Visualizing a Wavetella Network
CPSC 533C Project Proposal

Ken Deeter
kdeeter@cs.ubc.ca

February 13, 2003

Abstract

I propose to develop a realtime visualization system for the Wavetella mobile agent-based peer-to-peer filesystem. At minimum, the visualization invovles solving a dynamic graph layout problem. The use of agents allow the system to collect a variety of information about the network to potentially aid a user in interacting with the network. This project aims to explore which pieces of information are useful to the user, and in what ways this data can be displayed to the user to aid in a decision making process.

1 Introduction

One of the most popular distributed applications to arise out of the internet is that of peer-to-peer file sharing. These networks involve millions of nodes, and claim to make available terabytes of data. At the same time, the success of these networks depends directly on the number of participants, leading to a requirement that a user of the network must be able to interact with it in a reasonable manner.

Current end-user interfaces to networks such as Gnutella, Kazaa, and WinMX, hide much of the information about the network from the user. The only information about the network that is regularly presented is a sampling of end-to-end latency between a user's host and a remote host which hosts a file a user might be interested in. It is questionable whether this simple piece of information is all the user is interested in, or rather, all the user needs to make a decision about which files he or she wishes to request.

Some more recent interfaces introduce metrics such as a remote hosts' load or some equivalent of the "current queue length" metric. However presentation of this data is usually in tabular text format, rarely providing the user with a general overview of the state of the network relevant to his or her query.

The motivation behind this work is the hypothesis that users can benefit from more information being presented to them, and that this information can be communicated much more effectively (compared to current methods) in a visual manner.

2 Wavetella

Wavetella is a P2P filesystem based on the concept of mobile agents. For an operation such as a query of the network, agents are sent out to all participating nodes, instead of querying each node through a predetermined protocol.

The use of agents provides several advantages:

- The load on the network is more evenly dis-

tributed for a query, as moving the processing entity across a network reduces communication overhead between hosts.

- Information about the network can be collected in a distributed manner. For example, the load information for all the hosts can be collected by sending a self-replicating query agent out into the network instead of contacting each node one by one.

- Because new agents are sent out every time an operation is performed, changing the system behaviour for the entire network is trivial. In traditional distributed file sharing networks, software on *all* hosts need to be upgraded to add new features to the network. An agent based system, thus has much more flexibility.

3 Visualization of the P2P Sharing Network

Depending on the type of user the system is intended for, the visualization requirements change. The types of users of a typical filesharing network can be roughly categorized into technical and non-technical users. Technical users include programmers interested in the behaviour of the agents and system managers and admnistrators interested in agent behaviour and network topology. Non-technical users include end-users that may not generally be interested in network topology or agent behaviour, but rather, only have interest in the information which allows them to make better decisions about which files to download. The following is a sampling of information may aid an end-user in making such a decision:

- *Proximity*: an idea of how "close" a particular file is from the user. A user may make a reasonable correlation between the proximity (in terms of network hops) of a system and download throughput from a system.

  In another general case, multiple copies of a particular file may be available on multiple hosts in the network. The ability to see the spatial distribution (in network space) may give the user a better idea of which files may be downloaded with greater speed.

- *Load*: Although some current systems implement a *current queue length* type of indication of load on a particular host, a user may be able get a better idea from seeing a hosts' processing (CPU) load as well as its network load.

- *Hit concentration*: In most current interfaces, unless the list of results is explicity sorted by host identifier, it is difficult to percieve whether a particular host has many files that match a particular query. In a graphical display, this idea can be represented directly, without forcing the user to make an explicit inquiry to the system.

- *etc.*

Some of these attributes are file-specific, whereas others are host-specific, or even link-specific. The challenge of building a visualization system is to design it in such a way that it easily conveys all the types of data in one coherent picture. Though this may be done by displaying a dynamically generated network topology, it is not entirely clear that this is the best choice.

3.1 Impelementation

A basic graphical interface for the Wavetella system has already been developed by several graduate students. My implementation will likely start with the pre-existing one as a base, and enhance specific functionality in the visualization of query results. A

screenshot of the currently available system is provided in figure 3.1.

The current implementation presents a static network layout, based on a prepared XML representation of the network's topology. As such, it does not translate well to the dynamic case, and it is not clear whether the topology representation is even useful to the end-user.

My plan for an implementation will be similar in concept, but the display will not necessarily show a simple network topology. The display will be more tightly integrated with the underlying system, allowing it to dynamically show events as the system performs them. I plan to use the spacial position of display elements to encode one or more of the data attributes described in the previous section. Deciding which attributes are mapped to which visual properties is indeed the point of this work, but the belief is that something better can be acheived. As for an idea of the visual layout, some basic elements of the original will be carried over, but the graphical visualization component of the interface will likely be much more significant than in the original.

3.2 Secondary (backup) Ideas

3.2.1 Dynamic Graph Layout

Given that the Wavetella network has been successfully run in simulation for a system of around 500,000 nodes, a dynamic graph layout system that attempts to visualize such a large-scale network may be a useful tool for a technical user of the system.

3.2.2 Visualization of Abstract Agents

The behaviour of agents as they spread across the network can be complex and unintuitive. The task of programming such agents is often reduced to a process of trial and error, as there is no way to easily understand how a system behaves on a global scale. The ability to visualize this behaviour would be of great benefit to designers of agent systems. Most agent programming systems have common basic functionality, such as providing a facility for agents to move around in a network. One could envision a project that abstracts this functionality from all agent systems and attempts to provide a visualization for these behaviours. This system could be adapted to any agent system simply by providing some middleware that allows the visualization system to detect certain events in the agent system implementation.

4 Personal Background

This project is intended to fulfill the requirements for both this class and CPSC 527, Computer Communication Protocols. My practical experience with mobile agent systems is minimal, but I feel I have a reasonable understanding of the concept and the associated computing paradigm.

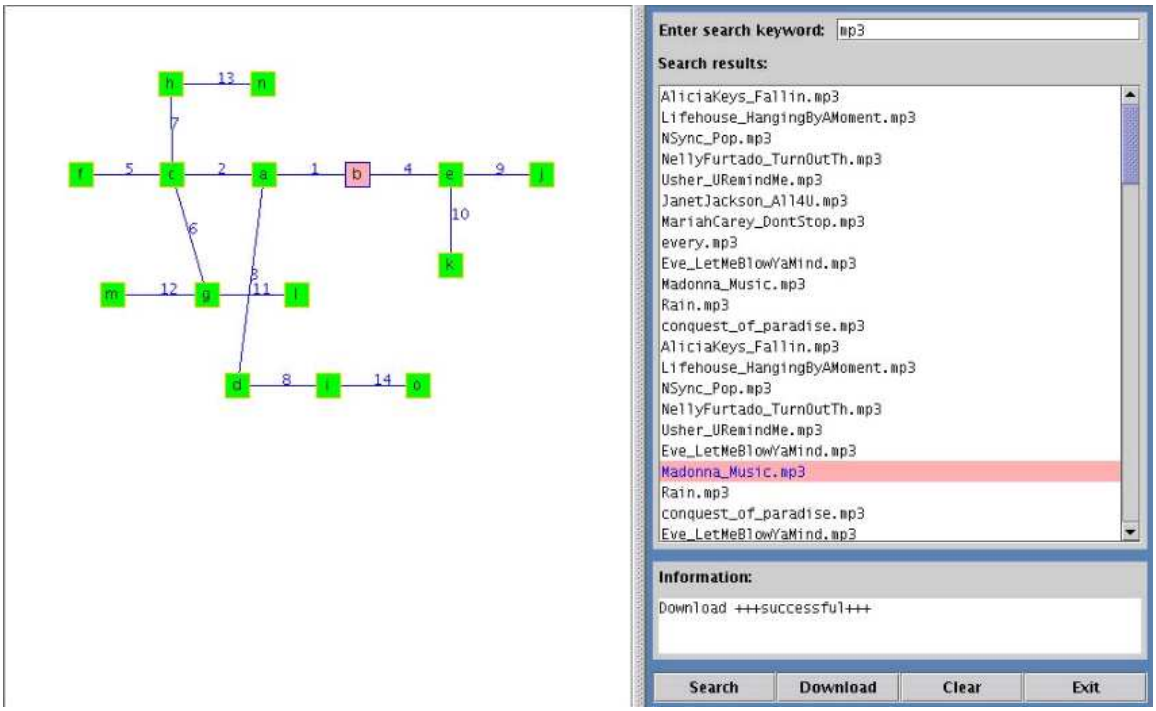

Figure 1: A screenshot of the previously developed Visual Wavetella system. The network topology layout is based on a static specification of the network topology.