Semantic Multicast: Intelligently Sharing Collaborative Sessions

Son Dao, Eddie Shek and Asha Vellaikal
HRL Laboratories,
Richard R. Muntz, Lixia Zhang and Miodrag Potkonjak
University of California, Los Angeles
and
Ouri Wolfson
University of Illinois,Chicago

We introduce the concept of semantic multicast to implement a large-scale shared interaction infrastructure providing mechanisms for collecting, indexing, and disseminating the information produced in collaborative sessions. This infrastructure captures the interactions between users (as video, text, audio, and other data streams) and promotes a philosophy of filtering, archiving, and correlating collaborative sessions in user and context sensitive groupings. The semantic multicast service efficiently disseminates relevant information to every user engaged in the collaborative session, making the aggregated streams of the collaborative session available to the correct users at the right amount of detail. This contextual focus is accomplished by introducing proxy servers to gather, annotate, and filter the streams appropriate for specific interest groups. Users are subscribed to appropriate proxies, based on their profiles, and the collaborative session becomes a multi-level multicast of data from sources through proxies and to user interest groups.

Categories and Subject Descriptors: H.3 [**Information Storage and Retrieval**]; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing—*abstracting methods, indexing methods, linguistic processing*; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*clustering, information filtering, relevance feedback*; H.3.4 [**Information Storage and Retrieval**]: Systems and Software—*distributed systems, user profiles and alert*

This work was supported by the DARPA Program on Intelligent Collaboration and Visualization under project No. N66001-97-2-8901.

Name: Son Dao, Eddie Shek and Asha Vellaikal
Affiliation: Information Sciences Lab, HRL Laboratories
Address: 3011 Malibu Canyon Road, Malibu CA 90265; email:{skdao,shek,asha}@hrl.com
Name: Richard Muntz, Lixia Zhang and Miodrag Potkonjak
Affiliation: Department of Computer Science, University of California
Address: Los Angeles CA 90095; email:{muntz,lixia,miodrag}@cs.ucla.edu
Name: Ouri Wolfson
Affiliation: Department of Electrical Engineering and Computer Science, University of Illinois
Address: Chicago, IL 60607-7053; email:wolfson@eecs.uic.edu

services, current awareness systems; H.3.5 [**Information Storage and Retrieval**]: Online Information Services—*data sharing, web-based services*

General Terms: Architecture, Collaboration, Semantics

Additional Key Words and Phrases: Collaboration data models, content extraction, distributed multimedia archives, incremental clustering, information dissemination, information filtering, information sharing, IP multicast, metadata, proxy-based services, user profiles

1. INTRODUCTION

Current networked collaborative activities usually consist of an "interaction stream" broadcast over a single network channel – if users have an interest in the session, they must participate in the entire event and process all broadcast information. Furthermore, if a collaborative session involves multiple interactions from interrelated working groups, a user must participate in the full broadcast from all groups to learn the interrelationship. This model of collaboration only supports two modes of operation: a user either actively participates in the session or the user does not participate at all. In actuality though, a collaborative session is one in which users participate in varying levels at varying times and multiple working groups, or interaction streams, concurrently overlap. Minimal support, if any, exists to decompose collaborative sessions among related working groups, filter and share information between groups, efficiently recall specific discussions, support occasionally disconnected users, or augment the interaction streams with hooks to online information repositories.

The semantic multicast concept introduced in this paper adds a level of "semantic indirection" to the IP multicast model by introducing a logical dissemination, enrichment, filtering and archiving structure for making the streams of the collaborative session available to the correct users at the right amount of detail using Internet multicast as the underlying communication mechanism. Section 2 discusses the architecture while Sections 3-5 discuss three main components of the semantic multicast concept, viz. semantic multicast graph discovery, collaboration data models, repositories and content extraction respectively.

2. ARCHITECTURE AND INFRASTRUCTURE

A collaborative session is comprised of the many streams of interaction arising in the session that are disseminated to all users in various forms, and the users with their specialized interests and abilities to participate in the collaboration. The goal of semantic multicast is to create a logical dissemination, filtering, and archiving structure for making the streams of a collaborative session available to the correct users at the right amount of detail and in an efficient manner. Given a collaborative session of many overlapping streams, a semantic multicast graph (SMG) establishes the flow of collaborative streams between interest groups. Nodes in the graph represent ranges and qualities of semantic topics present in the collaborative streams, and are allocated to a hierarchy of semantic proxy servers in the network that performs incremental data archiving and filtering for the semantic coverage of the node. Users subscribe to the output of appropriate proxies, based on their profiles, and the collaborative session becomes a multi-level processing and multicast

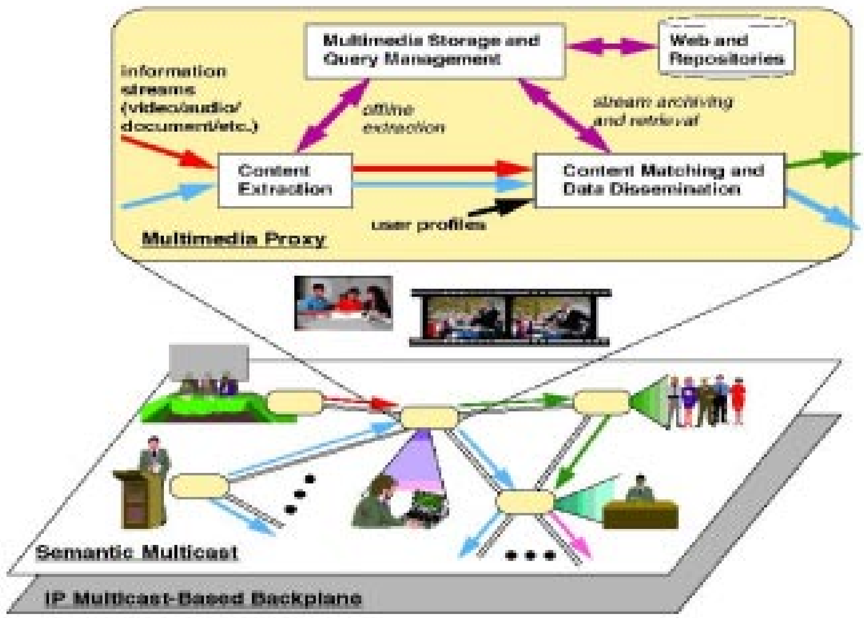

Fig. 1. Semantic Multicast Architecture

of data from sources through proxies and to users clustered based on their interest and capabilities. Figure 1 shows the semantic multicast architecture for filtering, archiving, and disseminating collaborative sessions as a combination of real-time and non-real-time as well as annotated and summarized information streams.

Internet multicast defines a decentralized and scalable mechanism for distributing entire data streams to dynamic user groups. Unfortunately, this network-level multicast protocol enforces a single level of indirection for delivering group data - a user either receives all data on the multicast stream or receives none. As a result, without integrated application support to partition and filter the information flows, multicast will never emerge to truly interconnect users at varying levels of detail and interest. Semantic multicast adds a level of "semantic indirection" to the IP multicast model while exploiting the group communication model of IP multicast as a basis for dissemination of information and control messages amongst the various logical entities in the semantic multicast environment.

3. DYNAMIC SEMANTIC MULTICAST GRAPH DISCOVERY

Given a collection of user requests against a collaborative session of many overlapping streams, a semantic multicast graph (SMG) is automatically constructed by a service assigner. A semantic multicast graph organizes user requests into a hierarchy of group requests that represents an increasingly specialized range and quality of semantic topics as one goes down the hierarchy. A SMG establishes a logical flow of collaboration stream between interest groups. The proper construction of semantic multicast graphs with an evolving user base and its materialization amongst a set of proxy servers is central to the semantic multicast service.

Semantic multicast graph discovery involves the clustering of users with similar profiles that summarizes the level of interest he/she has on topics in the collaboration. Terms in a user interest is specified in the same weighted similarity ontologies (i.e. structures that gather instance values and ranges into hierarchical similarity classes) as that used to specify the metadata summarizing the scope of content in information streams. Given a set of profiles (or stream metadata) we can compute the most specific generalization between elements of the set by using the similarity ontologies which may be provided by domain experts (such as a thesaurus) or automatically learned from samples of commonly co-occurring values [Dao and Perry 1996]. In addition, user request clustering is also contingent upon the users having similar computing and networking capabilities such as bandwidth budget so that information may be transcoded appropriately (e.g., key frames extraction from video) to limit resource utilization, and being located in a geographic neighborhood so that IP multicast trees can be prevented from overlapping if possible.

Currently, semantic multicast graph discovery consists of three main steps. When a new user request arrives, we (1) try to merge the request into an existing node in the semantic multicast graph that covers the request, (2) try to expand an existing SMG node that closely covers the request and merge them, and (3) as a last resort, create a new node containing only the new request. These steps generate a locally optimal clustering that is dependent on the order of user request arrival. To overcome this weakness, nodes in semantic multicast graphs are periodically merged and split offline in an attempt to achieve a globally optimal clustering.

Conceptual clustering of data has been studied extensively by the artificial intelligence community as a means to extract hidden inference information [Fisher 1987], and by the information retrieval community as a method to improve the quality of query results [Charikar et al. 1997; Chu et al. 1996]. Semantic multicast graph discovery differs in that the entities being clustered are user requests, and the goal of the clustering is more than straightly to optimize the accuracy of clustering but also to generate a structure that can be mapped into a proxy dataflow environment that can effectively support dissemination of information to geographically distributed users with vastly different resource availability.

4. COLLABORATION DATA MODELING AND REPOSITORY

The goal of semantic multicast is not only to support real-time collaborative activities but also the capture of collaborative work sessions in a distributed multimedia database for offline access and analysis. We utilize a formal model of collaboration [Rajan et al. 1995; Dommel and Garcia-Luna 1997] that defines the roles of participants, modes of interaction, and protocols for joining and leaving. This model serves not only as the basis for the collaborative work middleware control of a session but also as a part of the database schema for storing data captured during a session. Additional schema information is required to describe the semantic tags that are generated to support flexible, intelligent access to the data. Where necessary, control information (e.g., floor activity) will be captured also to support playback.

The establishment of a session requires selection of an existing schema for collaboration (from the database) or creation of a new one. This schema is expanded to include the semantic information and views that are materialized by the analysis tools that are to be invoked on the session streams. The analysis of the session raw data results in indices to portions of the data and also views on the data. One of the main goals is support of summary presentations that allow a participant to get up to speed, or quickly review a past session. There are several approaches to this: one is to compress data (e.g., video stream is compressed to a sequence of key frames) and a second is to "project out" information that is not germane to the current purpose. The latter exploits the user profile to determine topics of particular interest. The combination of all proxy data repositories is a distributed multimedia database. Since it is possible that data streams originating from various sources for the same interaction may get stored at different locations, a session schema is implemented as a "global view" that spans multiple proxy repositories and the corresponding playback capabilities will allow synchronization of related component streams across distributed repositories.

5. CONTENT EXTRACTION AND CORRELATION

Collaborative sessions typically consist of several types of synchronized data streams, primarily raw audio, video, text, and graphics data, along with other application-specific data types. As part of its operation, semantic multicast aims to filter, archive, fuse and disseminate these data streams. However, many multimedia data streams in their raw forms are not amenable to automated semantic interpretation [Zhang et al. 1995] and typically have to be enhanced with other features, which are either manually created/attached or are extracted by analyzing the raw data. Some of the enhancements and operations required to support the above functionalities include the following: (1) content-based data stream analysis at proxies in the network for metadata generation and filtering, (2) addition of a synchronized metadata channel accompanying raw data streams to assist filtering at the distributed proxies and archives, (3) creation of indices based on data semantics to allow fast retrieval of desired session information as determined by user interests and (4) summary creation to provide synchronization, review and browsing support for the logically disconnected users.

Multimedia data stream annotations, indices and summaries are based on the collaborative information model or schema as described in the earlier section and are closely matched or tied to end-user interests expressed in terms of a user profile. Depending on the sophistication of the underlying processing operations, these enhancements are produced in real-time, near real-time or non real-time with the subsequent increase in processing complexity typically resulting in greater precision in the operation involved [Zhou et al. 1998]. For example, processing operations which are carried out at network proxies to support online user-interest based dissemination and filtering will need to meet stringent real-time capabilities.

Content analysis of video and audio utilize a layered scheme with application-dependent models necessary at higher layers for classification based on application semantics. Video structuring operations require the segmentation of video into logical subunits comprising of a series of related scenes with common content [Tonomura et al. 1994]. In an approach involving mixed-media cues, features from video, audio and other streams of data are utilized jointly to trigger analysis and locate interesting events. For example, video processing techniques are used in conjunction with speech signal analysis to automatically locate boundaries of shots, scenes and conversations. In addition to annotation, another important functionality is the indexing of the video in a given archive. Annotations that are generated as described above can be utilized for indexing. Similarly, techniques to extract effective video summaries or abstracts that are concise versions of the video at different temporal resolutions in a real-time environment will also be part of the semantic multicast system.

6. CONCLUSIONS

In this paper, we have described the design of the semantic multicast service. Semantic multicast extends various other approaches for dynamic information sharing, including Internet multicast [RFC 1889 ; Floyd et al. 1995], information dissemination techniques, multimedia representation and storage [Dommel and Garcia-Luna 1997; Botafogo and Mosse 1995], and multimedia content extraction [Wactlar et al. 1996; Ahanger and Little 1996; Yeung and Yeo 1997]. A prototype of this large-scale shared interaction infrastructure is currently being evaluated for a distance learning application. Semantic multicast is designed to support effective dissemination of collaborative sessions over space and time (i.e., real-time and non-real-time) and among a diverse user community. It promises to bring significant benefit to many applications, including, but not limited to, distance learning and disaster assessment.

Distance learning is defined as learning that takes place when the students and the instructor are separated by space or time. Several aspects to distance learning are closely related to our concept of semantic multicast. A typical learning scenario is a multi-level one-to-many process wherein there is a single source (teacher) and multiple recipients (students) of shared text, image, audio and video signals. Study groups based on common interests and geographical location need to be supported by effective group communication mechanisms such as IP multicast. The concept of study groups also means that certain groups of students will be furthering the standard course distribution with background material and exercises. The application is rich in semantics as the students and/or study groups can have diverse interests with definite overlaps, which clearly fits our paradigm. Finally, distance learning can be benefited by the indexing and summarization facility of semantic multicast which allow archived material to be effectively retrieved at a later time.

Disaster assessment is concerned with bringing medical, monetary, and/or safety reinforcement to areas devastated by natural (or man-made) events, such as earthquakes, floods, and civil unrest. In typical scenarios, reinforcement teams are sent to independently administer relief to isolated points of the disaster area – these teams independently interact but minimal collaboration occurs across the teams or upward to higher levels of decision making. It is usually not until after the disaster has been settled that the amount of redundant effort and shared lessons/information is identified. Semantic multicast allows different teams to be connected to the same overall collaborative session so that the close interactions among members of one team are continually analyzed, abstracted, summarized, and disseminated to other teams encountering similar circumstances and needs.

ACKNOWLEDGMENTS

The authors would like to acknowledge valuable discussions, especially about semantic multicast graph concepts, with Brad Perry.

REFERENCES

AHANGER, G. AND LITTLE, T. 1996. A survey of technologies for parsing and indexing digital video. *Journal of Visual Communication and Image Representation 7*, 1, 28–43.

BOTAFOGO, R. AND MOSSE, D. 1995. The MORENA model for hypermedia authoring and browsing. *IEEE Multimedia*, 42–49.

CHARIKAR, M., C. CHEKURI, T. F., AND MOTWANI, R. 1997. Incremental clustering and dynamic information retrieval. In *Proceedings of the Conference on Theory of Computation* (1997).

CHU, W., YANG, H., CHIANG, K., MINOCK, M., AND ET AL. 1996. CoBase: a scalable and extensible cooperative information system. *Journal of Intelligent Information Systems: Integrating Artificial Intelligence and Database Technologies 6*, 2-3.

DAO, S. AND PERRY, B. 1996. Information dissemination in hybrid satellite/terrestrial networks. *IEEE Data Engineering Bulletin 19*, 3, 12–19.

DOMMEL, H.-P. AND GARCIA-LUNA, J. 1997. Floor control for multimedia conferencing and collaboration. *ACM Multimedia 5*, 1.

FISHER, D. 1987. Knowledge acquisition via incremental conceptual clustering. *Machine Learning 2*, 139–172.

FLOYD, S., JACOBSON, V., LIU, C., MCCANNE, S., AND ZHANG, L. 1995. A reliable multicast framework for lightweight session and application layer framing. In *Proceedings of ACM SIGCOMM* (August 1995).

RAJAN, S., RANGAN, P., AND VIN, H. 1995. A formal basis for structured multimedia collaboration. In *Proceedings of IEEE Multimedia Computing and Systems* (May 1995).

RFC 1889. RTP: a transport protocol for real-time applications. Internet Draft, http://www.ietf.org/rfc/rfc1889.txt.

TONOMURA, Y., AKUTSU, A., TANIGUCHI, Y., AND SUZUKI, G. 1994. Structured video computing. *IEEE Multimedia 1*, 3, 34–43.

WACTLAR, H., KANADE, T., SMITH, M., AND STEVENS, S. 1996. Intelligent access to digital video: The Informedia project. *IEEE Computer 29*, 5, 46–52.

YEUNG, M. AND YEO, B. 1997. Video content characterization and compaction for digital library applications. In *SPIE Storage and Retrieval for Image and Video Databases V* (1997), pp. 45–58.

ZHANG, H., LOW, C., AND SMOLIAR, S. 1995. Video parsing and browsing using compressed data. *Multimedia Tools and Applications*.

ZHOU, W., VELLAIKAL, A., AND KUO, C.-C. 1998. Online scene change detection of multicast (MBone) video. In *Proceedings of SPIE Multimedia Storage and Archiving Systems III* (Boston, November 1998), pp. 271–282.