

Agent Design Considerations within Distributed Information Retrieval Systems

Jonathan Carter

Dr. Ali Ghorbani, Ph.D., Department of Computer Science

Dr. Bruce Spencer, Ph.D., Department of Computer Science

Faculty of Computer Science, University of New Brunswick, Fredericton, NB, E3B 5A3

Abstract. This paper presents the design and implementation of an agent society within the confines of IBM's Aglet advanced programming interface specification. The society exists with the sole purpose of finding MP3's throughout a given network. A sociohistorical context of the problem of finding MP3's is provided. The core design issues of agent tracking, agent messaging, social structure, agent mobility methods, and agent security are addressed and resolved. An end-user interface is established along with a detailed discussion of design strategy.

1 Introduction

Due to the inherent legal controversy surrounding MPEG's encoding technology[1], many system administrators of commercially owned Internet sites are removing found MP3's from their FTP and WWW sites for a myriad of reasons.

Most commonly, these files are removed to minimize unwanted network traffic and to keep resources allocated to business-related tasks. Many other sites devoted to MP3 distribution are continuously shutdown and restarted due to economic constraints or technical issues.

In both cases, the physical location of MP3's are changing on a daily basis. As such, users seeking MP3's often find broken web links, shutdown FTP sites, and obsolete MP3 file listings. Users can commit several hours searching for a favourite song that inevitably has moved or simply been deleted.

The project's primary objective is to solve this problem through a distributed information retrieval approach using mobile agents[3]. This solution will be expressed through a design strategy that addresses many of the common issues inherent in mobile agent development. The design strategy is reflected through a physical implementation that is realized using the IBM Aglet Software Development Kit[2] written in the Java programming language.

The project's secondary objective is to propose various artificial intelligence (AI) techniques that can be used to improve the speed by which

the proposed mobile agent searches for a requested MP3. These techniques will be considered from the point-of-view of the mobile agent's traversal technique and recognition of a particular song.

The paper is organized in the following way. The following section introduces the end user perspectives of the final application from both the consumer and advertiser perspective. Specifically, the first subsection deals with the configuration of the consumer agent. The second subsection deals with the agent validation from both the consumer and advertiser perspective. Finally, conclusions are summarized in the last section.

2 End User Perspectives

Generally, users are attempting to find a desirable MP3 with a known artist or known song title in a quick and convenient way. A primary user objective is to make this task as easy as possible for both the consumer and the advertisers associated with MP3's. The consumer is defined as the user that is seeking a given MP3. The advertiser is defined as the user that seeks to inform the user of their specialized site that brings together agents that are interested in the particular music genres. Both groups of users use the sites in slightly different ways. As such, their roles and responsibilities dictate that they must interact differently with the sites.

2.1 Agent Configuration

After an authorized account has been established for a consumer, the consumer agent must be individually tailored to suit the needs of the consumer. As seen in figure 1, the user is presented with a dialog box for inputting of various configuration parameters. Primarily, the consumer customizes the agent through the choice of travelling patterns. A travelling pattern dictates how an agent plans on traversing sites within the confines of an established and possibly growing itinerary.

The first and simplest method of travel involves the sequential pattern. As the name suggest, a sequential pattern involves simple linear movement from one proposed site to the next in search of the consumer's MP3 through communication with fellow consumer and advertiser agents. After all sites have been visited, the consumer agent is directed back to its point-of-origin where it waits for the original user. Fortunately, this origin is manually set by the user. As such, the agent can be directed to return to any online site while the user remains offline. This frees users from having to perform manual searches. Instead, they can turn their computer off while their agent performs an exhaustive search of a society of agents.

The second travelling pattern involves the circular pattern of consumer agent movement. This pattern can be considered a descendant of the

sequential travel pattern. In this case, the inherited behaviour of travelling down a queue of sites is still prevalent. However, at the end of the queue, the consumer agent immediately returns to the first member of the queue and repeats the process again. This cyclic activity continues until time has expired as specified by the user. At such time, the reset of the current cycle is completed and the agent returns to the point of origin.

Normally, one would suspect that activity should halt upon the agent discovering that no more MP3's will be found within the given loop. Instead, the behaviour has been defined such that the activity continues in the interest of maximizing the possible number of URL's corresponding to the given song. Due to the nature of the problem, it is believed that this approach is better due to the unreliable nature of possibly expired URL addresses. Eventually, the agent returns to the predefined point-of-origin.

Primarily, this type of travelling pattern is ideal for advertising agents. An advertiser is going to be interested in marketing their site to specific demographics of agent society. As such, it is assumed that agents of a particular demographic will visit agent sites contained within familiar circuits for that demographic.

Advertisers may not want to promote their sites on other sites based on political or social factors. From this logic, an advertiser may want a highly controlled path for agent travel along with a repeated advertising presence within these familiar circuits. Thus, the circular plan of travel is ideal for advertisers that want extreme control over content and direction.

The last travelling pattern is a direct descendant of a circular travelling pattern. This plan reflects some level of intelligence that the other plans failed to reflect. Like the circular plan, the intelligent pattern moves in a circular motion of travel until time has expired for the agent. In order to move from site to site, this pattern ranks all unvisited sites based on a probability of success model. This probability of success dictates the likelihood of connecting with the site combined with a previous history of successful searches. If a site connection fails, the probability of success in connecting with the site is reduced by an unfixed percentage and it is flagged as unvisited.

As an example, consider an itinerary involving sites A, B, and C that have a 90

After the agent travel pattern has been established, an additional brain parameter to the agent must be set. By definition, the brain of an agent acts as the knowledge base for the agent. It stores all associations between songs and URL addresses. The brain resides as a secure agent on the host that is defined as the point-of-origin. The brain is used by agents to assist one agent in helping another in searching for MP3 URL's. Naturally, a high level of security must be maintained to verify that an agent's brain is communicating with the owner agent.

The user can perform various utilities on the agent's knowledge base. These actions are usually performed after the agent has been established and the agent has already learned some associations between MP3 requests and the respective sites. The user may choose to start over fresh and delete all current associations. This is useful for situations of known obsolescence on the user's part. Naturally, obsolescence will occur with time as the sites corresponding to a given song continue to change. If the agent's knowledge base is several weeks old, it is very likely that all associations are invalid and the agent may contribute to globally false associations. As such, it makes sense to perform routine purges.

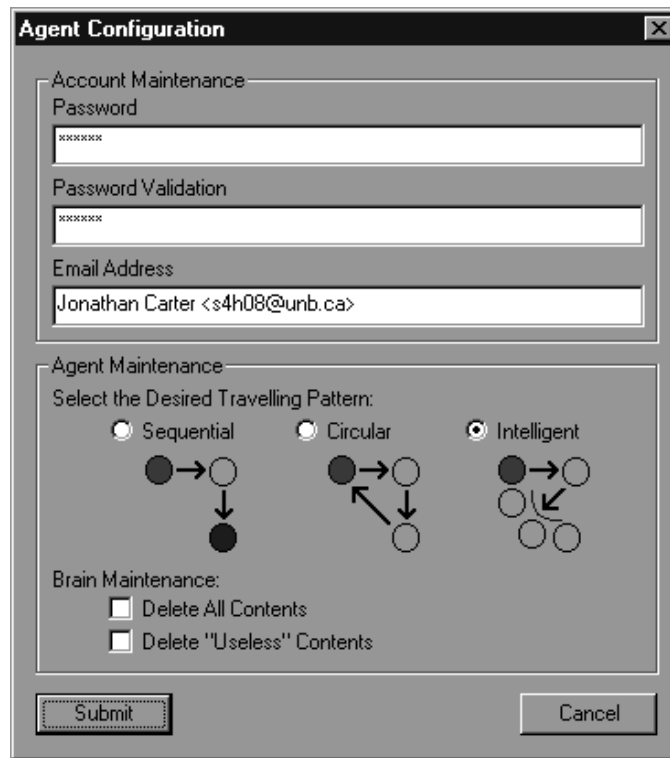


Fig. 1. Consumer Agent Configuration Snapshot

2.2 Agent Validation

After the account and agent maintenance have been performed for a fledgling agent, a new agent or an existing agent enter the agent validation stage. Within the agent validation phase, the user is given the opportunity to communicate their personal desires to the agent.

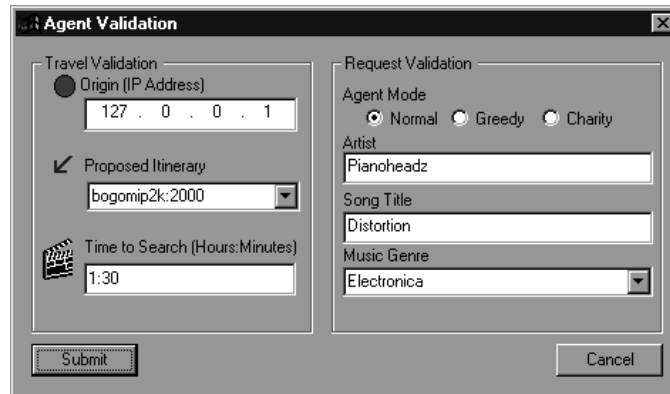


Fig. 2. Consumer Agent Validation Snapshot

These desires provide the necessary distinctions between each consumer agent with respect to goals and overt behaviours. The screen snapshot is depicted in Figure 2.

The advertisers are responsible for letting consumers know about their site's existence and the types of music their site specializes in. Agents that are seeking a particular style of music are designed to go to sites that specialize in that style of music so that agents can quickly contact other agents that share a similar interest and possibly a similar knowledge base of that style of music. Some of the agents on the advertised site may be exclusively put in charity mode by the advertisers in order to promote their style of music.

Like the consumer agent, an advertiser agent must be fully configured before the advertiser can reflect their own personal desires in the agent. After the agent configuration section, the advertiser can now interact with the advertising agent to set the advertising message they wish to broadcast. This is depicted in Figure 3

As seen in this figure, the advertiser must establish the point-of-origin, as required by all agents. As well, the advertiser must manually establish the itinerary of sites to visit. It is best to allow the advertiser complete control over which sites are going to be visited as the itinerary may be based on marketing, political, or other business decisions that are of far greater importance than allowing for a free roaming advertiser agent. The equivalence of the expiration time of the consumer agent is the time to advertise within the advertising agent. After an advertiser's time has expired, it returns to the point-of-origin. Lastly, the actual message of the advertisement must be defined. This message includes the address of the site along with the genres of music it is trying to attract.

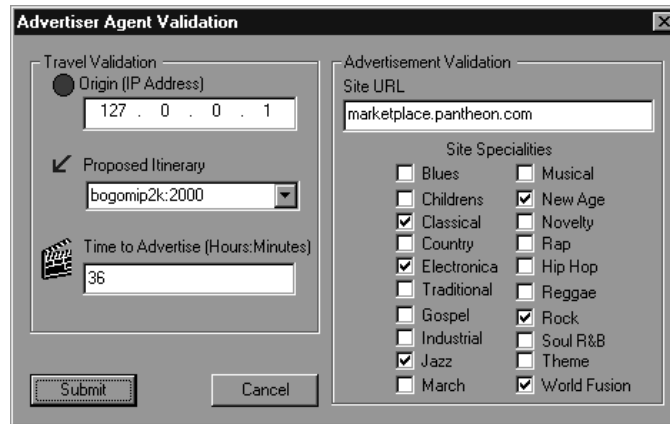


Fig. 3. Advertising Agent Validation Snapshot

3 Conclusion

In this paper, an agent society has been proposed that addresses the need of finding addresses of MP3's distributed over the Internet. The agents of the society discover the new addresses of MP3's through the use of advertiser agents that promote use of their site for particular genres of music. Artificial intelligence is integrated into the society through the travelling behaviour of the various agents along with the recognition and storage capabilities of the addresses themselves. The travelling behaviours include the foundational sequential, circular, and intelligent planners.

References

1. Advertising Age - Interactive Daily, "Uniform Resource Locators (URL)," Napster Fans Still Buy CD's, <http://www.adage.com/interactive/daily/archives/20000807/id20000807-16.html>; accessed 8-7-00.
2. Danny B. Lange and Mitsuru Oshima, *Programming and Deploying Java Mobile Agents with Aglets*, chapter 10 pp. 171-177, Addison-Wesley, Reading, Massachusetts, 1998.
3. Nicholas R. Jennings, "On agent-based software engineering," *Artificial Intelligence*, vol. 117, pp. 277-296, 2000.