PERSONAL MEDIA ALERT SYSTEMS: PERSONALIZATION AND DISSEMINATION OF BROADCAST CONTENT WITH A P2P MICROPAYMENT SCHEME

Yih-Farn Chen, David Gibbon, Zhu Liu, Behzad Shahraray, and Bin Wei

AT&T Labs-Research
180 Park Ave
Florham Park, NJ 07932, USA
{chen, dcg, zliu, behzad, bw}@research.att.com

ABSTRACT

Media consumers who are overwhelmed by multitudinous news content demand a system that would sift through tens or hundreds of broadcast TV channels on a daily basis to capture the most important clips that match users' interests and deliver these personalized clips for easy viewing on a typical TV or home PC. In this paper, we present a personal media alert system that extracts video segments for the interested users. In addition, alerts with pointers to content stored on the home server can be sent to mobile users who are authorized to access the content at home. To make content sharing feasible and scalable, we also propose a P2P payment scheme that would allow a consumer to have access to media clips aggregated on other consumers' personal media systems.

1. INTRODUCTION

Media consumers who are interested in the latest news today are bombarded with information from news channels, blogs, and pod-casts. While the broadcast TV news channels still provide the best edited content, the consumers today frequently have little time or patience to watch an entire news program searching for relevant news clips. The rapid increase in the processing power and the storage capacity of digital video recorders and personal computers, new chips for multi-channel recording, and the deployment of high speed broadband infrastructure have prompted us to build a Personal Media Alert (PMA) system to alleviate this problem for consumers. The system would sift through tens or hundreds of broadcast TV channels on a daily basis to capture the most important clips that match the keywords and profile of the user and deliver these relevant segments for easy and efficient viewing on his or her TV or home PC. Unlike a similar service we previously built where media acquisition and processing are all handled in a centralized place [1, 2], the PMA service does not require any license from content providers since the content is all recorded locally with the user's own subscription service. In addition, alerts with pointers to content stored on the home server can be sent to mobile users who are authorized to access the content recorded at home through a service such as Sony's LocationFreeTV[3]. Furthermore, the enormous individually obtained media segment collections constitute huge content resources that can be shared among users through P2P networking. To facilitate fair sharing of the content, we need to consider a P2P payment scheme that would allow a consumer to have access to media clips stored on other consumers' Digital Video Recorders (DVR) while making payments to both the original content provider and the other party that stores the media content.

To meet the requirements described above, we propose a new personal media alert system where content is acquired and processed locally on peers. Meanwhile media alerts are created based on a central index and alert content is disseminated through a P2P network for scalability. There are quite a few related systems for the purpose of extracting personalized content, such as those alerting services on the Internet, but very few can handle the required personalization for video sources [4].

In the rest of the paper, Section 2 describes how the personalized media alerts are created, and Section 3 discusses several issues involved for disseminating media alerts through a P2P network. An ongoing prototype is reported in Section 4, and finally, we draw our conclusion in Section 5.

2. PERSONALIZED MEDIA PROCESSING SYSTEM

Figure 1 illustrates the structure of the personal media alert system we proposed. A variety of content is acquired and processed on every user's machine, and the content index is shared with other users such that media alerts can be generated for everyone even if the content is not available locally. In this section, we describe a system (a node or a peer) that automatically acquires interesting video clips according to a user's personal interests in a typical home PC environment. Each node contains the following three modules. The content acquisition module records selected

broadcast TV programs from a variety of sources, including Direct Broadcast Satellite, Cable TV, and Digital TV terrestrial broadcasting; the content processing module transcodes the acquired content into different formats to serve a wide range of devices and access bandwidths, and it also extracts the embedded hierarchical content structure for query and browsing purposes; the media alert creation module constantly monitors new content, either on the local node or on other peers, and creates multimedia alerts that match the interest profiles of the users.

2.1. Acquisition Module

The acquisition module is built on the Microsoft XP Media Center Edition (MCE) platform. An MCE machine is a full fledged PVR (Personal Video Recorder) system, having the capability to schedule a single or a series of recordings of TV programs with integrated EPG support. MCE records a television show in the DVR-MS file format, which is similar to Microsoft's Advanced Streaming Format (ASF), and the enhancements allow the creation of key PVR functionalities, including time-shifting, live pause, and simultaneous record and playback. Without any extra endeavor, the metadata from the Electronic Program Guide (EPG), the closed caption (CC), audio, and video streams are seamlessly captured in one file. With the growing popularity of MCE machines, consumers can easily capture interesting broadcast content at home.

2.2. Processing Module

The source of the processing module is a DVR-MS file, and the processing results are stored in two local databases: the media database for serving the content and the content index database for answering content queries.

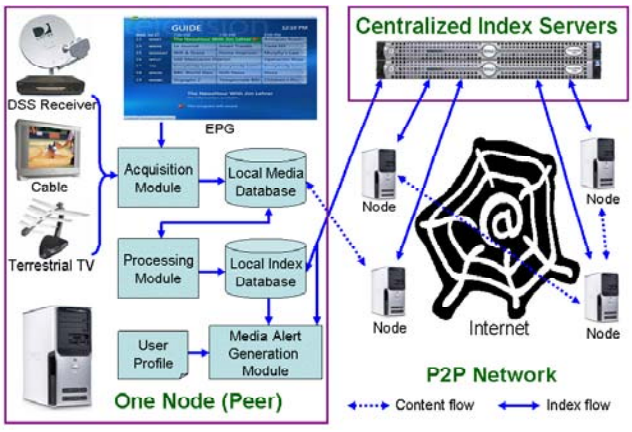

Fig. 1. Media acquisition and processing system.

The metadata in DVR-MS includes some key information about the show, for example, the program ID, the broadcast time, and some brief description of the content coming from the EPG. The metadata is extracted and saved in the content index database.

While a high quality motion video is usually more pleasing and may carry more information, in some cases and on some devices video replay is either not an option, or may be too costly. A different visual presentation of the video content can be done by selecting a subset of representative frames to convey the visual information. We use the algorithm discussed in [7] for performing the content-based sampling (CBS). This algorithm detects abrupt and gradual transitions in the video sequence, and the set of frames retained generates a compact representation of the video program.

CC contains rich content information about the program, and we use CC to search for the appropriate video clips. CC is normally not synchronized with the audio, which noticeably affects the quality of video browsing and query. Large vocabulary automatic speech recognition (ASR) is used to generate transcripts for the audio stream. After ASR, parallel text alignment is performed to align the timing information from the automatic speech transcription with the more accurate CC transcription. Alternatively, a parallel text alignment algorithm is used to import high quality off-line transcripts of the program when they become available. CC and ASR transcripts do not have case information, and the case restoration module of the processing module uses a rule-based capitalization algorithm trained from multiple sources, including AP newswire data and online stories published by national media companies. To better index and present the content, named entities, including country names, person names, locations, titles, etc. are extracted from the textual stream.

All content index information is combined to create a page/paragraph structure of the media. Each paragraph is composed of one scene cut frame and a set of related CC sentence. Such structure effectively represents the video data in a manner that is easy for users to browse the content nonlinearly. The content index is saved in a local index database, and it is also copied to a centralized index server for sharing with other peers.

Due to the different network bandwidth and video rendering capabilities on accessing devices, the DVR-MS files are repurposed into three WMV formats, including standard definition (SD) video (2Mbps/640x480/29.97fps), VHS-quality video (600Kbps/320x240/29.97fps) and low bitrate (LB) video (150Kbps/224x168/15fps). All the media content is saved in the local media database.

2.3. Media Alert Generation Module

The users provision their topics and associated keywords as well as the program sources for each topic in their profiles. Each topic in the profile can have different keywords and can use a different subset of the available program sources. An alert is only created if the keywords for a topic in the interest profile match content in the program sources

associated with this topic. Keywords for topics are correlated against closed caption text, speech recognized audio segments and other metadata like EPG. This is described in more detail in [5].

The media alert module periodically matches new content with user profiles. In the current implementation, a task runs at specific times in the day for each user and identifies new content that matches the program sources and keyword requirements for each topic. The first step in this task is to flag the alert content for each user. All new content since the last alert time is written out to an XML file, which includes the media content index. The relevant clips are extracted and then sent out to the users using a scaled-down version of the AT&T Enterprise Messaging Network [4].

Other approaches could entail real-time word spotting of the closed caption text as the content is being acquired. The alerts could effectively be sent immediately. Of course, the clipping segmentation and indexing algorithms would be less effective since it would not have the advantage of analyzing the entire broadcast.

3. P2P MEDIA ALERTS

While technology advances would soon allow a personal computer or media center at home to sift through tens of channels at home, we need to consider a P2P scheme if we want to allow a user to have broader access up to tens of thousands or even millions of TV programs captured on other peers' machines. Several technical and legal challenges exist for such a scheme to be applied to a personal media alert processing system:

- *Peer Content Discovery:* How does a local peer find programs recorded on other peers' machines with segments that match his or her interests?
- *P2P Media Delivery:* How does a peer receive a particular media clip stored on multiple machines, with potential node and network failures on the P2P network?
- *Digital rights management*: Unlike a personal media alert system where a user already obtains the right to record and view the programs and clips stored on the personal DVR, a clip accessed from other machines may require obtaining additional license rights from the content owners.
- *Incentives for sharing*: Obtaining clips stored on other machines requires consumption of both processing power and bandwidth from the peers that contribute the content. Incentives must be provided to promote sharing of media content among peers.

In the next two sections, we describe our current design with a centralized content index database, followed by a proposal that uses an entirely distributed scheme.

3.1. Content Discovery and P2P Media Delivery

Several P2P content delivery platforms are available today. For example, Konitki [8] is a P2P delivery management platform that allows normal PCs to publish, serve, download, and play rich media of all kinds. Through *adaptive multirate serving*, each computer can simultaneously connect to multiple servers that host a particular clip and request small parts of the file from each server. The Knotiki client can monitor the download process and dynamically drop existing connections and make connections to new servers for better performance.

Another example is Swarmstreaming [9] from Onion Networks, which not only enables users to participate in a distributed grid to provide fast downloads over broadband networks, but it also allows progressive playback of media files. This means that users can watch videos while they are still being downloaded.

While both Knotiki and Swarmstreaming support the identification of media content on the P2P network, they do not support an advanced search function like that of MediaAlert [4] to locate clips of interest that match a user profile. Our solution adopts a hyprid approach: the personal media alert client on each peer consults the global content index servers (hosted by the service providers) to find out which peers have media segments that match keywords of interest, the information is then used to initiate a streaming session to download the required clip. For clips of high interest, the streaming solution allows each peer that started the download process to act as a server to begin serving the same clip to other interested peers. The built-in THEX (Tree Hash Exchange Format) self-healing technology in Swarmstreaming also detects and automatically repairs any corruption that might occur during a file transfer.

3.2. Micropayment Scheme

To respect digital rights and to promote fair sharing, obtaining a media clip from a peer may require a *payment* be made to both the content owner and the peer that contributes its processing time and bandwidth. Since most of our media alert news clips range from 30 seconds to 3 minutes, the payment to content owner is potentially quite limited and may not justify the overhead cost of transaction fees. In addition, the service provider may or may not have direct business relationships with the particular content owner since we are potentially reaching millions of channels and content owners.

A good reference for pricing is today's iTunes video download service provided by Apple. A user downloading a typical TV show averaging about 30 minutes is charged U.S. $1.99. In our case, a typical media clip which is about 2 or 3 minutes may incur a cost less than a quarter ($.25), which puts such transactions in the realm of micropayments. In addition, since the resource usage for such a media clip incurs only a small overhead on the peer (but not negligible when many requests were made for the same clip), we need

a mechanism with even less overhead than a typical micropayment system. We are currently looking at two payment solutions for this problem:

- *Peppercoin* [10]: In this probabilistic payment scheme, a fraction of the micropayments received by the content owner are determined, via a procedure known as *cryptographic selection,* to qualify for upgrade to a macropayment. In this manner, the merchant transforms a large collection of small micropayments into a smaller collection of macro-payments, of the same total expected value. The content owner pays much less for processing the resulting macro-payments, since there are fewer of them.
- *WhoPay* [11]: WhoPay is an anonymous, fair, secure, and scalable P2P payment scheme that allows each peer to purchase *coins* from a trusted source (broker) and then issue coins of their own for payments among peers to reduce transaction load on the broker. Simulation results of WhoPay indicate that as peer availability increases, broker load decreases and peer load increases; coin transfers among peers dominate the peer load.

4. SYSTEM PROTOTYPE

We implemented the media processing and indexing modules on Microsoft Media Center Edition machines with 2.4 GHz hyper-threaded CPU, which are common in the consumer market. CC extraction, CBS, and transcoding to low bit rate video are conducted in real time while the video is being acquired, and video index is created within a couple of minutes after the program is aired. ASR requires 0.75 times real time, transcoding into VHS-quality video needs 1 times real time, and transcoding into SD video demands 3 times real time. The more computationally expensive processing components have to be done offline and they are optional depending on the peers' processing capabilities.

Fig. 2. Watching PMA on a Windows MCE machine

P2P media transfer including streaming was developed, and we plan to integrate payment schemes in the system. We also consider using a more mature P2P system for content delivery. Figure 2 is an example of our PMA system showing how a home user can watch media clips on his Windows Media Center PC. The left side panel shows the topic categories. The user is browsing clips that match his keywords in the category "Companies". The user can switch among categories and watch different clips easily using his remote control at hand.

5. CONCLUSION AND FUTURE WORK

We presented a system that aggregates personalized content individually obtained from TV broadcast programs and allows participating peers to find and share mutually interesting news clips through the support of P2P networks. To encourage fair trading of the content and discourage illicit content sharing, we proposed a micropayment scheme which enforces content consumers to make payments to both the peers who provide the content, and the original content owners. Since the cost of each transaction may be relatively small, our micropayment scheme reduces the number of payment transactions while ensuring full payment over a long period of time. Our system may be built upon an existing mature P2P delivery platform for content distribution.

6. REFERENCES

[1] Z. Liu, D. Gibbon, and B. Shahrary, "Multimedia Content Acquisition and Processing in the MIRACLE System," *CCNC 2006,* Las Vegas, Jan. 2006.

[2] D. Gibbon, Z. Liu, and B. Shahrary, "The MIRACLE Video Search Engine," *CCNC 2006,* Las Vegas, Jan. 2006.

[3] LocationFreeTV, http://products.sel.sony.com/locationfreetv, Sony corporation.

[4] B. Wei, B. Renger, Y. Chen, R. Jana, H. Huang, L. Begeja, D. Gibbon, Z. Liu, and B. Shahraray, "MediaAlert—A Broadcast Video Monitoring and Alerting System for Mobile Users," *The 3rd International Conference on Mobile Systems, Applications, and Services (MobiSys),* Seattle, June 2005.

[5] L. Begeja, et al, "eClips: A New Personalized Multimedia Delivery Service," *Journal of the IBTE*, Vol. 2, Part 2, April 2001.

[6] A. Wierzbicki and K. Goworek, "Peer-to-Peer Direct Sales," *IEEE International Conference on Peer-to-Peer Computing,* Konstanz, Germany, August 2005.

[7] B. Shahraray, "Scene Change Detection and Content-based Sampling of Video Sequences," *Digital Video Compression: Algorithms and Technologies 1995,* Proc. SPIE 2419, Feb. 1995.

[8] Kontiki, http://www.kontiki.com.

[9] Swarmcast Accelerator, http://swarmcast.net/swarming.

[10] R. L. Rivest, "Peppercoin micropayments," *Proc. FC '04*, A. Juels (Ed.), LNCS 3110, pp. 2-8, Springer 2004.

[11] Kai Wei, Yih-Farn Chen, Alan J. Smith, Binh Vo, "WhoPay: A Scalable and Anonymous Payment System for Peer-to-Peer Environments", *Proceedings of the 26th International Conference on Distributed Computing Systems*, Lisboa, Portugal, July 2006.