Activity Recognition for the Digital Home

Jeonghwa Yang, Georgia Tech
Bill N. Schilit, Google Research
David W. McDonald, University of Washington

Patterns mined from home networks can support smarter applications.

Home networked devices enable a wide range of daily activities including multiplayer gaming, movie downloading, and music streaming as well as modern conveniences such as home automation, wireless networking, and Internet access. Nevertheless, for most of us the futuristic digital home we see in movies isn't a reality.

We spend lots of time figuring out how to get our devices to do what we want and to keep them properly working and tuned. Ideally, these devices should support as well as enable the activities we like to do. For example, a user should be able to request "my favorite radio station" rather than have to input something like "192.168.1.100."

One step toward realizing a smarter digital home is efficiently modeling and recognizing human activities. We've developed a prototype system to detect various digital media and information access activities using the consumer electronic devices people already have. Cameras or other hardware sensors aren't necessary; instead, the network itself is the sensor, and the system monitors associated data flows among the devices and uses a template to match the data against generic classes of activities.

HOME ACTIVITY RECOGNITION

Activity recognition is a key feature of many ubiquitous computing applications ranging from office worker tracking to home healthcare. In general, activity recognition systems unobtrusively observe the behavior of people and characteristics of their environments, and, when necessary, take actions in response—ideally with little explicit user direction.

In the home environment, such systems can, for example, remind users to perform missed activities or complete actions (like taking medicine), help them recall information, or encourage them to act more safely.

Our digital home research has focused on recognizing common activities including

- browsing the Web,
- reading an online newspaper,
- watching movies and Internet TV,
- listening to music and Internet radio, and
- playing networked console games.

These activities might not be as important as monitoring medication usage, but they represent areas where people struggle with technology.

One application of home activity recognition is to observe users' interests based on their interaction with digital media and provide reasonable suggestions for future activity. For example, an application could inform you that a presidential debate is on TV based on your previous viewing of such debates.

Patterns mined from home activities can be used to support a wide range of similar over-the-shoulder applications. For example, such information could help define defaults such as which rooms to play music in or what stations to "preprogram" on Internet radio; it could likewise help users refind familiar songs on their media players, much as the history feature on Web browsers makes it easier to refind interesting websites.

Home activity recognition and data traffic monitoring can also help people understand complex network behavior—say, why they can't get an Xbox Live connection or why Internet access is slow (perhaps someone is watching Apple TV in the living room). Information about household activities can even be used to recommend changes in behavior—for example, to reduce TV viewing and spend more time playing aerobic games on the Wii.

These examples only hint at what's possible. Several researchers are exploring other aspects of this problem space. For example, Monika Henziger and colleagues describe how digital TV closed captioning could be used to generate Web queries ("Query-Free News

Search," *World Wide Web,* vol. 8, no. 2, 2005, pp. 101-126).

DIGITAL HOME NETWORKS

During the past several years, the computer and consumer electronics industries have introduced numerous home-networked communication and entertainment devices. Network-attached disks store and stream audio and video to PCs and network stereos, phones and game consoles connect through home routers to the Internet, and networked printers and picture frames print and display digital photos from laptops and Wi-Fi cameras.

All these devices support networking capability using wired (IEEE 802.3) or wireless (IEEE 802.11 a/b/g) technologies and IPv4. Many also use HTTP or the Real-Time Transfer Protocol for media transport and higher-level protocols such as universal plug and play for device discovery and UPnP AV (audio and video) for media control and management.

An important feature of UPnP is that it provides a standard way of describing devices on the network. Each UPnP device publishes an XML description that includes the name, manufacturer, model, and serial number as well as a list of embedded services—such as "rendering" image or audio files—available through URLs. Because each of the service classes is standardized, it's possible to determine which devices can work together for various uses even if they have different manufacturers.

HOME NETWORK MONITORING

In our research, we use the UPnP discovery protocol to detect UPnP-enabled devices on a home network and then create a UPnP *instance* table that contains a description of each device and its capabilities along with its physical (MAC) address.

We augment this static description of the network with a UPnP *operational status* table that lists which devices are sending packets to each other, the devices' streaming status (playing, paused, stopped), active media file names (like "lovesong.mp3"), and other information. The table contents change as devices connect to each other and play music, movies, and so on. To populate this table, we use both a low-level traffic flow monitor and a high-level UPnP event monitor.

A few home routers now can report data traffic using RFlow, a protocol based on Cisco Systems' NetFlow product for enterprises. RFlow records the MAC addresses for source and destination devices, along with a packet byte count, for a given period of time. Our experimental home network's RFlow-capable router sends samples to an SQL database server at uniform time intervals, indicating which devices have active data flows—for example, music streaming from a PC to a network stereo.

Home activity recognition and data traffic monitoring can help people understand complex network behavior.

Although RFlow provides information about data traffic between devices, it doesn't say much about that data's content. To obtain eventing information, we exploited UPnP event notification technology, which uses a publish-subscribe model to send status and control-variable changes between a media server or renderer to registered parties—this is useful for a remote-control point, for example.

We used Tools for UPnP Technologies, developed by Intel's Digital Home Group, as well as our own event subscriber application to capture media events for the operational status table. These events include the media content's format and uniform resource identifier, which can be used to obtain further information such as ID3 tags.

ACTIVITY TEMPLATES

The UPnP instance and operational status tables provide a medium-level description of home network transactions and capabilities. To make this information meaningful to people, we created high-level activity descriptions such as "read online newspaper," "watch Internet TV," "listen to Internet radio," "listen to music," "watch movie," "look at photo," and "play Xbox game." We define activities by media type, but other criteria such as device type could also be used.

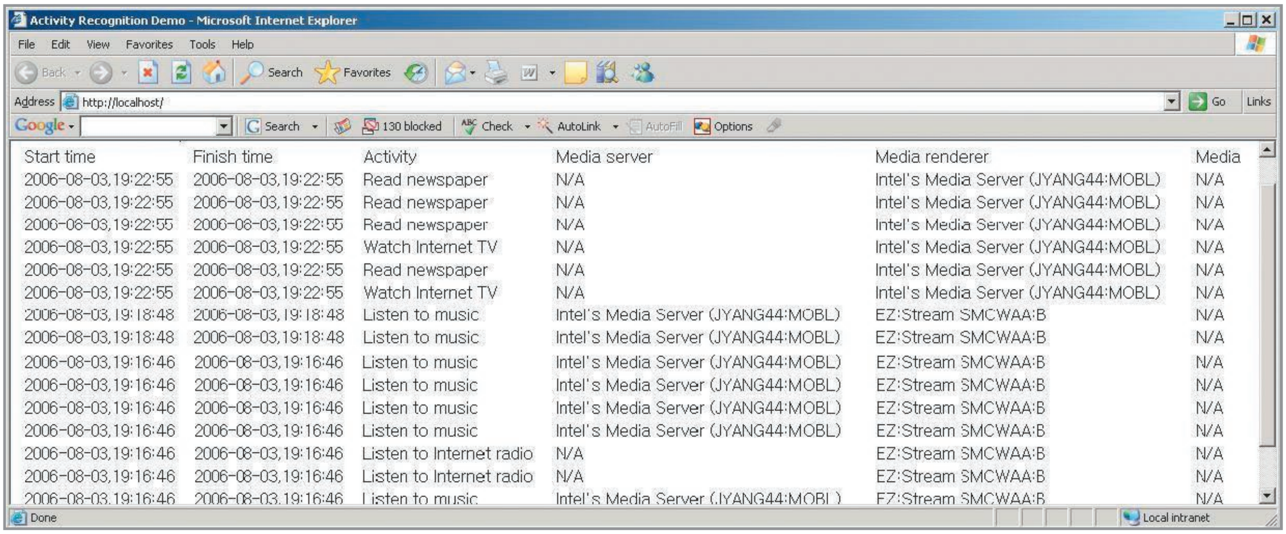
For each activity, we developed an *activity template* that lists attribute names and values including the source device, destination device, and media of data transferred along with numerous other parameters that characterize the activity duration and type.

We manually created our activity categories and templates, but it's possible to automatically generate activity tuples—for example, using potential device pairings based on client-server protocol information. However, user-meaningful activity labels would still need to be added manually. A privacy filter could also obscure some or all details of activities.

RECOGNITION ENGINE

We also developed a *recognition engine* that infers activities based on network traffic data and UPnP device information matched with generic classes of activities defined in the activity templates. When the engine observes a new traffic flow, it analyzes the source and destination devices involved and the time stamp, then refers to the UPnP instance and status tables to determine which devices are a media server and renderer.

The recognition engine also examines application-level events involving those two devices around the time of the network flow and matches this data against the activity templates. It records matching templates at each sampling interval and combines templates from one interval to the next



Start time	Finish time	Activity	Media server	Media renderer	Media
2006-08-03,19:22:55	2006-08-03,19:22:55	Read newspaper	N/A	Intel's Media Server (JYANG44:MOBL)	N/A
2006-08-03,19:22:55	2006-08-03,19:22:55	Read newspaper	N/A	Intel's Media Server (JYANG44:MOBL)	N/A
2006-08-03,19:22:55	2006-08-03,19:22:55	Read newspaper	N/A	Intel's Media Server (JYANG44:MOBL)	N/A
2006-08-03,19:22:55	2006-08-03,19:22:55	Watch Internet TV	N/A	Intel's Media Server (JYANG44:MOBL)	N/A
2006-08-03,19:22:55	2006-08-03,19:22:55	Read newspaper	N/A	Intel's Media Server (JYANG44:MOBL)	N/A
2006-08-03,19:22:55	2006-08-03,19:22:55	Watch Internet TV	N/A	Intel's Media Server (JYANG44:MOBL)	N/A
2006-08-03,19:18:48	2006-08-03,19:18:48	Listen to music	Intel's Media Server (JYANG44:MOBL)	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:18:48	2006-08-03,19:18:48	Listen to music	Intel's Media Server (JYANG44:MOBL)	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:16:46	2006-08-03,19:16:46	Listen to music	Intel's Media Server (JYANG44:MOBL)	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:16:46	2006-08-03,19:16:46	Listen to music	Intel's Media Server (JYANG44:MOBL)	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:16:46	2006-08-03,19:16:46	Listen to music	Intel's Media Server (JYANG44:MOBL)	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:16:46	2006-08-03,19:16:46	Listen to music	Intel's Media Server (JYANG44:MOBL)	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:16:46	2006-08-03,19:16:46	Listen to Internet radio	N/A	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:16:46	2006-08-03,19:16:46	Listen to Internet radio	N/A	EZ:Stream SMCWAA:B	N/A
2006-08-03,19:16:46	2006-08-03,19:16:46	Listen to music	Intel's Media Server (JYANG44:MOBL)	EZ:Stream SMCWAA:B	N/A

Figure 1. Web-based network activity visualizer. An activity history log combines multiple short, similar activities into a single human-understandable session.

so long as the source and destination devices are identical.

When a template no longer matches the network monitoring data, the engine maintains it for a "maximum pause time" before writing it to an activity history log. We do this to combine multiple short, similar activities into a single human-understandable session. Similarly, a template must match for a "minimum duration" to be written to the log.

The activity history log can take various forms—for example, it can be displayed on a computer or TV much like a program schedule. For our work, we developed the real-time Web-based activity visualizer shown in Figure 1 using JavaScript and PHP.

Modeling and detecting human activities using home network data traffic patterns is a step toward creating a smarter digital home that provides useful information and services in a more user-friendly way.

Primarily using UPnP technology, we created a network map of devices and their characteristics and recorded dynamic data flows along with high-level events. Based on these static and operational values, we manually defined and automatically matched human-level activity templates. We believe that a learn-by-demonstration method or simply sharing activity templates among a large user community could improve the manual steps of our process.

Jeonghwa Yang *is a PhD student in the School of Computer Science at the College of Computing, Georgia Institute of Technology. Contact her at jeonghwa@cc.gatech.edu.*

Bill N. Schilit *is at Google Research. Contact him at schilit@computer.org.*

David W. McDonald *is an assistant professor in the Information School at the University of Washington. Contact him at dwmc@u.washington.edu.*

Part of the work described in this article was done while the authors were with Intel Research.