, and who will be responsible for the same. A conceptual depiction of how the system maybe deployed is provided in figure 4.13. The figure shows the connectivity achieved by deploying HARM on the Internet, between three major parties, the contractor, owner, and the designer, and two of the contractors at remote sites A and B. The HARM set up and configuration described earlier allow two possible administrative modes, the centrally administered mode , and the shared and distributed mode. These are discussed below.

Centrally Administered Mode In this mode of administering the system, the HARM module, as well as all project information will be maintained on the main HTTPd server shown in the figure. The responsibility of maintaining the HTTPd *server*, and the entire module will rest with the contractor. Other parties on the project will need to have at least one of the computers on their local area network (LAN) or stand alone workstations as shown in figure 4.13, running Mosaic. This computer will serve as a World Wide Web *client.* Any of these parties may, however, provide information updates for archives relevant to their role in the project.

This is achieved by providing a special directory for incoming file transfers (shown as the FTP directory in figure 4.14), which acts analogous to an in-box. The computer system administrator of the Web server, usually referred to as a Webmaster, at the contractor's home office would then be in charge of relocating the data files to their appropriate place in the directory structure shown in the figure.

Shared and Distributed Mode This mode of administration entails a more active participation in the maintenance of the module from other parties in the project. Here again, the HARM module itself will be maintained on the contractors' home office HTTPd server. The key difference here is that other parties on the project will have the responsibility of providing information *directly*, on parts of the project relevant to their involvement.

This maybe achieved in either of two ways. As shown in figure 4.14, an external party such as the designer may be given *write permissions* to the directory on the server that houses the drawings archive for a given project. The term write permissions refers to the ability of the user to access and directly transfer files to and from a directory, on to the disk space of a computer. In other words, to be able to write information onto a memory location on the computer (Hahn, 1993). Information in such an archive would then be jointly maintained by the contractor's office and the relevant participant in the project.

Alternatively, in the above example the designer may choose to maintain his own HTTPd server at his home office. In such a case, hyperlinks pointing to the relevant information on the remote server are simply built into the respective document in the HARM main module. As explained earlier since the process of retrieving information from a remote location is completely transparent and the due to the seamless characteristic of the interface, the actual physical location of the information makes no difference to the look and feel or use of the module.

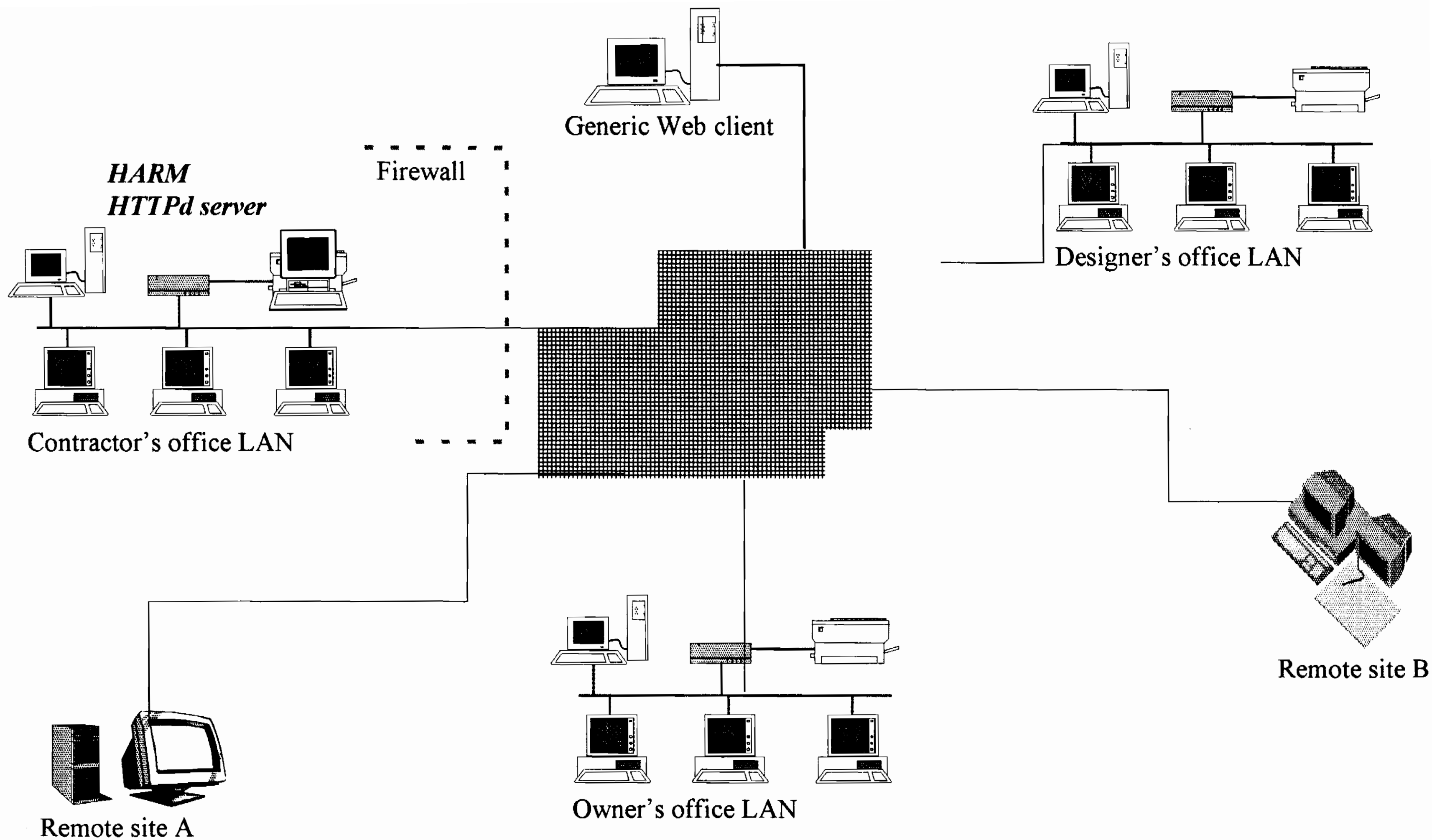

Figure 4.13 : Conceptual depiction of deployment and usage of HARM

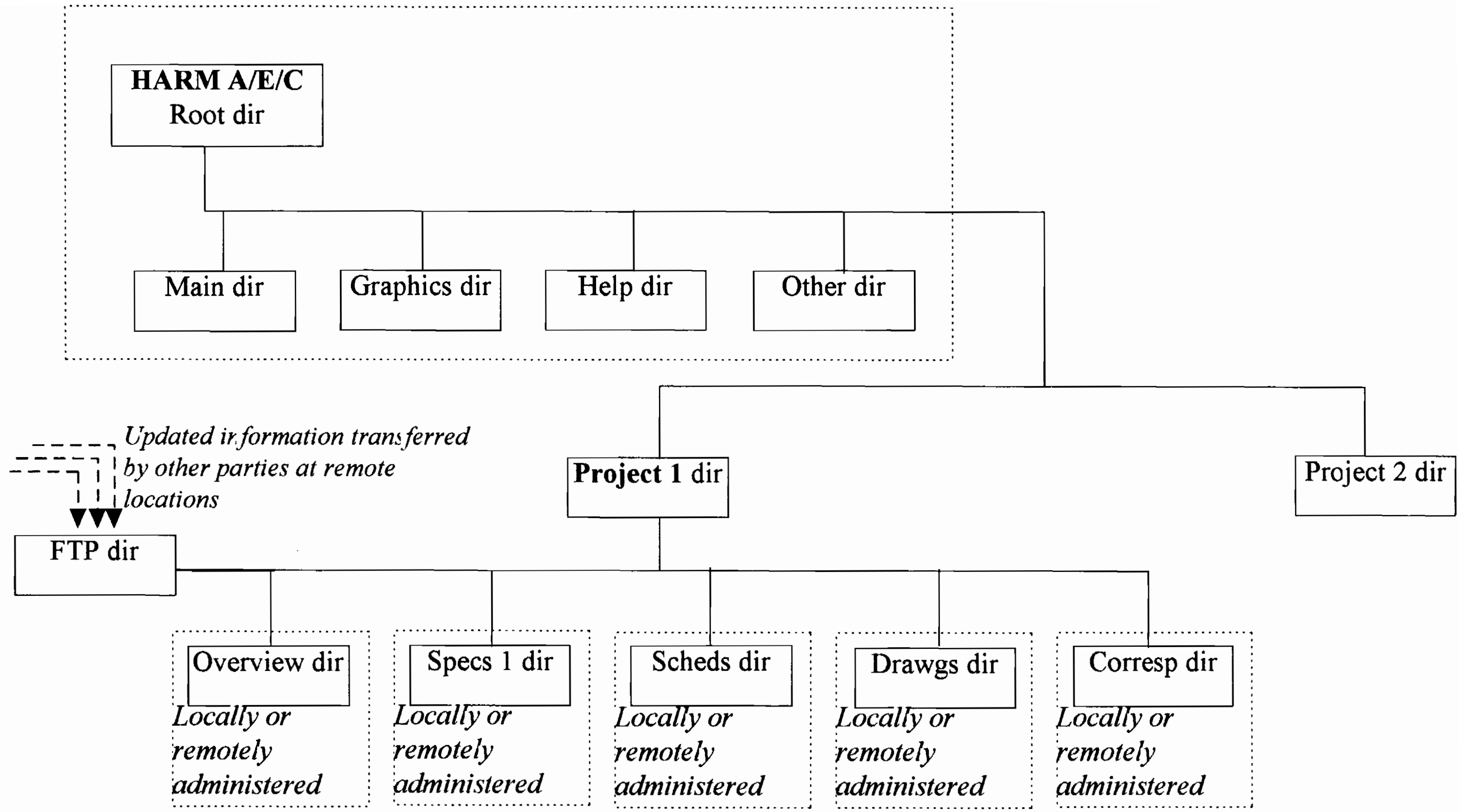

Figure 4.14 : A depiction of procedures for central and distributed administrative modes

4.6 Security Concerns

With the growth in use of the Internet for commercial and industrial purposes, there is some concern about the relative lack of security in transmitting sensitive information on the Internet (Allen, 1994; US DOCS, 1994). This section presents a brief discussion on the main security related concerns in using a hypermedia information system such as HARM on the Internet, and more specifically the World Wide Web (WWW).

There are certain valid security concerns with respect to using the WWW for information transmittal as discussed in this thesis. These are summarized below[2];

- As such, computers connected to the Internet have inadequate defense mechanisms to protect them from unwarranted access by unauthorized users. The computer programs that enable machines to be set up as information servers are complex, and tend to contain certain errors or bugs. Such bugs that continue to exist in most server software create weak spots in server software such as NCSA Mosaic HTTPd, that make them vulnerable to unauthorized persons who may seek to break into the system.

---

[2] Source: "Security issues for the WWW, on-line documentation at URL http://www.swcp.com/~mccurley/danger/generic.html

- Base protocols used by the Internet such as TCP/IP were designed to encourage cooperative usage of Internet resources, rather than to guard against adversarial security violations by unauthorized users. Thus historical restrictions on security technology have inhibited the improvement of the situation by preventing deployment of more secure systems.

- Web browsers retrieve data from the Internet, which is then acted upon by the users' local machine. In other words, data that is foreign to the local computer is used as program input which makes the former very susceptible to importing destructive computer viruses.

- Use of a Web browser for remote information tends to reveal information about the users' local machine such as, the IP (Internet Protocol) address of the machine, the users' name and e-mail address, all which may be used by a potential unauthorized person to break into the former's computer or system.

There are two aspects to security that are particularly relevant to the use of hypermedia computing on the Internet as represented by HARM. The first has to do with the integrity of data files that will be transferred over the Internet using Internet tools such as FTP and e-mail. The second aspect relates to providing restricted access to all or part of the HARM module, and the information it provides access to. It is possible to build sophisticated security measures through specially written computer programs for that purpose. As mentioned at the outset of this thesis, although the latter is beyond the scope of this thesis, a brief discussion on these two aspects is provided here.

4.6.1 Data Transmission

Developmental work on the TCP/IP suite of protocols by the Internet Task Force have ensured that the integrity of data transmitted across the Internet is preserved (Barry, 1994). This means that data files that are transferred across the Internet are virtually guaranteed to reach their destination without being damaged or corrupted by the process of being transferred itself. While there are built in security measures in this suite of protocols to prevent tampering with the data during transmission, there is no official administrative body that oversees such data transactions, which therefore remain somewhat susceptible to security violations.

Internet Task Force officials have pointed out the fact that given the open architecture on which the Internet is based, the above is an unavoidable fallout (US DOCS, 1994). Also given the fact, that this open architecture has been largely instrumental in the virtual platform independence enjoyed by Internet users, the downside appears to be acceptable. The Computer Emergency Response Team (CERT), has also observed that deliberate unlawful security violations are in the vast majority of cases the work of eccentric individuals who pursue such activity as an intellectual challenge, rather than with the intent to commit industrial espionage.

The issue of secure data transmission across the Internet is address to some extent by message encryption[3]. Encryption works by encoding the text of a message with a key. In traditional encryption systems, the same key is utilized for both encoding and decoding, whereas in the new public key or asymmetric encryption systems, keys are employed in pairs, one being used for encoding and another for decoding. In this system everyone owns a unique pair of keys. The key for encoding messages, called the public key, is widely distributed and used. The second key, called the private key, is known only to its owner, who uses it to decrypt his or her incoming messages (Schneier, 1994; Stein, 1995).

Thus with this system, a person sending a message to a second person encrypts the message with that latter's public key. The message can only be decrypted by the owner of the private key, guarding it from interception. Since organizations using the Internet for commercial purposes have a critical need for secure transmission on the Web, there is

[3] Source: "The World Wide Web Security FAQ", Lincoln D. Stein, on-line documentation at URL http://www-genome.wi.mit.edu/WWW/faqs/www-security-faq.html

very active interest in developing schemes for encrypting the data that passes between browser and server.

4.6.2 Guarding Access to the Module

It would be possible to restrict access to the module by employing a specially written Unix shell script for the purpose. Access to the module or information housed in certain designated directories would then require user authentication by asking for a password. The developers of NCSA Mosaic HTTPd have pointed out, however, that such a procedure would provide no more and no less security than is available to electronic mail messages sent across the Internet (Mosaic Technical Documentation, 1995). Security measures of a higher level would most certainly entail the services of a professional programmer. In turn, it would also mean providing each group of users with a unique identification and password, and much additional effort in administering and tracking secure usage of the system.

In essence, guarding access to HARM or a like system would entail restricting access to the HTTPd server on which the information would be located. A few of the essential considerations that would require taking into consideration are summarized below (Stein, 1995).

- *Operating System* To some extent the degree of security that can be achieved for the server depends on the operating system employed by the machine. In general more powerful operating systems such as Unix, are more prone to security breaches. In Unix systems, the large number of built-in servers, services, scripting languages, and interpreters, are particularly vulnerable to attack as there are several portals of entry for unauthorized persons to take advantage of. Less capable systems, such as Macintosh systems and Microsoft Windows based machines, are less easily exploited. It maybe noted that to an extent also, the security of the system will depend on the server software employed[4]. Here again, the vulnerability to security lapses increase as the number of features offered by the server software increases.

- *Security Precautions* There are certain simple but effective security precautions that can be taken for Web servers running on Unix systems. These include limiting the number of login accounts available on the server, and regularly removing the accounts of inactive users (Hahn, 1993).

- *File Permissions* To maximize security, a well defined and regulated policy for both the document root (where the HTML documents are stored on the server), and the server root (where log and configuration files are kept) should be

[4] Source: NCSA Mosaic on-line documentation at URL http://ncsa.uiuc.edu/mosaic/

formulated. This helps in keeping the server from intrusions by either local or remote users. A possible way to accomplish this is to create a user account for the Web administrator/Webmaster and a "WWW" group for all other users on the system, who will be authoring HTML documents. On Unix systems this can be achieved by editing the /etc/passwd file and making the server root the home directory for the user. The server root should be set up so that only the www user can write to the configuration and log directories and to their contents. These directories should preferably not be readable outside the readable by the www group. NCSA HTTPd enbles the system administrator to restrict access to parts of the directory structure that houses the HTML documents to Internet browsers with

certain IP addresses, or to remote users who can provide a correct password. It maybe noted that this feature enables the shared and distributed mode for administering HARM as discussed earlier in section 4.5.2 of this chapter.

- *Access Restrictions* As mentioned earlier, the data files archived under HARM are located on the server machine. Access to these files can be restricted in either of the following ways, each of which has been mentioned earlier;

Restriction by IP address, subnet, or domain, by which individual documents or directories are protected in such a way that only browsers connecting from certain IP addresses can gain access to them for making changes.

Restriction by user name and password, whereby documents or directories are protected so that the remote user has to provide a name and password in order to gain access.

Encryption using public key cryptography as discussed earlier in this chapter.

Fianlly, it is worth noting, that in the context of HARM, the construction project data that is proposed for transmission across the Internet, may not always qualify as being sensitive data. Indeed a number of construction industry personnel, including construction managers, designers, and architects who were asked to identify types of project data they felt would qualify as being sensitive, indicated that mainstream project data as may be accessible through HARM did not merit elaborate security measures.

The issue of security will probably be best addressed on a case by case basis, depending on the policies of the particular organization employing this technology, and the degree of security concerns associated with the project on hand.

This chapter has presented the proof of concept module HARM, that was built to demonstrate the viability of utilizing a hypermedia based storage and retrieval system for the distribution of A/E/C project information. The concluding chapter of the thesis that now follows brings the discussion to an end after a brief outline of the anticipated future of the Interrnt, and its relevance to this research.

5

Conclusion

This chapter of the thesis provides the concluding notes regarding the proof of concept module itself, and the technology on which it is based. The future direction of the Internet with respect to the World Wide Web and hypermedia computing is also discussed. The scope for future work pertaining to the module is mentioned, and an outline of the scope for further research in the area of World Wide Web and hypermedia computing on the Internet closes the discussion.

5.1 Concluding Notes

As pointed out earlier, the hypermedia module that serves as a proof of concept for this thesis provides access to information on a single project, rather than multiple projects as proposed by this thesis. However, the point to be taken into consideration is that the module has been designed to serve as a front end for a indefinite number of projects. The limitation evident in the context of this thesis reflects solely on time constraints in performing the research than on the functionality of the module. A functional limitation of the module is, however, imposed by the indirect procedure for updating information in the module.

Although the development of hypermedia computing on the Internet has advanced rapidly in the past few years, and new improvements are being made at a very fast pace, there remain certain limitations to using the technology discussed in this thesis. These include the relatively slow data transmission speeds for large file sizes such as large CAD files, or when using connections to the Internet such as SLIP.

However, as evident from the figures discussed in chapter 3, the use and development of the Internet is growing exponentially and improvements in these areas should be forthcoming in the near future. Some of the expected future developments are outlined in the following section.

5.2 The Future of Internet

There are a wide range of views regarding the developments expected to take place on the Internet in the future, and the direction that the technology will take. Plans have been described for making the Internet available through public computer stations, or kiosks, which may make accessing the Internet as easy as making a telephone call from a phone booth.. It is expected that television may soon be linked into the Internet, greatly expanding the possibilities for information dissemination and entertainment[1]. Video teleconferencing and interactive television are just two of the features that are expected to appear in force at that time. Virtual reality, and programs which make possible novelties such as lifelike flight simulations, is also coming to the Internet -- for advanced entertainment and for instruction and training. Although a comprehensive discussion on this topic is not possible here, a note on some of the key issues concerning the future of the Internet is provided in the following sections.

[1] Source: "The future of the Internet", on-line documentation at URL http://www.theta.com/freedommag/future.htm

5.2.1 Directions and Developments

The arrival of easy to use and versatile Web publishing software has helped accelerate commercial publishing on the Web. Commercial use of the Web has also grown exponentially due to the lifting of previous restrictions. The original design goal of W3 to encourage collaborative work continues, allowing geographically dispersed people to work together in groups, sharing knowledge, modifying, annotating, and contributing as well as reading[2].

The development of object oriented databases for use in conjunction with the WWW also continues. Object-oriented features in the web will allow users to manipulate objects other than documents. Examples of this may be scientific data with which multiple users define, and interact with, objects of all conceivable kinds. The arrival of tools that allow interfacing of the Web with SQL (Structured Query Language) databases, has brought further interactivity to information systems that are being developed for the Internet.

Finally, significant developments are continuing in the tools becoming available on the Web. These include increasingly sophisticated search engines and web-wandering robots which build indexes of information retrieved.

---

[2] Source: "W3 seminar : Future directions", on-line documentation at URL http://www.w3.org/hypertext/WWW/Talks/CompSem93/Future.html

5.2.2 Enhancements in Protocols

The HTTP protocol was enhanced in early 1993 to allow a number of the features discussed in the previous section to be incorporated. This enhancement also allowed multimedia with format negotiation, authentication, more information about the client being available to the server, more feedback to information providers, and data entry by users. Other extensions which are being refined include markup languages that permit authoring of structured multimedia documents, and a forms language for complex on-line queries. Attempts are also underway to refine standards in file formats and communication protocols by a number of W3 development teams in conjunction with the Internet Engineering Task Force.

5.2.3 Connectivity and Bandwidth

The rapid growth in the use of the Internet, and in particular the exponential rise in the popularity and use of the WWW has caused the computer and telecommunications industry to start offering network connectivity at affordable costs[3]. This further spurred growth in the use of this technology for information dissemination, storage and retrieval,

---

[3] Source: "Editor's notes", on- line documentation at URL http://www.boardwatch.com/mag/95/jun/bwm1.htm

bringing about fears that the Internet will be overburdened with the amount of network traffic it has to handle. ISDN (Integrated Systems Digital Network), lines now provide SLIP/PPP links to the home - using existing copper in the ground with transfer rates of 115 kpbs. Major corporations such as Sprint, MCI, AT&T, IBM, Prodigy, CompuServe, Ameritech, Pacific Bell, have voiced current plans to offer network connections. Each would be expected to contribute additional backbone capacity, and MCI and Sprint backbones are already at the 155 Mbps level and are expected to be approaching 500 Mbps in the near future. However, there is some room for doubt regarding how this additional backbone capacity would be integrated with the existing set up, and how it may affect individual users on the Internet.

5.3 Further Research

The scope for further research is outlined in this section with respect to the proof of concept module, and more generally with respect to hypermedia computing and the World Wide Web.

There is scope for further research in four main areas of this thesis. These relate to further investigation on data types suitable for distribution, customization of the module, deployment and administration of the set up, and concerns relating to security.

Further research on which types of A/E/C project data most frequently require distribution would help increase the benefits realized from using a system such as HARM. Since the concept of using the Internet for data distribution is still in its infancy in the A/E/C industry, there is little empirical information in this area. Further research should preferably involve actual feedback from the industry, and attempt also to more closely define the needs in data distribution.

As stated earlier, there is considerable scope for customizing the module. However, further research in this area should preferably be aimed at investigating what features could be added without affecting the virtual independence from commercial software products as seen in the version of HARM discussed in this thesis. This would be in keeping with the principle adhered to in this thesis, of employing software that has become the de facto standard by virtue of its wide availability and use on the Internet. There is scope for further research in optimizing the set up of the module. Such research would preferably look in to issues relating to the deployment and administration of the module in actual industry settings. This would include but not be limited to defining roles and responsibilities for administering the module, and monitoring legal implications thereof, as digital communication of data replace paper based communication in some areas of an on going project.

Finally, research regarding the security concerns discussed earlier may attempt to better clarify which kinds of data require careful safeguards during transmission across public networks, and ascertain how important an issue this is. Suitable technology to address this issue could then be looked into and incorporated into the set up and use of the module.

With respect to the World Wide Web and hypermedia computing on the Internet and the use of resources on the World Wide Web, in the light of the discussion in the previous section, there are several areas that merit further research.

Future research could examine and demonstrate the applicability of World Wide Web tools that enable interfacing databases containing large volumes of project information with a hypermedia front end. This may include demonstrating how a forms interface on the Web maybe employed to enable direct data entry by the user, for providing information updates or feedback on a given issue from a remote location.

The utility of more sophisticated tools such as Web search engines or robots could be examined in the retrieval of specific A/E/C related information over the Internet.

With improvements in data transmission times over the Internet, future work may also lay stress on the real time capture of multimedia data from the site, and the possibility of viewing the same through a hypermedia front end on the Web. This may also include looking into the possibilities for incorporating video conferencing technology on the Web, to aid in communication between geographically dispersed participants on a project.

References

"Application of Information Technology in Construction." (1993). *Proceedings of the Fifth International Conference in Computing in Civil and Building Engineering. V ICCBE, ASCE.* New York, NY, 1571-1580.

"Applications of Information Technology in Construction." (1991). *Institution of Civil Engineers*, London.

Bates, B.J. (1989). Evolving to an information society : Issues and problems. In J. Salvaggio (Ed.). *The information society: Economic, social and structural issues.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Baxendale, T. (1994). Information systems development for contracting. *Automation in Construction., 2*(4), 291-302.

Bessom, R. M. (1975). Marketing's role in construction firms. *Journal of the Construction Division., ASCE, 101*(C03), 647-659.

Carson, S. C. & Rada, R. (1994) The new media. *Communications of the ACM, 37*(9), 23-26.

Chaffee, S. H. (1991). *Communication concepts : Explication.* Sage Publications, Inc. Newbury Park, CA.

Choi, K. C., & Ibbs., C.W. (1990). Costs and benefits of computerization in design and construction. *Journal of Management in Engineering., ASCE, 4*(1).

Choi, K. C., & Ibbs., C.W. (1990). CAD/CAE in construction : trends, problems, and needs. *Journal of Management in Engineering., ASCE, 6*(4), 394-415.

Croft, W.B. (1990) Hypertext and information retrieval : What are the fundamental concepts ?. *Proceedings of the European Conference on Hypertext, INRIA, France.*

Gardner, P. (1993). Information technology : Panacea or Poison ? *Proceedings of the Fifth International Conference on Computing in Civil and Building Engineering V-CCCBE Anaheim, CA*, 2, 1358-1365.

Davis, W. S., & McCormack, A. (1979). *The information age*. Massachusetts: Addison-Wesley

Dillon, M., Jul, E., Burge, M., & Hickey, C. (1993). Assessing information on the Internet. *OCLC Online Computer Library Center, Inc.*, , Dublin, Ohio: Office of Research.

Dizard, W.P. (1982). *The coming information age. An overview of technology, economics and politics*. New York: Longman.

"Filing of construction photos linked with database." (1993). *Proceedings of the Fifth International Conference in Computing in Civil and Building Engineering. V ICCBE, ASCE. New York, NY*, 718-721.

Finn, R. (1995). Superhighway or supercollider -- finding useful Internet information. *Scientist, 9*(7), 17-18.

Gerwick, B.C. & Woolery, J.C. (1983). *Construction and engineering marketing for major project services*. New York: John Wiley & Sons, Inc.

Gilster, P. (1993). *The Internet navigator*. New York: John Wiley & Sons, Inc..

Glushko, R. J. (1990) Hypertext : prospects and problems for crew system design. Crew System Ergonomics Information Analysis Center, CSERIAC-SOAR-90-1, *Search Technology, Inc.*, 4-10.

______ (1990) Hypertext : prospects and problems for crew system design. Crew System Ergonomics Information Analysis Center, CSERIAC-SOAR-90-1, *Search Technology, Inc.*, 79-88.

Grieshaber, D. (1994). *Documentation storage and retrieval system for construction sites*. Unpublished project and report submitted to faculty of the Department of Civil Engineering, Virginia Polytechnic Institute & State University.

Hahn, H. (1993). *A student's guide to UNIX*. New York: McGraw-Hill, Inc..

Hahn, H. & Stout, R. (1994). *The Internet complete reference*. Berkley, C.A: Osbourne McGraw-Hill

Hemmet, F. J. (1993) Application of Information Technology in Construction. *Proceedings of the Fifth International Conference in Computing in Civil and Building Engineering. V ICCBE, ASCE.. New York, NY*, 1571-1580.

Hodges, M. E. & Sasnett, R. M. (1993). *Multimedia computing: Case studies from MIT Project Athena*. New York: Addison Wesley Publishing Company.

Howard, H.C., Levitt, R. E., Paulson, B.C., Pohl, J.G., and Tatum, C.B. (1989). Computer Integration : Reducing Fragmentation in the AEC Industry. *Journal of Computing in Civil Engineering. ASCE, 3*(1), 18-31.

"Information integrated construction." (1993). *Proceedings of the Fifth International Conference in Computing in Civil and Building Engineering. V ICCBE, ASCE.. New York, NY,* 145-152.

"Infra technologies, Tools for innovation." (1993) *Civil Engineering, 63*(11), 68-69.

"Intelligent document generation system for construction planning." (1992*). Proceedings of. Conference on Artificial Intelligence Applications, IEEE. IEEE Service Center Piscataway, N.J*, 14-19.

"Intelligent interface for construction information." (1993*). Proceedings of the Fifth International Conference in Computing in Civil and Building Engineering. V ICCBE, ASCE. New York, NY,* 1366-1369.

International Organization for Standardization (1986). *Standard Generalized Markup Language, ISO* 8879-1986.

Jonassen, D.H.. (1989). *Hypertext / Hypermedia*. Englewood Cliffs, N.J. :Educational Technology.

Katterman, L. (1995). Using the Internet for technology transfer. *Scientist, 9*(12), 10.

Keyes. J. (1994). *The McGraw Hill multimedia handbook*. New York: McGraw Hill, Inc..

Kindborg, M. & Pettersson, R. (1991). Classification of Navigational Principles in Mulitmedia Systems. Multimedia Information. *Proceedings of a Conference at Churchill College, Cambridge, U.K. Bowker Saur,* London, U.K.

Lane, E. & Summerhill, C. (1993). *Internet primer for information professionals: a basic guide to Internet networking technology*. Westport, CT: Meckler Publishing..

LaQuey, T. & Ryer J. C. (1993*). The Internet companion: A beginner's guide to global networking*. Reading, Mass: Addison-Wesley Publishing Company.

Lattor, M. (1992) Internet growth (1981-1991); RFC 1296, Network Working Group Request for Comments. *Network Information Systems Center, SRI International, Menlo Park, California*.

Lucarella, D. (1990). A model for hypertext based information retrieval. *Proceedings of the European Conference on Hypertext, INRIA, France*,.

Lynch, D.C. & Rose, M.T. (1993). *The Internet system handbook*. Reading, Mass: Addison-Wesley Publishing Company..

McCaffer, R. & Baldwin, A. N. (1984*). Estimating and tendering for civil engineering works*. Great Britain: Granada.

Meltsner, K.J. (1995). Understanding the Internet: A guide for materials scientists and engineers. *Journal of the Minerals Metals and Materials Society, 47*(4), 9-13.

Michel, J. (1991). Multimedia for the 1990s: Technological trends and main future applications. Multimedia Information. *Proceedings of a Conference at Churchill College, Cambridge, U.K. Bowker Saur,* London, U.K.

Miyatake, Y., and Kangari, R.. (1993). Experiencing Computer Integrated Construction. *Journal of Construction Engineering and Management. ASCE, 119*(2), 307-322.

More construction for money. (1983). Summary Report of the Construction Industry Cost Effectiveness Project, *The Business Roundtable (BRT)*, New York, N.Y.

Naeim, F., Benuska, K. Martin, J. A. (1993). A multimedia information system for buildings and campuses. *Proceedings of the Fifth International Conference on*

Computing in Civil and Building Engineering (V-CCCBE), Anaheim, CA, 2, 1358-1365.

Nielsen, J. (1990). *Hypertext and Hypermedia*. Boston : Academic Press.

Office of Technology Assessment. Congress of the United States. (1987). International competition in services banking building software know-how. *OTA -ITE-328*, Washington, D.C: U.S. G.P.O.

Parfitt, M.K., Syal, M.G., Khalvati, M., Bhatia, S. (1993). Computer Integrated Design Drawing and Construction Project Plans. *Journal of Construction Engineering and Management. ASCE, 119*(4), 729-741.

Perkowski, J.C. (1988). Technical Trends in the E & C Business : The Next 10 Years. *Journal of Construction Engineering and Management. ASCE, 114*(4), 565-576.

Robinson, J. H. (1991). *3D computer modeling and animation for marketing project services, construction in the 21st Century.*

Rocha, L. A. & Morad, A. A. (1993) Intelligent systems for information retrieval in construction. *Proceedings of the Fifth International. Conference on Computing Civil Building, Engineering (V ICCCBE), ASCE. , Anaheim, CA,* 1594-1599.

Sanvido, V., Grobler, F., Parfitt, K., Guvenis, M., & Coyle, M. (1992). Critical success factors for construction projects. *Journal of Construction Engineering and Management., ASCE, 118*(1), 94-111.

Schneier, B. (1994). *Applied cryptography*. New York:. John Wiley & Sons.

Stein, L. D. (1995). *How to Set Up and Maintain a World Wide Web Site: The Guide for Information Providers*. New York: Addison-Wesley Publishing Co.

Stewart, A.T. (1993). Information Management in the A/E Environment. *Computing in Civil Engineering*, 372-377.

Teicholz, P. and Fischer, M. (1994). Strategy for Computer Integrated Technology. . *Journal of Construction Engineering and Management., ASCE, 120*(1), 117-129.

"The Internet -- networks joined together". (1995). *Chemical Engineering and News. 73*(13), 74-74.

Thomas, B. J. (1995). *The Internet for scientists and engineers: Online tools and resources*. Bellingham, WA: SPIE Optical Engineering Press..

Turk, C.M. (1992). Information Technology Planning for Construction Companies. *Computer Aided Engineering*, 59-67.

United States, Congress, House, the Subcommittee on Science of the Committee on Science, Space, and Technology, U.S. House of Representatives, One Hundred Third Congress. (1994). *Internet security*: US DOCS no. 103/128 Washington: U.S.G.P.O.

United States, Congress, House the Subcommittee on Science of the Committee on Science, Space, and Technology, US House of Representatives, One Hundred Third Congress. (1994). *Internet security*. US DOCS no. 103/187 Washington : U.S.G.P.O.

APPENDIX

HARM A/E/C Users Manual

This appendix is meant to serve as a users manual for the Hypermedia Archiving and Retrieval Module. The following sections outline the basic requirements and procedures for using the module.

1.0 Requirements

The following are the hardware, software and connectivity requirements for using HARM A/E/C. Since there is a considerable degree of flexibility in addressing these requirements, the sections below outline the minimum required in each case.

1.1 Hardware and Software

The basic software that will be required for using HARM is a hypermedia browser, like NCSA Mosaic. However, the user is not limited to any particular browser. The hardware system requirements for NCSA Mosaic, one among several possible hypermedia browsers, are provided below[1]. It should be noted that there are different versions of Mosaic for the three major computing platforms mentioned below. This does not affect the look and feel or usage of either the browser or this module itself.

Microsoft Windows Mosaic for the Microsoft Windows platform requires a minimum of an 80386SX-based personal computer with 4 megabytes of RAM (Random Access Memory), operating in enhanced mode. Versions of Mosaic are available for the Windows 95 and Windows NT. The latest version of Win32s with OLE software is also required for the Windows 3.x platform. It maybe noted that the latter can be obtained for free via anonymous ftp from the Mosaic for Windows WWW site. Access to this site is provided from within HARM.

Macintosh System Mosaic for the Macintosh requires a Macintosh personal computer, IISi series or better, running System 7 or later with 4 megabytes of

---

[1] Source: NCSA Mosaic On-line documentation at URL
http://www.ncsa.uiuc.edu/SDG/Software/WinMosaic/HomePage.html

RAM and a memory partition of at least 2 megabytes. In addition MacTCP 2.0.2 or later is required, (MacTCP 2.0.4 or later is recommended).

X Window System Mosaic for the X Window system can be used on most UNIX-based workstations (e.g., Sun SPARC, IBM RS/6000, SGI, DEC).

The system requirements outlined above will be sufficient to address the needs of other Internet hypermedia browsers such as Netscape, SpyGlass etc.

1.2 Viewers and Helper Applications

A hypermedia browser, by itself, will allow a user to access the textual information in the module. Depending on the browser being used the user will need to have some additional application software that will act as viewers (also known as helper applications). This will be essential for viewing file formats that are not supported by the browser program. These application programs must be located on the users' local machine or LAN. It is recommended that such helper applications be obtained for the file formats shown in table A.1, which are likely to be encountered through HARM.

Format	Description
GIF	Graphical Interchange Format
JPG	Joint Photographic Experts Group
AIF	Audio Interchange Format
MPEG	Motion Picture Experts Group
DOC	Microsoft Word Document file extension
DXF	Data Exchange Format
P3	Primavera Project Planner file extension

Table A.1 : File formats and their descriptions

It is possible to obtain free versions of helper application programs for all of the above file formats with the possible exception of the last two, which would probably require commercially available software. Hyperlinked access to such freely available software is provided from within HARM, through the on-line assistance.

1.3 Internet Connection

Access to the module will naturally require an Internet connection. At the basic level, individual users will require a personal computer or workstation, a modem with communications software, and a phone line connection. Access to the Internet will have to be obtained from an Internet service provider who maintains terminal servers that enable full access dial-up connections.

To enable interactive access to the World Wide Web and other Internet resources, the communications software must support a SLIP (Serial Line Internet Protocol) or PPP (Point to Point Protocol), which will make the users machine behave as a peer computer on the Internet. The user must dial into a computer or terminal server maintained by the access provider, running one of the above protocols to make this connection.

2.0 Accessing Information

Once an Internet connection has been achieved and a suitable hypermedia browser has been installed, the following steps are to be followed for configuring the browser, accessing the module, and accessing information through HARM.

2.1 Configuring the Browser

Most browser programs allow the user to specify and set certain preferences while using the program.. The following should be specified in the browsers' "Preferences ..." file, usually accessible from the file menu ;

- Set "Caching" to On
- Specify a path and directory for the browsers cache
- Set the number of documents to be cached to not more than 5
- Specify helper applications for the respective file formats, for instance AutoCAD as a helper application for viewing DXF file formats
- Specify the users e-mail address under "Network services"

2.2 Accessing HARM A/E/C

- Access the HARM A/E/C main page by entering the following URL (Uniform Resource Locator) at the browsers command line or "Open URL" option under the file menu;

  http://clark.ce.vt.edu/harm/harm.html

3.0 Using the Module

The module provides hyperlinked access to information contained within its own archives as well as external resources on the Internet. Hyperlinks are indicated by highlighted text (usually blue in color), buttons, or images. Clicking on a hyperlink provides access to more specific information on a particular topic or subject.

The following sections provide points to be noted on to navigating through the module, obtaining some assistance, and updating information in the module.

3.1 Navigation through HARM A/E/C

At any point within the module the user may access the universal menu bar shown in figure A.1. The menu bar allows the user to return to the main page (by clicking on the button marked 'M"), and jump forwards or backwards at a particular level, or within a certain archive.

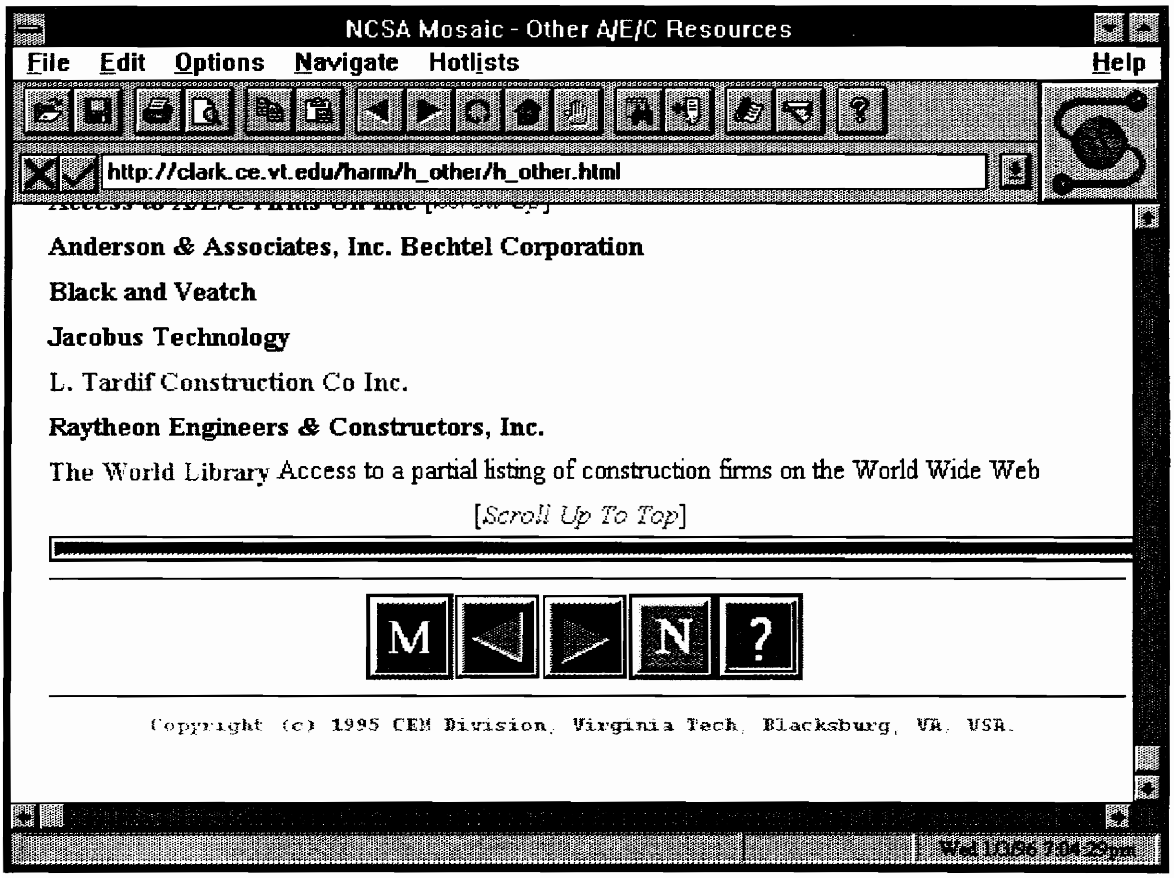

Figure A.1 The universal menu bar

Clicking on the button marked ‘N’ and ‘?’ takes the user to a navigational tool and the on-line help shown in figures A.2 and A.3 respectively. The navigational tool provides a map that shows the conceptual layout of information within the module. Clicking on a button in the map takes the user to that respective part or archive in the module.

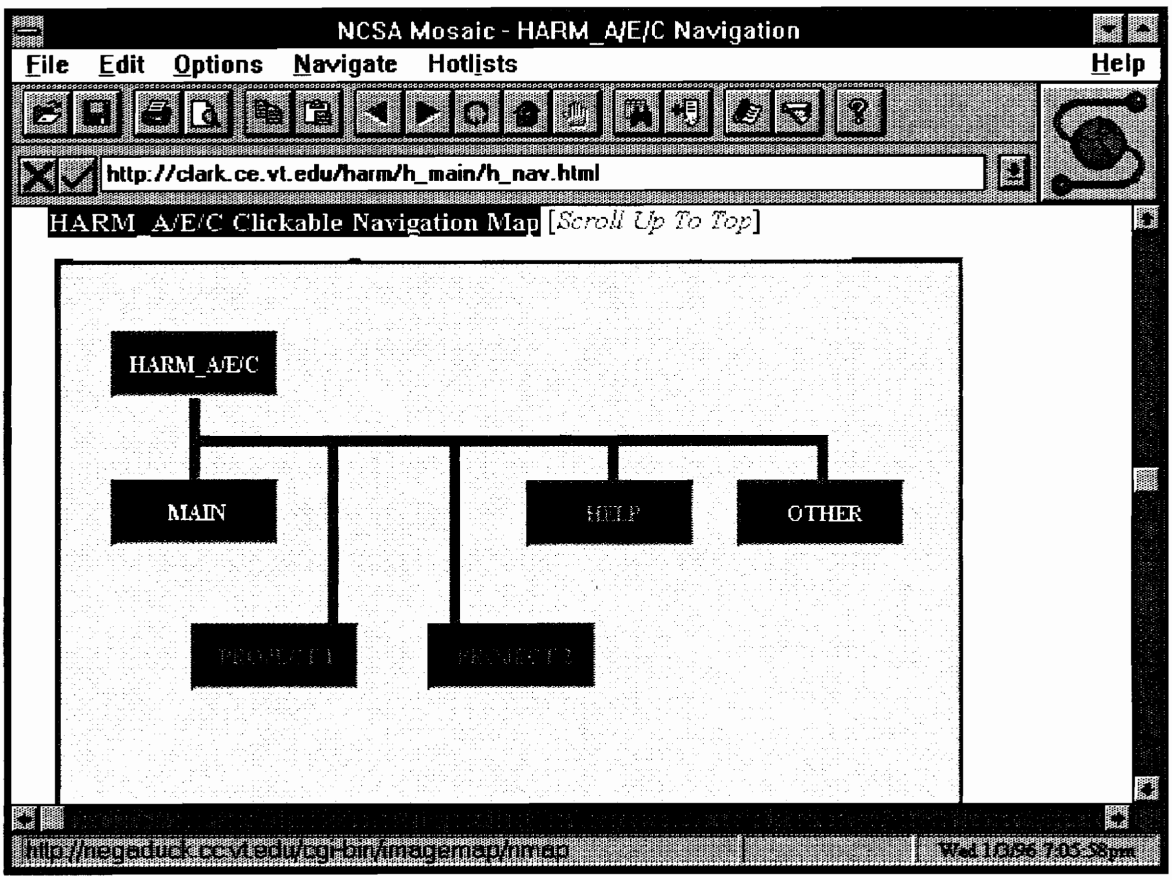

Figure A.2 The clickable navigation map

On-line assistance can also be accessed from this map by clicking on the button marked ‘Help’ in the map.

3.2 Getting Assistance

On-line assistance can be accessed as noted in section 3.1. A help document such as the one shown below provides hyperlinks to assistance on the module itself as well as on

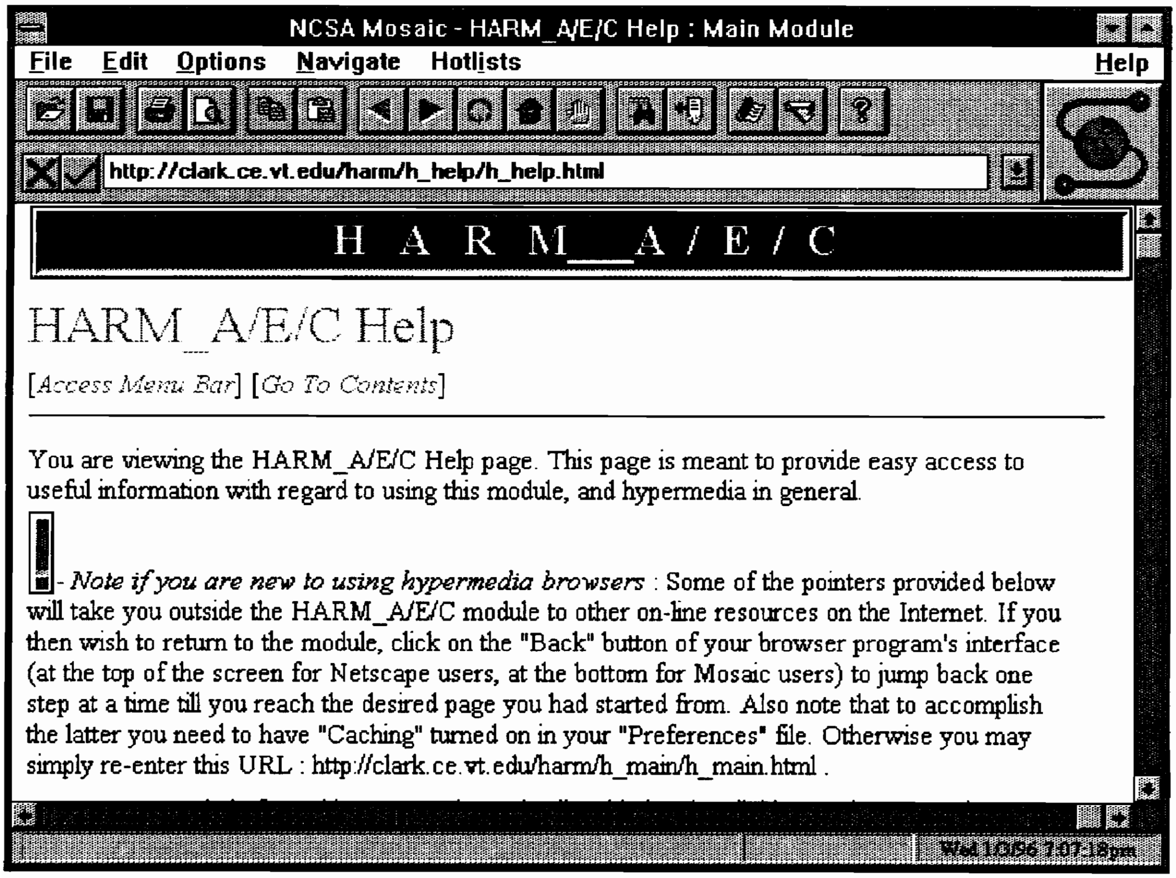

Figure A.3 On-line assistance

NCSA Mosaic, Netscape, viewers and helper applications, authoring in HTML, on the W3 Project, and on finding other information on the Web. These topics can be accessed by clicking on the appropriate hyperlink (not shown in the above figure).

3.3 Updating Information

Users will not be able to make any changes to the information viewed through HARM. A user may send in updated information on a given project, to be included in the module by using an ftp (file transfer protocol) tool to transfer files.

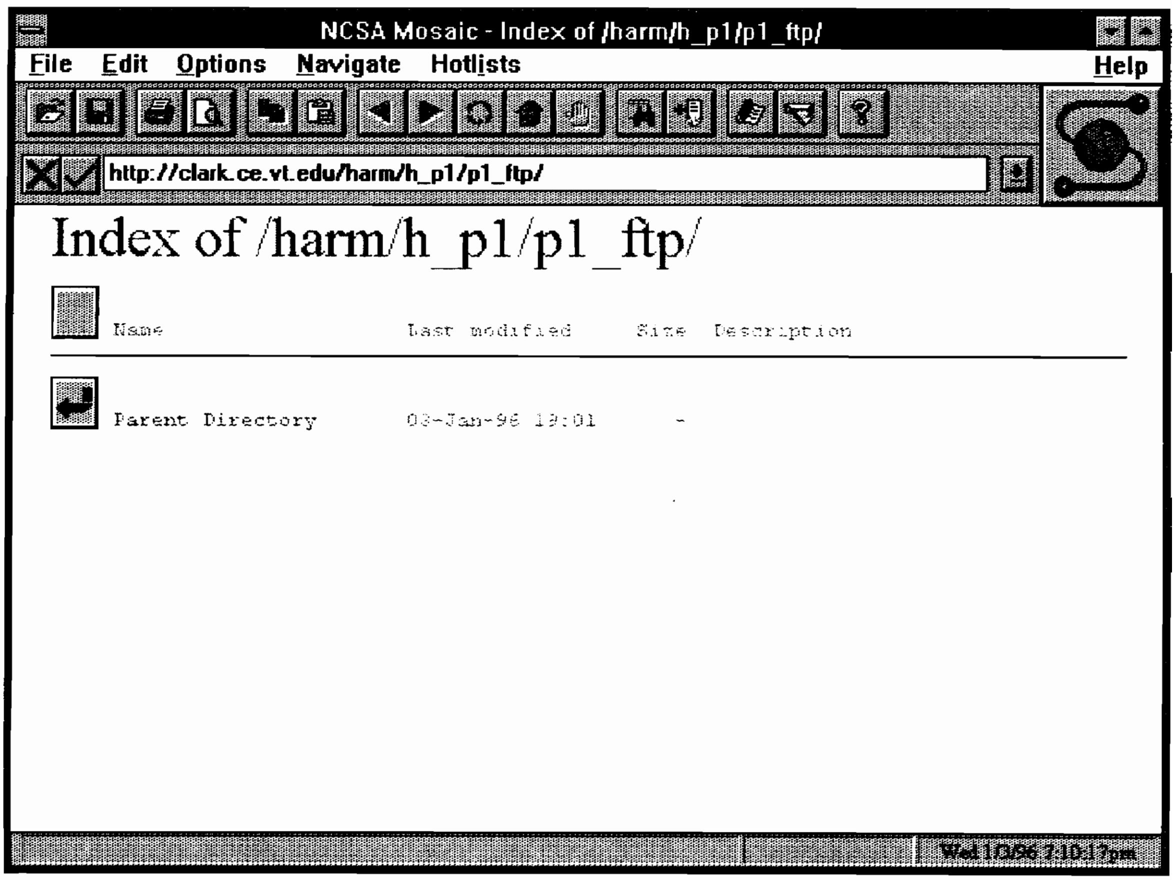

Figure A.4 The FTP directory for Project 1, for transferring updated files.

Information in the module would typically be updated in the following steps :

- Activate an ftp session to clark.ce.vt.edu
- Login with valid user name and password
- Transfer updated files to
- Apprise the HARM administrator of the transfer via e-mail, and request that the files be moved to the appropriate archive under that project.

Vita

Kalpajit De was born in the town of Dartford in Kent, United Kingdom. He began the first few years of his schooling in Hornchurch and Brentwood, England. He subsequently moved to India, where graduated from Don Bosco Park Circus High School, Calcutta. He earned his bachelors degree in Civil Engineering from the University of Madras in 1991. After graduation, he joined Development Consultants Limited, Calcutta, India, where he was employed as Design Engineer in the Nuclear Power Cell until July 1993.

He joined the Construction Engineering Division of the Via Department of Civil Engineering, Virginia Tech, as a graduate student in the Fall of 1993. He completed his masters degree in January 1996. He will be joining the San Francisco, California office of M.A. Mortenson Company as Field Engineer at the end of January 1996. His first assignment will be on the seismic upgrade of the Kaiser Hospital Project.