CONCEPTUAL MODEL FOR INVESTIGATION OF THE INFORMATION SERVICING IN A DISTRIBUTED INFORMATIONAL-LEARNING ENVIRONMENT

Radi Romansky, Iva Nikolova

Abstract: *This paper examines some aspects of the information servicing in a distributed information-learning environment called DILE. The problems concerning the educational and informational resources sharing between the dispersed users and a resource discovery mechanism in the context of the distributed resource sharing environments are discussed. We present a conceptual model and its organization for investigation of information servicing and preliminary performance evaluation of simple resource discovery mechanisms based on request propagation.*

Key words: *Distributed Learning Environment, Collaborative Learning, Group-work support, Distributed Resource Discovery Services*

1. INTRODUCTION

A distinctive feature of the distributed informational-learning environments (DILE) is their ability to provide support for group working and collaboration between dispersed users [1]. In this connection it is very important to provide a demanding combination of facilities for resource development and service provision: *interactivity*, *flexibility*, *responsiveness*, *availability*, *coherence*, *security*, *dynamic resource and group configuration*, *monitoring*. Informational services require: access to static and dynamic information regarding system component and services; a framework that fits well the heterogeneous and dynamic nature of DILE, including decentralized maintenance and operation; scalability and performance. Sharing relationships among the educational and informational resources supported by the environment may vary from mostly static to highly dynamic one [2]. In either case, the fact the users have little or no knowledge of the services and resources contributed by participants in the virtual organization poses significant obstacles to theirs effectively use. For this reason, informational services designed to support the initial discovery of the resources, as well as monitoring implementation are vital parts of DILE.

The current paper presents a conceptual scheme for modeling investigation of information servicing in a distributed learning environment that offers sharing mechanisms and multiple access to heterogeneous resources (CGI programs, static HTML, applets, servlets, tools for collaborative working like shared repository for arbitrary documents, notebooks, meta-tools, online instruments, computers, services, data and etc.), all connected through the Internet/Intranet and the middleware software layer, that support the basic services for the system components. An approach for investigation of discovery service, which aims to estimate simple techniques for discovering shared resource, is discussed.

2. PROBLEM STATEMENT AND CONCEPTUAL MODEL ORGANIZATION

Many solutions have been proposed and described for the network topology of the components of the collaborative learning environments [3] with a view of effectively representing of the educational resources and informational services. Each solution has its own advantages and drawbacks, which are presented in [4].

In the design of the DILE platform it takes into account the informational principle, presented in [5]: ***RESOURCES=M*N*B*T*P***, where: *M* – number of messages transmitted in the network environment; *N* – average number of destination hosts for each message; B – average amount of network bandwidth required for a message to each destination; *T*-

timelines for packets delivering to each destination; *P* – number of processor cycles required to receive and process each message. Obviously, in order to utilize the distributed services and resources, as well as to improve the system performance it is necessary to looking for a communication model between the components of the system, which reduce some of the above variables (for example, *N* and *P)*.

In this connection, the general principle during the design phase of the environment that supports the distributed model of learning is to divide the workload for the necessary services to a set of servers' nodes and to provide the necessary workspace of resources for the groups of distant users of the environment - students and tutors (Fig. 1a.). The distributed processing between the components of the structure is presented in Fig. 1b. Figure 2 illustrates the simplified scheme for distributed authentication and authorization. This allows for the secure remote access of distributed resources.

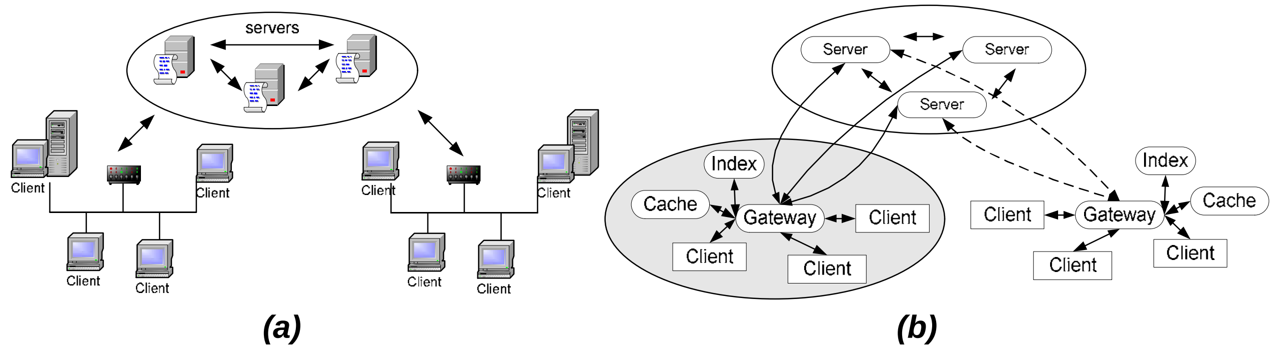

Figure 1.

The main advantages of such communication model are not far to seek: *scalability, coherence of authentically procedures, reliability and flexibility*. The drawbacks of this solution are the increasing the parameters *T* and *M*, due to the usage of many servers and the fact that these servers should change additional information for their synchronization.

In the context of such decentralized management one of the main problems of the information servicing is the discovery of the desired resources, giving its description. This is further complicated by the natural tendency of such environments to evolve over time (for example, with joining and leaving server nodes along their resources and users, changing the sharing policy and resource characteristics as accessibility, loading and etc.) as well as the resource heterogeneity. All these create significant difficulties for traditional centralized and hierarchical resource discovery services. Many research works in that area show, that in such dynamic environments it is more efficient to forward requests than to disseminate steal resource information. Request propagation is currently an active research topic [6,7,8]. The main goal of our study is to investigate a set of simple techniques for request propagation, in order to examine and to evaluate its efficiency for resource utilization in an information network for distributed learning.

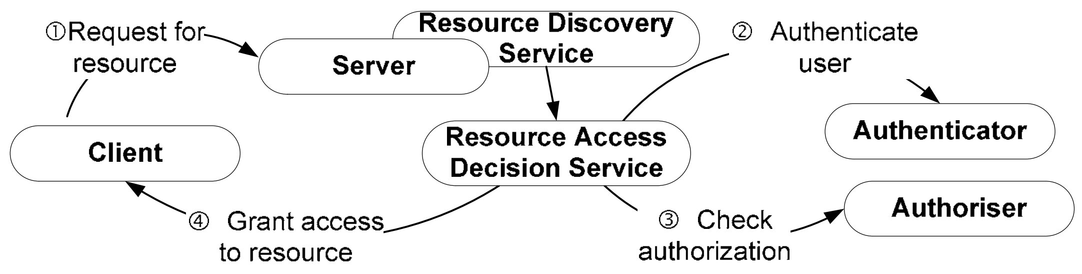

Figure 2.

Our conceptual model it is defined on the base of the following features:

1) We considered a collaborative learning group, where all online resources may be shared and used by multiple concurrent users. The group is collection of cooperating nodes providing a common set of services (file sharing, resource sharing, indexing, searching, discovering, membership and etc.).

2) Each node may request services from the others or may offer services to others. Thus, depending on the situation the nodes may support the *client* as well as *server* functionality.

3) The network infrastructure allows multiple requests to be serviced simultaneously by node. Particular nodes communicate in between via sending and receiving messages, marking the request type and the received answer.

4) Request may be processed locally or remote – as regards the propagation rules. Users send their requests to some known (typically local) node. The node responds with the matching resource descriptions, if it has them locally, otherwise it processes the request forwarding it to another node. The request is forwarded through the intermediate nodes until its TTL (Time-To-Live) expires or matching resources are found. If a node has information matching a forwarded request, it sends the information directly to the node that initiated the forwarding (rather than via intermediate nodes).

5) Within the collaborative group, the resource discovery algorithm is defined by two mechanisms: *the membership service* and a *request-forwarding strategy.* The nodes connect with each other and execute the Joint Algorithm. The algorithm is only meant as a protocol for new nodes that are not a member of collaborative learning group to joint it by connecting with exactly one node that is already a member of the group. The membership protocol is responsible as well for collecting, refreshing and providing a particular node with membership information about the other nodes. The request-forwarding strategy is used to determine to which from the locally known nodes requests should be forwarded. In addition to contact addresses, nodes can store also information about the nodes, which belong to the collaborative group, such as information about requests previously serviced. The trade off between the amount of information about neighbors and search performance generate a large set of alternatives – from random forwarding (no information about resources provided by other participating nodes) to one-hop forwarding (when nodes know exactly which node has the requested resource). Because of the dynamics of this information, nodes cache only the address/addresses where the relevant information was previously found. Nodes will forward these requests, which could not be serviced to the participants, selected amongst the nodes that are locally known.

3. PLANNING AND ORGANIZATION OF THE MODELING INVESTIGATION

Our objective is to evaluate the performance of the request processing. Therefore, we are planning to implement a simulation experiments to understand whether a decentralized resource discovery approach, applied in the context of distributed learning environments is appropriate in terms of *response time*, *response quality*, and *scalability.*

The performance of the information servicing from a point of view of the resource discovery, depends on the characteristics of the distributed information-learning environment and its usage. Management of the resources requires actualized data concerning the intensity of user's access, system load, type of the input/output streams and processes' efficiency. These characteristics it is necessary to be preliminary specified and formalized. For that purpose, some of the most important parameters for the performance investigation of the decentralized resource discovery approach should be analyzed: *(a)* A set of participating nodes and interconnection network – the connection graph strongly influences the resource discovery performance; *(b)* Distribution of services and resources on nodes – the amount of resource information hosted by nodes varies: some nodes share a large number of resources, others just one. These resource distributions can be modeled by means of stochastic processes; *(c)* A set of distinct types of resources and resource frequency – some of the resources are numerous and widely available, others are rare or unique; *(d)* The number of distinct requests for services, that every participants will generate in random distribution; *(e)* A system for measurement and collecting statistics for the distribution of the servicing times, waiting times and delays, the number of intermediate nodes via the requests had been forwarded**;** (*f)* Request propagation algorithms – simple neighbor selection policies that are based on randomness and or historical information about the results of prior serviced requests will be simulated and analyzed.

In modeling the information servicing, several assumptions are used to simplify our models:

- Nodes that are members of the collaborative learning group are joined to the information network in advance. It assumes no failures. The connection graph changes over time, influenced by location of resources, request-forwarding strategy, the number of user requests.
- Each node is implemented as a process that communicates with other nodes.
- Each node maintains two types of information: a) about a number of resources it contributes and b) about other nodes node in the system. The amount of information hosted by nodes may vary and it is stochastically distributed.
- Each node is empowered as both, a resource server and a client, depending on the role required at the time.
- Each node has possibilities for: examining if the request it is for it or not; terminating requests with expired TTLs; counting terminated requests; forwarding requests it cannot answer to a node selected from the locally known nodes.

The model parameters that will be conducted during our simulation experiments, with different servicing disciplines are presented in the Table 3.1.

Table 3.1

Model parameters, will be conduct during the simulation experiments	
Number of arrived requests	- for different number of nodes
Number of lost requests	- in the absence or rarity of some of the services
Number of forwarded for requests	- for different number of nodes (including requests that have been lost or locally processed
Distributions of servicing, waiting times and the delays	- for balanced (all nodes have equal number of resources) and unbalanced resource distribution (a significant part of nodes have no resources)
Processing load for the servers	- for balanced and unbalanced service distribution

4. CONCLUSIONS AND FUTURE WORK

In this paper, we outlined some of the basic aspects of the information servicing in a distributed informational-learning environment at the conceptual level. We have highlighted some issues around the investigation of the resource discovery service. The general objective of our study is to evaluate a decentralized resource discovery approach in terms of response time, response quality and scalability. To better understand the tradeoffs between communication costs and performance and to characterize the correlation between the resources sharing characteristics and the user requests patterns, we are considered to evaluate a set of simple resource discovery techniques by building their simulation models.

REFERENCES

1. Michaelson, R. "Web-based Group Work, in Proceedings of the 10th Annual CTI-AFM Conference". 1999, Brighton: CTI-AFM.

2. K. Czajkowski, S. Fitzerald, I. Foster, and C. Kesselman. "Grid information services for distributed resource sharing". In 10th IEEE International Symposium on High Performance Distributed Computing, 2001, San Francisco CA, IEEE Press.

3. Pandzic I. S., Joslin Ch. And Magnenat-Thalmann N. "Trends in a Collaborative Virtual Environment" International Conference on software, Telecommunications and Computer Networks-SoftCOM 2000. Split, Rijeka, Dubrovnik (Croatia), Trieste, Venice (Italy), October 2000, p. 893-901.

4. Broll W. "Bringing People Together – An Infrastructure for Shared Virtual Worlds on the Internet". "Proceedings of the IEEE WE-TICE'97". IEEE Computer Society Press.

5. Singhal S. and Zyda M. "Networked Virtual Environments: Design and Implementation". ACM Press, 1999

6. Rowstron, A. I. T., and Druschel, P. Pastry: Scalable, decentralized object location, and routing for large-scale P2P systems. In Middleware (2001), p. 329-350.

7. Kutten, S., and Peleg, D. Deterministic distributed resource discovery. In 19th ACM SIGACT/SIGHOPS Symposium on Principle of Distributed Computing (Portland, Oregon, 2000)

8. Stoica, I., Morris, R., Karger, D., Kaashoek, M. F., and Balakrishan "A scalable p2p lookup service for internet applications. In SIGCOMM, USA, 2001.

ABOUT THE AUTHORS

Assoc. Prof. Radi Romansky, Ph.D, Department of Computer Systems, Technical University – Sofia, Phone: +359 2 965-25-24, E-mail: rrom@tu-sofia.bg

Assist. Prof. Iva Nikolova, Department of Computer Systems, Technical University – Sofia, Phone: +359 2 965-26-80, E-mail: inni@tu-sofia.bg