

A comprehensive agent-based architecture for intelligent information retrieval in a distributed heterogeneous environment

Neal G. Shaw^{a,*}, Ahmad Mian^b, Surya B. Yadav^{c,1}

^a*Department of Information Systems and Management Sciences, College of Business Administration,
University of Texas at Arlington, Box 19437, Arlington, TX 76019-0437, USA*

^b*FILECONTROL.COM, Houston, TX, USA*

^c*Area of ISQS, College of Business Administration, Texas Tech University, Lubbock, TX 79409-2101, USA*

Received 1 June 1998; received in revised form 1 April 2001; accepted 1 June 2001

Abstract

The Internet has become the global infrastructure supporting information acquisition and retrieval from many heterogeneous data sources containing high-speed text and rich multimedia images, audio, and video. AgentRAIDER is an ongoing research project at Texas Tech University designed to develop a comprehensive architecture for an intelligent information retrieval system with distributed heterogeneous data sources. The system is designed to support intelligent retrieval and integration of information from the Internet. Current systems of this nature focus only on specific aspects of the distributed heterogeneous problem such as database queries or information filtering. Consequently, these concepts and others have never been successfully integrated into a unified, cohesive architecture. This paper discusses the design and implementation of the AgentRAIDER system and identifies areas for applications of the system in various domains. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Intelligent agents; Information retrieval; Heterogeneous data sources; Distributed decision-making

1. Introduction

1.1. Background

The Internet has become the global infrastructure supporting information acquisition and retrieval from many heterogeneous data sources containing high-speed text and rich multimedia images, audio, and video. Users attempting to extract valuable informa-

tion from such a complex environment largely fail, often either because not enough information is available or because of information overload [18]. The amount of information available via the Internet has become too large to be managed by an individual user. System developers must develop information retrieval systems that can assist users in accessing and managing the information base that is the Internet.

When developing such a system, however, many non-trivial issues arise for a system developer. For example:

- How can one system access data from multiple data sources, all of which are different in nature?

* Corresponding author. Tel.: +1-817-272-7398; fax: +1-817-272-5801.

E-mail addresses: nshaw@uta.edu (N.G. Shaw), amian@filecontrol.com (A. Mian), yadav@ba.ttu.edu (S.B. Yadav).

¹ Tel.: +1-806-742-2165; fax: +1-806-742-2099.