Fast, Detailed Inference of Diverse Daily Human Activities

Matthai Philipose, Sunny Consolvo, Tanzeem Choudhury, Kenneth Fishkin, Mike Perkowitz, Ian Smith
Intel Research Seattle
sharp@intel-research.net

Dieter Fox, Henry Kautz, Donald Patterson
Dept. of Computer Science and Engineering
University of Washington
{fox, kautz, patterson}@cs.washington.edu

ABSTRACT

The ability to detect what humans are doing has long been a key part of the ubiquitous computing agenda. We demonstrate a system that can detect many day-to-day human activities in considerable detail. The system has three novel components: sensors based on Radio Frequency Identification (RFID) tags, a library of roughly 38,000 human activity models mined from the web, and a fast, approximate inference engine. We show how the inferences derived can be used to drive two applications from the eldercare space. The Caregiver's Assistant, targeted at the eldercare professional, automatically fills out entries from a state-mandated Activities of Daily Living form. The CareNet Display, targeted at family caregivers, provides up-to-date information about the elder's activities.

Keywords

Activity, inference, RFID, eldercare

INTRODUCTION

Detecting what day-to-day activity a person is doing is central to many ubiquitous computing applications. Examples include care giving software which monitors patients' activities, detects anomalies and prompts patients, best-known-method management software for detecting, training and enforcing superior methods for performing, and personal assistive software such as activity-based reminders and custom configuration of devices. Many of these applications require that computers track activities being performed accurately and in some detail, that many activities, activity settings and ways of performing the activities be tracked, and that tracking occur in real time.

The requirement of detailed, diverse and rapid detection of activities has been a difficult one to meet. In fact, using conventional sensors (such as vision, audio and accelerometers) that directly monitor the human performing the activity, only a small set of activities has been shown to be robustly detectable. An alternate approach has been to monitor the effect the human has on their surroundings e.g. devices used, doors opened and spaces accessed. These systems have been successful in recognizing a larger set of activities, although the diversity, accuracy and detail of recognition have been modest: they tend to use coarse-resolution sensors of the environment that conflate many activities, hand-built or hand-labeled models of activities that are therefore necessarily limited in number, and (in some cases) deterministic models of activity that are not particularly resilient to variation and error.

Our work follows the latter (environment-tracking) approach: we model activities by the sequence of objects used in them, and their duration of use. To overcome the limitations of previous systems, however, we use novel sensors based on Radio Frequency Identification (RFID) tags to track object usage at a fine granularity, mine formal models of activities from human-written descriptions of activities, and use approximate probabilistic inference to rapidly track activities in spite of various kinds of uncertainty. We also show how the activities so inferred can be used to drive two applications in the eldercare space.

We describe our contributions in more detail below.

SYSTEM DESCRIPTION

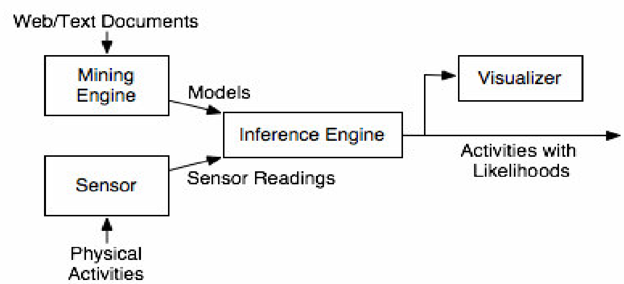

Figure 1: System Structure

Our system has the fairly conventional structure shown in figure 1. Given models of activities, and sensor readings, our inference engine deduces the likelihood of each model.

The usage model is as follows. Objects in the space where the activities may occur are tagged with RFID tags. People performing activities wear gloves or bracelets that have RFID tag readers built into them. As they perform their activities, the proximity of their hands to tagged objects they touch informs our engine that these objects are in use. The engine infers the activities in progress and periodically notifies clients of the most likely activities.

Radio Frequency Identification (RFID) Based Sensors

RFID tags are small 40-cent stickers with a built in antenna and circuitry. When queried by ambient readers, the tag harnesses energy from the querying signal to return an identifier unique to the tag. Our key contribution, which makes RFID useful in activity inferencing, is in adapting

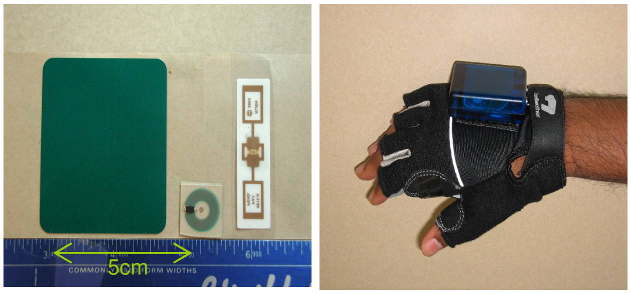

Figure 2: RFID Tags (L) and Glove-Based Reader (R)

RFID readers to enable object detection. For short-range (3-5in) tags we have produced a lightweight, wireless glove (shown in figure 2) that essentially detects touched objects. For long range (15-25ft) tags, we have developed a new technique that can detect when a tag moves, based solely on signal strength variation at the reader [2].

Model Library Mined from Text

Making Tea:

1. Fill a **teapot** from the **faucet**. Place kettle on the **stove** and boil.
2. Pour hot **water** into a **cup**, filling ¾ of the cup. Immerse **teabag** in cup for two minutes and dispose of teabag.
3. Add **milk** and **sugar** to taste.

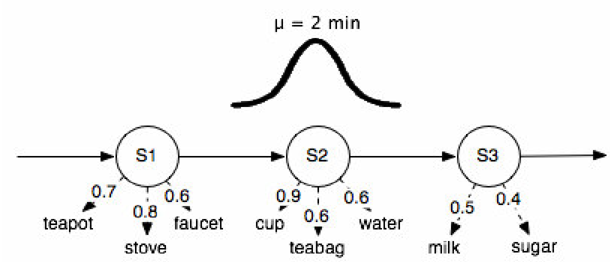

Figure 3: Human Directions (top); Mined Model (bottom)

We model activities as a descriptive name and a sequence of steps. Each step is associated with the objects that may be used during it, the probability of using each object, and (optionally) the duration of the step. The bottom of figure 3 shows the model for making tea. We note that human descriptions (such as recipes and how-tos) of activities mirror this structure: they have a title, followed by a number of steps. Each step mentions the objects involved and possibly a duration. A key contribution of our system is showing how to use very lightweight natural language processing techniques and term-similarity measures from the web to mine the models automatically [3]. By mining various web sites we have accumulated a library of roughly 38,000 models.

Inference

We compile our models of activities into Dynamic Bayes Nets (DBNs) and solve them approximately using a sampling-based technique called particle filtering. Our contribution [4] is in adapting techniques to account for durations of steps and partial orders between steps.

APPLICATIONS

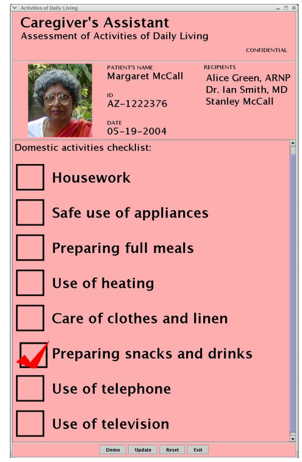

Figure 4: Caregiver's Assistant Digital ADL Form

We have applied our inference engine to the domain of care giving, eldercare in particular. We have focused on two applications. The Caregiver's Assistant (CA) is a digital version of a medical form known as the Activities of Daily Living (ADL) form. The paper version of the form is filled by many caregivers during visits to the patient. The form records the ability of patients to perform ADLs; recording is known to be tedious, error-prone and intrusive. The CA (figure 4 shows an interface) automates the process. A companion application, the CareNet Display [1] provides similar information to family caregivers. Early results from trials indicate that actual users find the display useful.

SUMMARY

We have described a combination of novel sensing, model mining and inference technologies that have the potential for enabling a range of activity-based ubicomp applications.

REFERENCES

1. Sunny Consolvo, Peter Roessler, Brett E. Shelton, Anthony LaMarca, Bill Schilit, Sara Bly. Technology for Care Networks of Elders. IEEE Pervasive Computing, Vol 3, no 2. June 2004
2. Fishkin, K., Jiang, B., Philipose, M. and Roy, S. I Sense a Disturbance in the Force. Ubicomp 2004.
3. Perkowitz, M., Philipose, M., Fishkin, K., Patterson, D. Mining Models of Human Activities from the Web. WWW 2004.
4. Philipose, M., Fishkin, K., Perkowitz, M., Patterson, D., and Haehnel, D. The Probabilistic Activity Toolkit: Towards Enabling Activity Aware Computer Interfaces. Intel Research Seattle Technical Memo IRS-TR-03-013, December 2003.