Supporting Adaptive Remote Access to Multiresolutional or Hierarchical Data for Large User Groups

David Gotz
gotz@cs.unc.edu
University of North Carolina at Chapel Hill
CB #3175, Sitterson Hall
Chapel Hill, NC 27599 USA

Categories and Subject Descriptors

H.5 [**HCI**]: Multimedia Information Systems; H.1.1 [**Models and Principles**]: Systems and Information Theory—*Value of Information*; C.2.4 [**Computer-Communication Networks**]: Distributed Systems—*Client/Server*

General Terms

Algorithms, Measurement, Performance

Keywords

Adaptation, Multimedia, Streaming, Remote Data Access

1. INTRODUCTION

Advances in storage and processing technologies now allow content providers and scientists to capture, simulate, or create immense collections of data. Unfortunately, advances in networking technologies, while impressive, have not kept pace. Although we can build multi-gigabit networks, achieving those line speeds end-to-end remains almost impossible. Distributing content to large groups of independent users makes that task even harder.

As the gap between the amount of data we can capture, store, and process and the resources we have to transmit that data increases, the problem of scalable and adaptive distribution becomes increasingly central to the performance of high-bandwidth and high-volume applications such as data visualization, teleimmersion, and media streaming.

In my research, I address the distribution of multiresolutional or hierarchical datasets to large groups of heterogeneous and independent users. The systems of interest in my work share a number of common properties:

- **Coherent access patterns.** I assume that systems access data in a coherent manner as opposed to random access. This means that given a current region of interest within a dataset, a system is more likely to access data nearby to that point of interest than data that is far away.
- **Large independent client pool.** I am concerned with systems that support a large set of independent clients. The clients have heterogeneous resources available for communication and have independent application needs.
- **Multiresolutional or hierarchical data structure.** The systems must utilize data that can be stored in multiresolutional or hierarchical data structures to allow for adaptive data access in response to dynamic resource limitations. Adaptive behavior can take place along multiple dimensions.

MM'04, October 10-16, 2004, New York, New York, USA.

This class of application is common across a number of seemingly disparate applications. For example, consider a digital museum providing the world with online and interactive access to an image-based model of a historic space. Imagine teams of doctors intelligently accessing scores of media streams to help treat remotely located patients. Imagine a multinational company with a visualization tool that displays massive, remotely stored CAD models to independently working employees around the world.

Among the fundamental questions that must be addressed are the following: How can we structure these massive datasets to facilitate access over a relatively narrow communication infrastructure? How can we compactly and intuitively express multidimensional adaptation policies for application-level goals? How can we efficiently evaluate adaptation policies? How can we support independent application-level needs via a distribution scheme that scales to support large user groups?

In my dissertation, I develop concepts and abstractions to begin addressing these questions. Together, these concepts form a framework for enabling the navigation of remotely stored datasets by large groups of heterogeneous and independent users. My framework defines a generic graph-based representation abstraction suitable to the wide variety of datasets used by this system class. I map this representation abstraction to a channel-based transmission scheme. I develop scalable methods for performing utility-driven adaptation of each client's data stream which allow for the distribution of multiresolutional or hierarchical data to large user groups.

2. ACCESS FRAMEWORK

My framework for scalable and adaptive access to multiresolutional or hierarchical data consists of four major components. First, I propose an abstract data representation that is powerful enough to express a set of representation properties essential to distribution and adaptation, yet flexible enough to apply to a large class of applications. Second, I propose a channel-based transmission scheme for scalable distribution. Third, I propose algorithms for adaptation which provide the means for mapping application-level goals to specific communication operations. Finally, I discuss the development of performance models to measure the implications of different engineering decisions on functioning systems.

2.1 Representation

I define an abstract data representation based on a graph structure embedded within a multidimensional space. I use nodes in the graph to represent individual elements of information. Edges connect nodes to express syntactic relationships between elements (such as encoding dependencies) and to represent the data required to resolve the information at one node given the resolved information from another. Furthermore, edges are colored to form groups called clusters. These clusters form units of data that are grouped together at the access level of a system.

The representation graph is embedded within a utility space, defined by the available dimensions of adaptation in the dataset. The utility space is used to define spatial metrics that measure the utility of specific nodes to the application given current system conditions. Current conditions are represented with a point of interest within the utility space, a prediction vector that anticipates future needs, and a set of scale factors used to manage the tradeoffs between the available dimensions of adaptation.

2.2 Transmission

To achieve scalable delivery to large user groups, I designed my framework to remove any per-client work from the central server. I use a channel-based transmission scheme which depends on a subscription oriented service, such as broadcast or multicast. For smaller user groups, unicast may also be appropriate.

In the multicast case, my design maps each cluster from the data representation abstraction to a unique transmission channel. The data from each cluster is sent out repeatedly, returning to the start of the associated data stream once the final byte has been transmitted. A server therefore continuously transmits a large set of channels, one for each cluster. This method does not require any per-client work by the server. Interested clients subscribe to a set of channels, independently controlling their incoming data stream by choosing appropriate channels for subscription.

2.3 Adaptation

I have developed algorithms for adaptation that make use of my abstract data representation graph and the multidimensional utility space within which the graph is defined [3]. I define a set of generic adaptation operations that can support multidimensional and multimedia adaptive behavior.

My adaptive framework uses a spatially defined utility metric and a cost metric that combine to measure the utility-cost ratio of adaptation alternatives. I define adaptation as an attempt to maximize the utility-cost ratio; an attempt to obtain the most useful information at the lowest cost. The maximization process is performed independently on each client, guiding the management of the set of subscribed channels.

Through this adaptation framework, individual clients can adapt their incoming data stream to reflect both network-level, and application-level conditions. For example, the size of the subscribed channel set can be used to control bandwidth utilization. Application-level data requirements can be reflected in the utility metric and changing needs are satisfied via channel subscriptions.

2.4 Performance Model

There are several engineering parameters which must be specified while implementing a working system. For example, the list of parameters includes bandwidth allocation, data representation granularity, network delivery method, and many others. The values chosen for these parameters can have a significant impact on overall system performance.

I have performed a brief evaluation on the impact of certain parameters in the context of digital museums [2]. In current work, I am developing models that measure and predict the performance of of my framework in a generic context. This model will help designers intelligently engineer their systems.

3. CONTRIBUTIONS

My dissertation research presents a framework for adaptive remote access to multiresolutional or hierarchical data to large user groups. Within this context, there are several specific contributions:

- **A Generic Representation Abstraction.** I distill the common elements of data representation into a generic abstraction. Any database that can be mapped to this abstraction is compatible with the communication and adaptation potions of my framework.
- **Utility-Based Adaptation via a Multidimensional Spatial Metric.** I present algorithms for expressing and evaluating the utility of information through spatial utility metrics. This spatial expression of utility can be much more intuitive than complex rule-based methods, especially for systems with many dimensions of adaptation.
- **Simple Server Design.** The channel-based communication framework leverages a multicast/broadcast infrastructure to allow interactive access to theoretically unlimited user group sizes.
- **Performance Model.** The performance model can be used to measure and predict system performance across a range of values for several engineering parameters.
- **Driving Application for Multicast Development.** My system prototype, designed to stream image-based rendering datasets to large user groups, is sensitive to several end-user properties of multicast delivery including latency, loss rates, and join/leave delays. This application can motivate future multicast development and performance testing.

4. RELATED WORK

Other researchers share my goal of defining a generic adaptation model. For example, an alternative utility measure was proposed in [7]. Several design principles were outlined in [4].

My methods for supporting scalable delivery for interactive applications through channel set management were inspired by several projects in the area of broadcasting video-on-demand, including Pyramid Broadcasting [6].

The driving application in my research is a remote version of Sea of Images [1]. My representation allows multidimensional adaptation in much the same way as layered video representations [5].

5. REFERENCES

[1] D. G. Aliaga, T. Funkhouser, D. Yanovsky, and I. Carlbom. Sea of images. In *Proc. of IEEE Visualization*, 2002.

[2] D. Gotz and K. Mayer-Patel. A framework for scalable delivery of digitized spaces. *To Appear in the International Journal on Digital Libraries*. Special Issue on Digital Museums.

[3] D. Gotz and K. Mayer-Patel. A General Framework for Multidimensional Adaptation. In *Proc. of ACM Multimedia*, 2004.

[4] M. McIlhagga, A. Light, and I. Wakeman. Towards a design methodology for adaptive applications. In *Proc. of ACM/IEEE International Conf. on Mobile Computing and Networking*, 1998.

[5] R. Rejaie, M. Handley, and D. Estrin. Layered quality adaptation for internet video streaming. *IEEE Journal on Selected Areas of Communications (JSAC)*, Winter 2000. Special issue on Internet QoS.

[6] S. Viswanathan and T. Imielinski. Metropolitan area video-on-demand service using pyramid broadcasting. *Multimedia Systems*, 4:197–208, 1996.

[7] J. Walpole, C. Krasic, L. Liu, D. Maier, C. Pu, D. McNamee, and D. Steere. Quality of service semantics for multimedia database systems. *Database Semantics: Semantic Issues in Multimedia Systems*, 1999.