To be presented at *The Second Workshop on Personalization in Future TV (TV '02)*
at the *Second International Conference on Adaptive Hypermedia and Adaptive Web-based Systems (AH 2002)*

Real Time Television Content Platform:

Personalized Programming Over Existing Broadcast Infrastructures

Kelly L. Dempski

Accenture Technology Labs
161 N. Clark St.
Chicago, IL 60601 USA

Kelly.L.Dempski@accenture.com

Abstract. The success of personalized and interactive television programming has been largely hindered by the prohibitive cost of rolling out the necessary infrastructure to a large population. Most interactive TV infrastructures are based on a thin client model and it is very costly to update the transmission infrastructure. It is also costly to supply each customer with a set top box capable of receiving the interactive programming. As content providers are struggling to implement these architectures, the adoption of powerful next generation consoles is sharply increasing. This, combined with the recent introduction of inexpensive consumer electronics devices such as TiVo and ReplayTV, creates an opportunity to exploit media creation and manipulation capabilities in a "thick client" model using the existing broadcast infrastructure. This paper explains the concept of the Real Time Television Content Platform and discusses how this concept can deliver custom experiences with very few infrastructural changes.

Keywords. Personal Video Recorder (PVR), Data casting, Digital asset management, Personalized content.

Introduction

There are many that believe that the value of convergence is in the ability to create personalized programming [1]. The definition of "personalized" can be quite broad, ranging from content that is easily accessible on-demand, to content that is generated on the fly based on user preferences and interaction. Currently, the greatest impediment to delivering any type of personalized media is the delivery mechanism itself. Interactive services are limited by the relatively low performance of the current generation of set top boxes. On demand services are limited by the lack of ubiquitous broadband service.

Since its inception, television has used a broadcast model to reach the widest audience for the least cost. This model is very appropriate in a world of limited, generalized content, but it leaves very little room for the delivery of personalized or customized content. Over time, content providers addressed the varied tastes of viewers by supplying more and more broadcast channels. Currently, cable and satellite networks feature channels such as the Cartoon Network, the History Channel, and others that appeal to niche interests. This is a double-edged sword. These channels supply a wide range of content, but they also create churn by forcing the viewer to surf through hundreds of channels to find something they want to watch. There is more content available, but it has become more difficult to find it.

Content providers have addressed this problem by attempting to create and deploy new transmission paradigms that break out of the broadcast and channel model. One approach has been to change the infrastructure itself. An example is Video on Demand (VoD) over the Internet. The switch to a data networking model creates more opportunities for narrowcasting, but the difficulty here is expense [2] The cost of rolling out the backend infrastructure, the transmission infrastructure, and the client hardware is frequently much higher than the foreseeable revenues from video on demand. These new systems also introduce new technical problems such as latency, scalability, and quality of service. Also, high bandwidth networks are nowhere near as ubiquitous as cable or satellite service. These are very difficult problems, and they must be overcome for the new systems to gain acceptance.

The focus on new transmission methods allows the content provider to deliver services to homes through low performance "thin client" set top boxes. The dependence on thin-client solutions stems from the historical fact that televisions have until very recently been very thin clients. Also, "thick client" set top boxes have been expensive and difficult to roll out to the mass market. Therefore, any changes to services have necessarily been done to the backend or transmission infrastructure. This paper proposes an alternative solution. Instead of making changes to the transmission channel, we suggest leveraging the power of a new set of inexpensive consumer electronics devices such as gaming consoles, home media servers, and personal video recorders (PVRs). We have generalized the functionality of these devices to a hypothetical composite device dubbed the Real Time Television Content Platform (RTTCP). At present, no single device incorporates all of the functionality of the RTTCP, but several devices such as Sony's PS2 [3] and Nokia's MediaTerminal [4] have begun incorporating these features into one platform. There are many companies making significant investments in the development of very inexpensive, yet very powerful consumer devices. A device like the RTTCP would be capable of creating and manipulating media at the destination point. This allows the content provider to broadcast a set of primitives to everyone over the broadcast channel. The RTTCP can receive these primitives and use them to create personalized content. The result is the ability to create narrowcast content over the broadcast channel, resulting in very inexpensive personalized content. The remainder of this paper will discuss two instantiations of this approach using two very different types of media as the basic primitive.

The First Prototype: An Interactive Car Commercial

Many television commercials feature products that are actually computer-generated models. In many cases, it is less costly to manipulate a virtual model than to film an actual product. As with film, the computer has become a very powerful tool to easily manipulate objects and generate special effects in commercials. Currently, the virtual model is used to produce one commercial that is then broadcast to everyone because it is very time consuming to render realistic scenes. This is changing. Today's newest graphics chips are capable of rendering realistic scenes very quickly. The first example of RTTCP content demonstrates the 3D graphics processing power found in next generation gaming consoles such as the Microsoft Xbox [5] or the Sony PS2. Both are very inexpensive consumer electronics devices capable of producing very realistic imagery in real time. This first prototype shows how this power is used to generate custom content based on 3D models and other media that are broadcast to the client machine. The RTTCP is used to generate a custom car commercial based on the viewer's interests and interactions. With the appropriate architecture, the RTTCP can create virtual commercials that can replace broadcast content when appropriate.

Before any content can be rendered, a system must be devised that can interject the custom content into the normal flow of the broadcast stream. This is done very simply through the thin data channel of the standard television stream. This data channel exists in the vertical blanking interval of the television signal and is used for data services such as closed captioning and Intel's Intercast. In this first scenario, the data stream includes command messages that can be received and interpreted by the RTTCP. These command messages tell the RTTCP that a commercial break is coming and where it can download the appropriate content. When the RTTCP receives this message, it can contact the provider and download the appropriate media primitives. Alternately, the primitives could be sent in the data stream if the primitives were small enough. If the receiver were not RTTCP compliant, it would ignore the commands and do nothing. Figure 1 below shows how this might work.

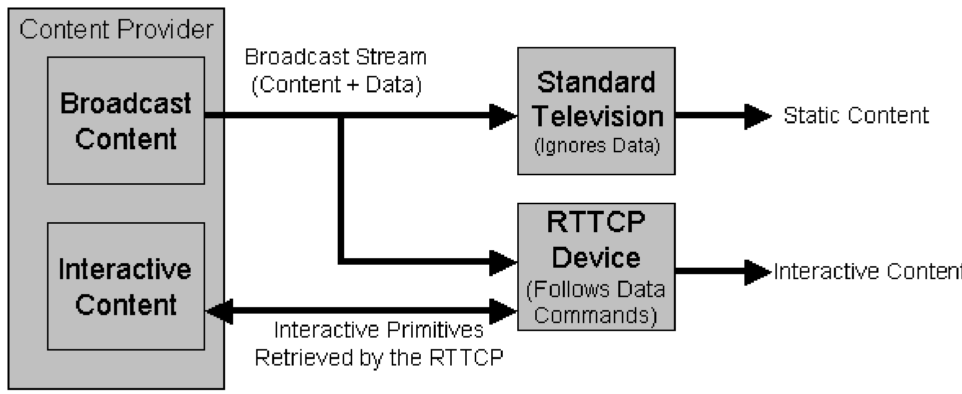

Figure 1: Broadcast data drives the RTTCP and is ignored by other devices.

When it is time for a commercial break, the broadcast channel sends a generalized, broadcast version of the commercial to maintain compatibility with older devices. The overall system does not force viewers to upgrade their hardware. Viewers with standard televisions will see a standard commercial. Viewers who have an RTTCP device will see something very different. When the programming switches to a commercial break, the RTTCP will take over the television set and render a custom commercial based on the 3D models that it has obtained. If the viewer has no interest in the commercial, the rendered scene will continue for the duration of the commercial and then end. It will be very similar to the standard broadcast version. If the viewers choose to, they can interact with the content via their remote controls. Figure 2 below shows scenes from a virtual car commercial.

Figure 2: Screenshots from a very realistic virtual commercial rendered in real time by the RTTCP.

In this commercial, a viewer can interrupt the normal flow of the ad and change the color of the paint, change the color and style of the interior, and try out different accessories. The commercial is fully interactive and fully configurable. It is not a series of prerendered video scenes. When the viewer is done, the commercial shows them the suggested price based on their selections, ends, and the RTTCP rejoins the standard broadcast stream. If the viewer has spent a long time experimenting with options, the RTTCP will record the broadcast stream to allow for time shifting.

The ability to render virtual imagery on a television screen is not a new idea. This has been done since the earliest game consoles. However, recent advances in graphics technology had made it possible to render extremely realistic scenes in real time on very inexpensive equipment. Figure 3 below demonstrates this with side by side screenshots of a real NASCAR race and a virtual scene from NASCAR Heat, a recently released video game for the Xbox.

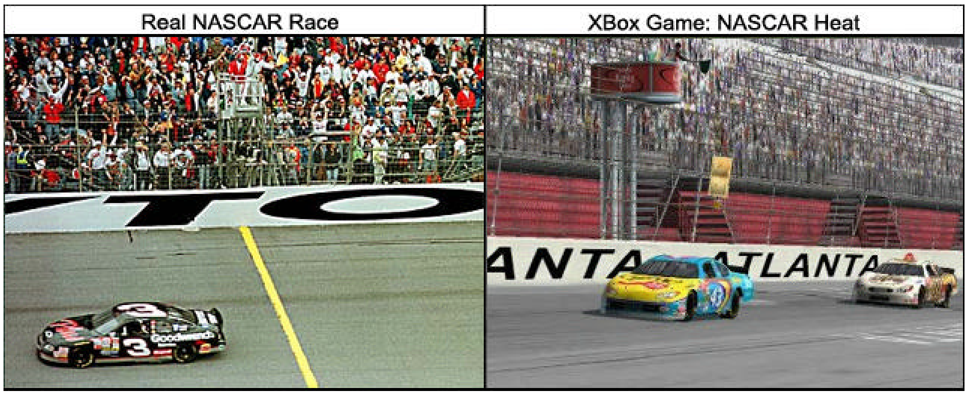

Figure 3: Real cars and their virtual counterparts.

This new technology creates the opportunity to render content that rivals real content in terms of quality and fidelity. These graphics are no longer blocky and cartoonish. Interactive rendered content of many products will have the same production quality as a static video image. The rendered content will be far more interesting because it can be modified in real time.

The Benefits of Client Side Rendering

The ability of the viewer to interact is only half of the equation. Most viewers watch television in a sedentary way. They do not want to be forced to continually interact. With this in mind, the car commercial prototype has been built to interact with the viewer in very subtle ways. Normally, commercials are very generalized. Even if a viewer wants a given product, chances are that they want a slightly different model or style than what is shown in the commercial. The RTTCP enabled commercial remembers the viewer's options and shows that exact model on every subsequent showing of the commercial. A viewer need not interact every time the commercial is shown. Instead, the commercial is tuned to show them exactly what they are interested in. In effect, it continually reminds them of how much they like the product.

The ability to render completely customizable content could completely change the nature of commercials. Commercials today are mostly static. They cannot learn from the viewer. Content rendered on the fly can be very dynamic. For instance, if a viewer picks all of the performance options in a given commercial, the next showing of that commercial can portray a higher performance model of the car. Likewise, once the system knows that the viewer like a blue car, it can show every model in blue. This idea also extends beyond simple advertising. Once viewer buys a car, the commercial time slot could be used to render scenes that give maintenance tips or other content that serves to build the relationship between the vendor and the customer.

This car commercial is just one example of one of the core capabilities of the RTTCP. This ability to create individualized content enables narrowcasting in the most literal definition. One of the biggest obstacles to true narrowcasting is scalability. The limitation to scalability is typically in the transmission medium. It is very difficult to deliver many custom streams in a timely manner. This architecture is based from the beginning on using the broadcast infrastructure. Broadcasting content to millions of viewers is a solved problem. Broadcasting primitives to a powerful device creates a system that is both extremely scalable and highly personalizable. This system is capable of delivering unique experiences to a very large number of viewers with almost no infrastructural changes.

Much of the discussion above has ignored forthcoming advances such as MPEG7 and other media delivery innovations. These innovations may augment portions of the architecture and become core components of the RTTCP. They have been omitted from the discussion to show that the RTTCP concept can work independently of these new technologies and standards. The second prototype continues on this thread. The car commercial relies on the RTTCP's ability to create media. The second prototype shows how it can manipulate media.

The Second Prototype: MyNews

The car commercial is a demonstration of how the graphics power found in the current generation of gaming consoles can be harnessed to create new media. However, in many cases it is preferable to have actual video of actual events and objects. In this case, personalization can be achieved through manipulation of the media. To do this, the RTTCP includes a second piece of technology found in new consumer electronic devices – the personal video recorder. Personal Video Recorders (PVRs) such as TiVo [6] and ReplayTV [7] are a relatively new phenomenon that combines inexpensive computing and storage in a consumer device that is capable of recording broadcast content according to a viewer's preferences. Users typically use the device to create their own personal channels. PVRs consolidate programming by recording shows from multiple channels and presenting them through one consistent interface. In effect, the PVR strips away all notions of channel or timeslot, allowing users to watch what they want, when they want.

The MyNews prototype takes the idea of a personalized channel one step further. The goal is to create a PVR enabled experience that delivers a single personalized program in addition to the personalized set of programs. It does this by recording small segments based on viewer preferences and compiles these segments into one program. This approach is well suited to news programming, but can also add value to sports, music, infotainment, reality shows, and other forms of content that can be naturally segmented.

The opportunities for personalization go far beyond simple compilation, as the viewer preferences can also be used to drive program flow, targeted advertisements, etc. This concept drives personalization to the next level – moving from personalized channels to personalized programs using the capabilities of the PVR and the existing broadcast infrastructure.

The MyNews prototype is built on the idea that a PVR based agent can be used to assemble a customized news program based on broadcast content. Currently, both PVRs and content broadcasters are closed systems. Therefore, this prototype assumes that the news broadcaster has some control over the software on the RTTCP. The MyNews agent would run as an added service on top of the usual PVR functionality.

This agent stores viewer preference data (described in the next section) on the local PVR. It then "watches" the broadcast news stream from the news source. News content is broadcast over the same transmission medium as usual. In this way, the same content also feeds viewers who do not have the agent or the PVR. The news provider includes codes in the data segment of the broadcast that define the subject matter of the current story. The agent uses this information to record and compile only the stories of interest. It also uses the same information to record segues between stories to give the compilation the proper flow and production quality. The agent can operate in two modes. It can record all the programming and sort the news stories according to preference. It can also record only the news items that fall within a certain threshold and create a concise news program of a user specified length. In either case, the recorded program includes a menu for random access to all the news items. Finally, the preference data can also be used to insert commercials that more closely match the viewer's interests.

The end result is a news program that includes only the stories that are of interest to the viewer. The compilation is customized, but the snippets are blended together such that the quality of the experience is the same as any single produced show. This system enables the content provider to deliver high quality content over existing broadcast channels and uses local manipulation to customize the content. The following sections describe the components that make this possible.

A Look at the MyNews Preference Interface

Before one can talk about personalized programming, one needs to figure out the best way to get personal preferences from the user. This involved striking a balance between the relatively broad and deep taxonomy of news subjects and the need to create an interface that was easily navigable on a television. The result was a sparse interface that represents news topics in terms of categories, places, and people. Screenshots of this interface are shown in Figure 4 below.

Figure 4: Three screens from the MyNews user interface. Items are rated using a simple remote control.

A full taxonomy of news items could be fairly broad and detailed. It might have thousands of different items. For most content providers, the development of this taxonomy is an ongoing problem. Even if the taxonomy were complete, forcing the user to input preferences based on the full taxonomy would be overwhelming for the typical user. For these reasons, the preference interface presents only a small representative sample of the taxonomy to the user. The user rates items of interest on a –2 to +2 scale. Any items that are not rated are assumed to have a neutral rating of 0. In cases where the rated item represents a deeper hierarchy, all child items are given the same rating. These preferences provide the first set of data needed to build a viewer profile. The system builds a richer profile using two additional techniques.

First, the system infers demographic data from the user’s selections. This demographic data can be used to infer a more detailed profile. For example, a user who picks “Business” and “Alan Greenspan” may be more interested in US stock market data than someone who picks “Business” and “Tony Blair”. Likewise, the person who picks “Paul McCartney” is probably of a different demographic than the person who picks “Britney Spears”. The system can use these inferences to rate subcategories and to build a more detailed profile of the user [8]. Because of this, many of the items in the interface might be chosen to develop a demographic profile more than an explicit rating about that one topic.

Secondly, the system can refine its detailed profile by monitoring user interactions. A user’s actions, such as skipping a given story, can drive the system to update the profile. For example, one of the categories on the interface is “Sports”. Internally, the system contains the subcategories of “baseball”, “soccer”, etc. When the user rates “Sports” as +2, that rating is applied to all subcategories. If the user skips every baseball story the system decrements the preference for “Baseball”. Eventually, usage patterns will tune the profile to accurately reflect the user’s preferences.

This shallow system of preference gathering satisfies two design goals. First, it enables the system to develop a rich profile without asking too many questions. In fact, the user could skip the preference interface entirely and rely on usage patterns to build the profile. Secondly, it acknowledges the fact that content tagging is an ongoing problem. This system allows an evolution from a sparse backend taxonomy to something richer and more detailed over time. This will allow content providers to roll out applications before final systems are in place. Systems can begin to generate revenue before the taxonomy is actually finalized.

Matching Preferences to Content

Compiling client side preferences is only half the battle. The content itself must be tagged using the same taxonomy. Many news providers are currently working on content tagging systems for the purposes of archiving and retrieving their assets. This system could also be used for a MyNews type of application. In order to match the user preferences with the content, information about the current news story would need to be supplied along with the content. In many cases, this information already exists in electronic form and could be included in the broadcast stream. At the very least, some metadata exists in the form of production and scheduling directives. The prototype application reads these tags and matches them against the user preferences.

Currently, the prototype uses simple vector dot product techniques to match content to the user's preference profile. The method is very simple, but seems to be very effective in tests involving several hours of tagged broadcast news footage. If needed, the matching scheme could be updated as tagging and preference data becomes more complex. Currently, the matching methods do not employ any collaborative filtering to help develop preference or matching data. This could lead to improvements or more interesting applications.

One final point to be addressed is how to match the preferences to an individual viewer in a shared television environment. This problem can be solved in a number of ways. One way would be a simplified "login" using a simple set of buttons on a remote control. The MyNews demo largely ignores this problem. The assumption is that the preferences are household preferences. In most cases, a family or a couple will watch the evening news together. In this case, the preferences should be the preferences of the set of viewers instead of a singular viewer.

The development of both the indexing systems and the preference systems is an ongoing task. There is still much research to be done to determine the best taxonomies, indexing methods, and retrieval methods. However, the MyNews prototype is flexible enough to accommodate an evolving tagging scheme. A first release might rely on a very shallow scheme that concentrates only on high level categories. Over time, further details could be added and the agent could be updated accordingly over either a network or with additional data on the broadcast stream. The next generation of PVRs will be network-ready. One advantage of a networked device is that the software can be easily updated by the service. The tagging and matching schemes can continually evolve without requiring the user to update the hardware or manually patching their software. It is noteworthy to mention that evolving broadband networks may ease the transport of content and preference data, but the broadcast channel will remain the most appropriate and ubiquitous transport for the video content for the foreseeable future.

Value to the Consumer and Content Provider

Personalized content is always valuable consumers because it enables them to consume better content in less time. This is especially true with news content because of the possibility of a low signal to noise ratio. A system like MyNews creates a better and more informative experience for the user because it strips out unwanted content and allows the viewers to concentrate on what they feel is most important.

This creates added value for the content provider in terms of viewer loyalty and brand appeal. This could be translated into subscription-based models where the general broadcast is available for free, but the personalized content costs a small subscription fee. This is analogous to the current model for the TiVo service. In this context, there are many possible business models that could focus on getting revenue from the consumer.

Another model is to focus on the more traditional revenue routes such as advertising. The existence of a viewer profile (and inferred demographic data) creates many opportunities for targeted advertising. There are many possibilities for business models that draw more revenue from the advertiser. These types of models concern privacy advocates because this could lead to abuse of the preference data. However, all the content manipulation occurs on the client PVR. Therefore, it is technically possible to create targeted advertising without the preference data ever leaving the viewer’s home. This is demonstrated by the car commercial prototype. The car can be modified and tuned to the viewer’s preferences without the preference data ever leaving the home. This would complicate the revenue model, but it does create possibilities that do not necessarily abuse privacy.

Future Work

The capabilities of a platform such as the RTTCP represent new challenges in the areas of content tagging, user profiling, privacy, usability, and a host of other areas. Future development of the RTTCP will concentrate on developing more applications that highlight the capabilities and provide a test bed for solving these new challenges. There is also a significant amount of work to be done in educating the content providers about the capabilities that are made possible by this new breed of inexpensive "thick client" devices. The consumers have already begun to embrace this technology, but the adoption and uptake of the RTTCP will depend heavily on the acceptance of the content providers.

Conclusion

The MyNews and car commercial prototypes are two of several Accenture prototypes that fall under the broader project name "Real Time Television Content Platform". These prototypes concentrate on how client side processing can be used to augment current broadcast infrastructures. MyNews is an example of how quality personalized content can be supplied with very little cost by leveraging devices that are already in the homes of the content consumers. The car commercial demonstrates the power of the device to actually create new content at the client side. These two capabilities applied separately or in combination create many new opportunities for content providers. These prototypes are not meant to singular examples of the technology. They represent the most basic examples of how low cost client side technology can be used to create rich personalized multimedia experiences.

References

1. Stroud, Jim. 2001. TV Personalization: A Key Component of Interactive TV. *The Carmel Grou*p.
2. Hua, Kien A., Ying Cai, Simon Sheu. 1998. Patching: A Multicast Technique for True Video-on-Demand Services. *ACM Multimedia 98*.
3. http://www.us.playstation.com/
4. http://www.nokia.com/multimedia/mediaterminal.html
5. http://www.xbox.com/
6. http://www.tivo.com
7. http://www.replaytv.com
8. Gena, Cristina, Liliana Ardissono. 2001. On the Construction of TV Viewer Stereotypes Starting from Lifestyles Surveys. Workshop on Personalization in Future TV 2001.