Activity Recognition for the Smart Hospital

Dairazalia Sánchez, Monica Tentori, and Jesús Favela, *CICESE*

A hidden Markov model uses contextual information to recognize user activities and provide opportunistic services to hospital staff.

Although researchers have developed robust approaches for estimating, location, and user identity, estimating user activities has proven much more challenging. Human activities are so complex and dynamic that it's often unclear what information is even relevant for modeling activities. Robust approaches to recognize user activities will require identifying the relevant information to be sensed and the appropriate sensing technologies.

In our effort to develop an approach for automatically estimating hospital-staff activities, we trained a discrete hidden Markov model (HMM) to map contextual information to a user activity. (For other approaches to estimating work activities, see the related sidebar.) We trained the model and evaluated it using data captured from almost 200 hours of detailed observation and documentation of hospital workers. Here, we discuss our approach, the results, and how activity recognition could empower our vision of the hospital as a smart environment.

The iHospital

The iHospital is our vision of a highly interactive smart environment saturated with heterogeneous computing devices. These devices range from handheld computers that can capture and access limited amounts of information, to PCs that can be used at fixed sites for longer time periods, to conveniently located semipublic displays that let users share and discuss information. The iHospital staff can interact with the environment using a set of specialized services that account for contextual information. To help realize this vision, we have developed several ubiquitous and context-aware applications such as a context-aware mobile communication system, a context-aware map, and an activity-based application that support the following functionalities.[1,2]

Awareness of people and artifacts

Hospital staff must navigate hospital premises to gather the information they need to conduct their work; this navigation is time-consuming. We developed and incorporated into the iHospital a context-aware application aimed at saving the staff time by making them aware of the presence, location, and status of artifacts and people, either using a digital floor map or an automatically updated list of people and artifacts. To introduce location awareness, we use a component that estimates users' positions throughout the hospital.[3]

Context-aware communication

Hospital staff can use contextual information as a trigger to retrieve or send messages. For example, a nurse can send a message to the physician responsible for a patient in the next shift when the patient's laboratory results are ready. The sender doesn't need to know which physician will be attending the patient or when the laboratory results will be ready. The application, on the basis of contextual information, automatically sends the message.

In addition, hospital staff can transfer information from public spaces to personal devices, share

Estimating Work Activities

Previous approaches to inferring people's activities range from estimating low-abstraction activities (such as chewing or sitting) to estimating more complex activities.

One approach accurately infers activities 73 percent of the time by detecting user interaction with particular objects.[1] This approach tags every item of interest with an RFID tag and reads the tags through an RFID-detecting glove. In a similar study—part of the Georgia Tech Aware Home project—researchers used sound recordings to monitor and infer activities in the home.[2] They placed arrays of microphones to infer activities such as chewing with 75 percent accuracy.

The SEER (Selective Perception Architecture for Activity Recognition) application estimates activities with a higher level of abstraction (such as when a user is attending a conference).[3] It uses audio, video, and computer interaction obtained from sensors distributed in an office environment to train a layered HMM. SEER obtained 99.7 percent accuracy.

Another study involving an office environment used layered HMMs to model individual and group actions.[4] It modeled individual activities using audio and video and used the individual interactions to estimate the group action. It identified individual actions with 89.7 percent accuracy and group actions with 70.3 percent accuracy.

Researchers have also proposed using production rules to identify complex activities of hospital workers.[5] The presence of artifacts and people with RFID tags can trigger the inference rules coded by a knowledge engineer. However, creating these rules is time-consuming and requires considerable expertise and knowledge of the setting.

Earlier, we tested an approach that uses a back-propagation neural network trained to map from contextual information (the location of hospital staff, artifacts in use, people with whom they collaborate, and time of day) to activities performed by hospital workers.[6] Thus, we used some neurons—location, artifacts, role, and time—as inputs and activity as the output. We took the code for each contextual variable from the results of a workplace study. We transformed the information collected into inputs and outputs to the corresponding neurons. For example, whenever an artifact or person is involved in certain activity, we assign a one to the input vector; otherwise, we assign a zero. This approach can correctly predict hospital workers' activities 75 percent of the time. Although this is sufficient for developing some context-aware applications, it isn't enough to support hospital work.

References

1. M. Philipose, K.P. Fishkin, and M. Perkowi, "Inferring Activities from Interactions with Objects," *IEEE Pervasive Computing*, vol. 3, no. 4, 2004, pp. 50–57.
2. J. Fogarty et al., "Predicting Human Interruptibility with Sensors," *ACM Trans. Computer-Human Interaction*, vol. 12, no. 1, 2005, pp. 119—146.
3. N. Oliver, A. Garg, and E. Horvitz, "Layered Representations for Learning and Inferring Office Activity from Multiple Sensory Channels," *Computer Vision and Image Understanding*, vol. 96, no. 2, 2004, pp. 163–180.
4. D. Zhang et al., "Modeling Individual and Group Actions in Meeting with Layered HMMs," *IEEE Trans. Multimedia*, vol. 8, no. 3, 2006, pp. 509–520.
5. H.B. Christensen, "Using Logic Programming to Detect Activities in Pervasive Healthcare," *Proc. Int'l Conf. Logic Programming*, Springer, 2002, pp. 185–199.
6. J. Favela et al., "Activity Recognition for Context-Aware Hospital Applications: Issues and Opportunities for the Deployment of Pervasive Networks," *Mobile Networks and Applications*, vol. 12, nos. 2–3, 2007, pp. 155–171.

information between heterogeneous devices, remotely monitor other computers, and share handheld applications. For example, as figure 1 shows, two colleagues carrying their PDAs and discussing a clinical case using the public display could seamlessly transfer information between their personal information space (for example, the PDA) and the shared space (for example, the public display). The user selects digital files on his or her PDA and can transfer them to a nearby device. The user picks the target device from a list, which the PDA generates based on proximity.

Other researchers have developed technology for helping hospital staff collaborate—such as when consulting with a specialist or requesting help in transferring a patient. For example, Vocera is a hands-free voice communication system that lets users contact fellow hospital workers by name, role, or location.[4] This set of ubiquitous and context-aware services, which run in several heterogeneous devices, let the iHospital augment the hospital staff's collaboration.

Adapted and personalized information

To provide relevant information to the hospital staff, iHospital services take into account contextual information, such as the user's identity, role, and location; the time of day; and the status of clinical information (such as the availability of lab results). For example, when a physician is in front of a patient's bed, the physician's PDA automatically displays the patient's medical record. The iHospital services also consider the users' context and personalize the information accordingly. For example, when a physician approaches a public display, it shows only the physician's patients, personnel calendar, messages, and the location of others with whom he or she might want to interact.

Figure 1. Two physicians discussing a clinical case. One manipulates information directly with a public display while a colleague remotely interacts with the display from his PDA.

Multitasking abilities

The iHospital services help the staff manage their activities and information in a centralized way[1] using an activity-based service such as mobileSJ. The device can define computational activities and identify associated information resources, such

Figure 2. Activity-aware information access to the patient's medical record.

as contacts, emails, and pending issues. We call these computational activities *e-spheres* (see figure 2). When a user switches between e-spheres, each e-sphere can quickly retrieve its workspace state (the display layout, status of resources such as documents, and windows' overlay order) and contextual information (such as open documents). In addition, mobileSJ lets users share activities and resources and communicate with each other through short-message service or phone calls.

Using HMMs to estimate activities

Despite the benefits that the iHospital services provide, carrying out a single activity typically involves using several services, each with different functionality and data presentation. For example, a physician discussing a clinical case with an intern might need to share the patient's medical record, highlight areas of an x-ray image, and consult a related journal article. This would require having both the physician and intern browse the documents needed to inform their discussion, launch and login to each application, and establish a session between their handhelds. Because an activity in hospital work lasts an average of five minutes,[3] much of this time could be spent launching the right service in their PDAs and setting up the computational infrastructure (such as the login for each application used) instead of focusing on the primary goal—the clinical care of patients.

However, taking into account contextual information detached from the user's current activity could lead to the presentation of services and information disembodied and dissociated from the users' current goal. One way to overcome such a difficulty is to provide ambient-intelligence environments with knowledge of the activities the hospital staff is executing. Thus, activity estimation is central to the design of activity-aware applications that let ambient-intelligence environments discover the information relevant to the task at hand, thus empowering our iHopital vision.

We used HMMs for activity recognition because they consider past information when determining the probability of switching from one state (activity) to another. You can view an HMM as a doubly embedded stochastic process. It has an underlying hidden stochastic process that you can observe only through another set of stochastic processes, which produce the sequence of observations. So, one process can use observations and past information to inform the hidden process, which can then model an activity.

Data collection

To train and test our HMM, we used data from a workplace study conducted in a public hospital.[3] The study, which took place over nine months, required having the researchers shadow five nurses, five medical interns, and five physicians for two complete working shifts (approximately seven hours each). The total time of detailed observation was 196 hours and 46 minutes, with an average time of observation per informant of 13 hours and seven minutes.

The study recorded various actions, capturing for each a time stamp, the location, details about the nature of the action, a list of artifacts used, the contents of conversations, and the participants' role—patient, physician, and so forth (see figure 3a). An observer generated a new time stamp when new events altered one or more contextual variables.

We used the researchers' and hospital staff's analysis of this information to develop a coding scheme that describes the activities performed. For example, as figure 3a shows, the actions in the first row of the observation report describe a nurse recording a diagnosis in a patient's chart. This is an information management (IM) activity. The other activities identified are clinical case assessment (CCA), patient care (PC), preparation (P), coordination (C), and

classes and certification (CC). Although each of the activities is quite different, the contextual information that characterizes each one is usually similar. For example, if a physician is in front of a patient's bed with medical interns and a nurse, the physician could be discussing a clinical case, coordinating activities, preparing medicine, or providing patient care. So, it isn't obvious which information we should use to differentiate the activities.

Data representation

We transformed the collected data into inputs and outputs for the HMM. The inputs included the contextual information of the people involved and the artifacts being used during the activity's execution (see figure 3b), an activity transition matrix (figure 3c), and a priority vector (figure 3d). The activities we were estimating represent the hidden states. We trained three HMMs to estimate the activities, one per role (physician, medical intern, or nurse). (Putting them all into one HMM would have increased complexity and decreased performance.)

The priority probability vector (that is, the initial state distribution) represents the priority for executing one activity over another. We can view this vector as the percentage of the time hospital staff spent performing different activities. To calculate this vector, we measured the total time (per day per subject) that hospital staff spent on the activities to be estimated.

Similarly, we calculated the transition matrix from the study's information. The matrix portrays the state transition probabilities, which contain the probability of performing an activity or being at state q_i and switching to state q_j. To calculate this matrix, we measured the transitions from one time stamp to another.

On the basis of the contextual variables we used in a previous study as inputs for a neural network (NN) (see the sidebar for more information),[5] we defined a *contextual-variables matrix* for the HMM. We conducted an occurrence analysis (per activity per role) to disregard the variables with a low occurrence average. For example, a typewriter is a relevant artifact when predicting whether a physician is formalizing notes. Nurses don't have access to typewriters, so including this artifact when training the nurse's HMM won't improve the results. So, the input matrix included only variables with a prominent occurrence average.
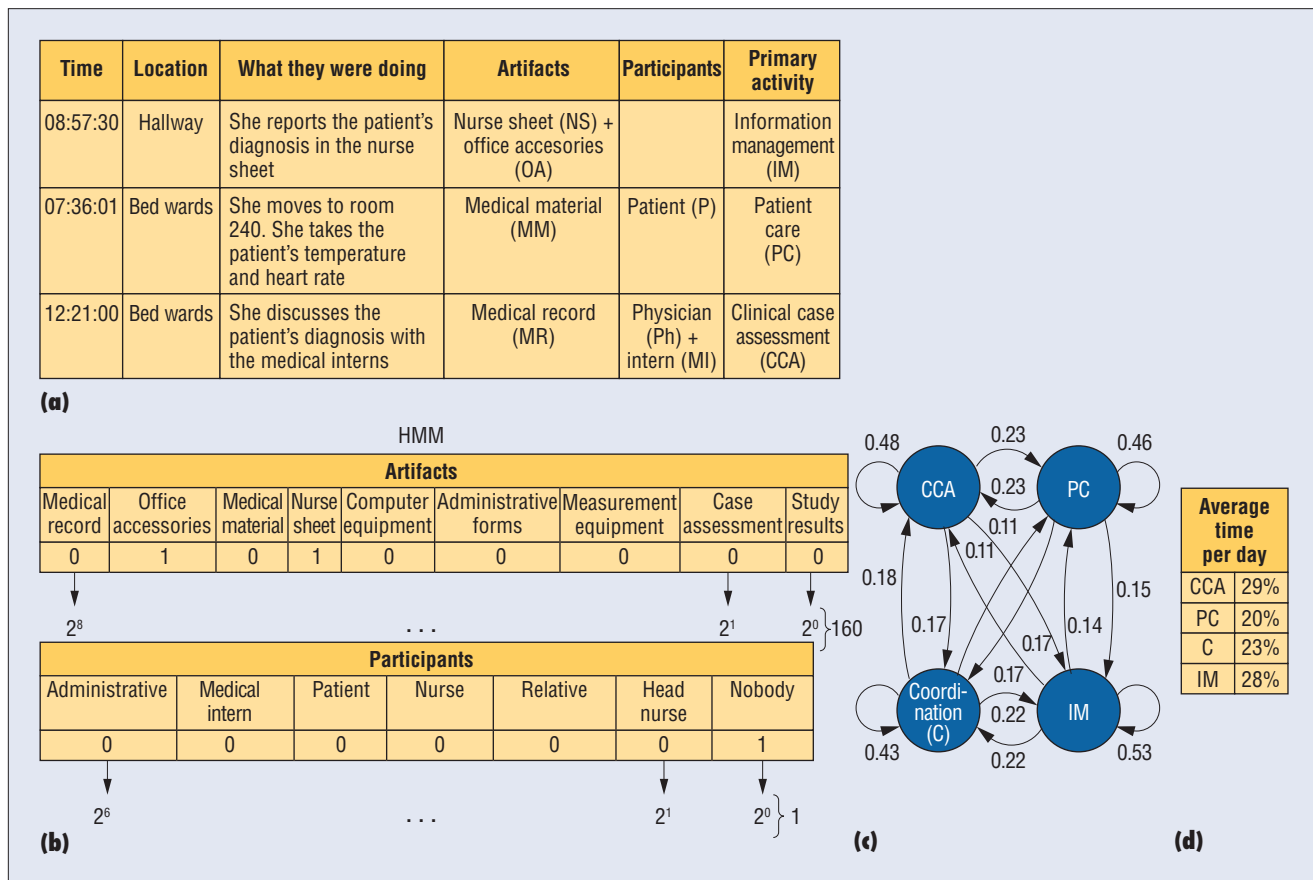
Time	Location	What they were doing	Artifacts	Participants	Primary activity
08:57:30	Hallway	She reports the patient's diagnosis in the nurse sheet	Nurse sheet (NS) + office accesories (OA)		Information management (IM)
07:36:01	Bed wards	She moves to room 240. She takes the patient's temperature and heart rate	Medical material (MM)	Patient (P)	Patient care (PC)
12:21:00	Bed wards	She discusses the patient's diagnosis with the medical interns	Medical record (MR)	Physician (Ph) + intern (MI)	Clinical case assessment (CCA)

(a)

(b)

(c)

(d)

Figure 3. A sample of the inputs to the hidden Markov model: (a) a coded fragment of the observations, (b) input variables, (c) a graph of the activity transitions for physicians, and (d) a physician's priority probability vector (that is, the priority for executing one activity over another).

Relevant contextual information

We analyzed which of the contextual information was relevant for training the HMMs. Using the location and time of day decreased activity-recognition accuracy[6] because hospital activities aren't executed at a predefined time of day or location. So, we eliminated these as input.

A two-level parallel-layered HMM

On the basis of a layered HMM model proposed elsewhere,[7,8] we developed a two-layer HMM to estimate hospital-staff activities (see figure 4). The first level comprises two HMMs, trained in parallel. We used matrices of people and artifact interactions as inputs. We then fused these two models' outputs to create the merged output matrix, which we used to feed the second-level HMM.

Once we had the HMM and the variables relevant to model the hospital staff's activities, we prepared the samples obtained from

the case study to conduct different simulations with the proposed HMM. We randomly gathered 1,258 samples from the 2,538 samples from the case study to form two data sets: one for training the model and one for testing it. To avoid allowing an activity's frequency to influence the training, we balanced the training and testing sets by taking only 230 samples (on average). After reducing the data sets, we split the data for training and for testing, as table 1 shows. We used

Table 1. Data for training and testing, according to role.

Activity	Nurses Train/Test	Medical interns Train/Test	Physicians Train/Test
Clinical case assessment	150/80	150/80	140/60
Patient care	150/80	90/60	140/60
Coordination	110/60	150/80	140/60
Preparation	150/80	—	—
Information management	67/35	150/80	140/60
Classes and certification	15/10	15/10	—

the HMM Matlab toolbox for the parallel-layered HMM's activity estimations.

Results and discussion

The layered model estimated staff activity with an error of 6.08 percent for nurses, 7.92 percent for physicians, and 8.17 percent for interns (the average error rate was 7.4 percent). Table 2 shows the confusion matrix of the medical interns' activities. The rows in this matrix correspond to the values of the
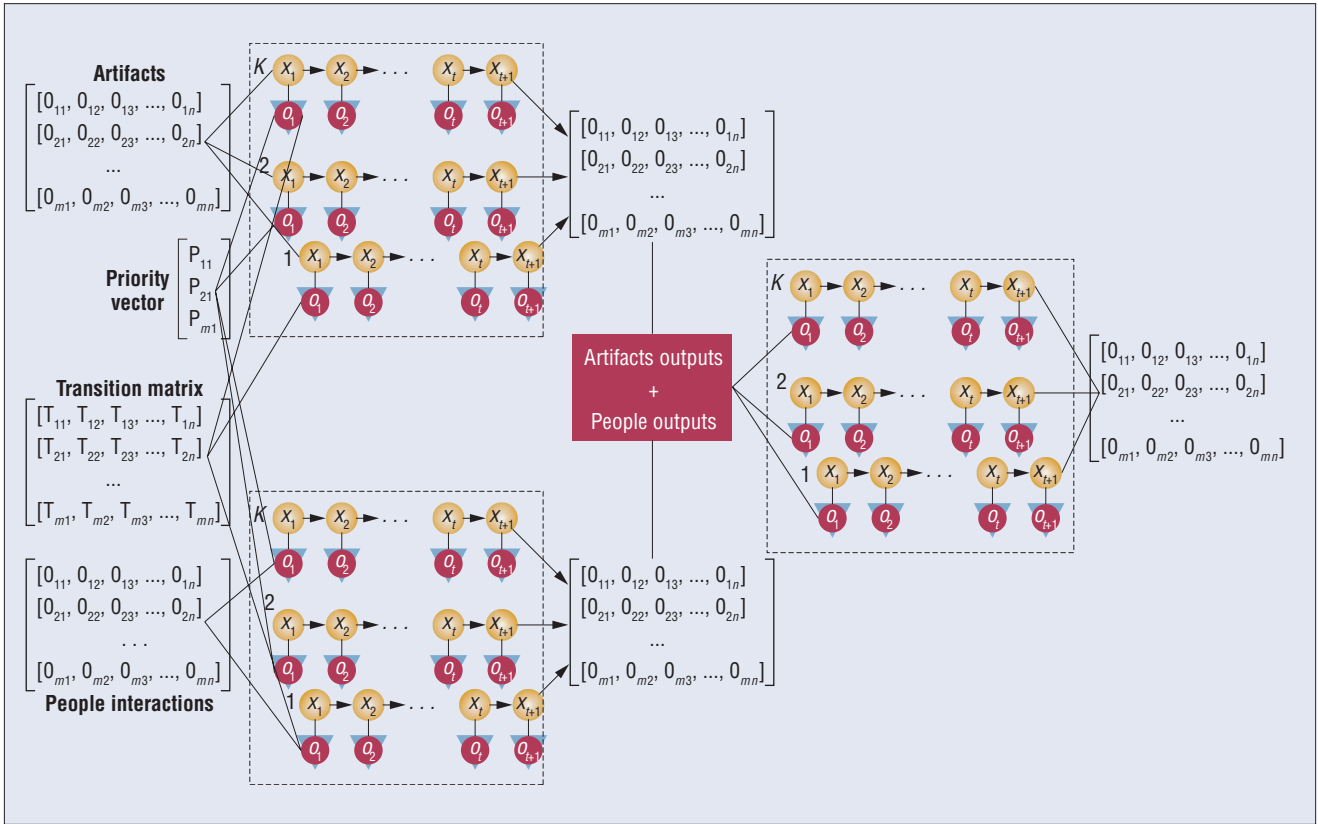
Figure 4. Our two-level parallel-layered HMM.

Table 2. Confusion matrix of interns' activities.*
The bold indicates a correct estimate.

Activity	HMM estimate				
	CCA	PC	C	IM	CC
Clinical case assessment (CCA)	**98.75**	1.25	0	0	0
Patient care (PC)	3.33	**96.67**	0	0	0
Coordination (C)	0	0	**100.00**	0	0
Information management (IM)	0	0	26.50	**73.75**	0
Classes and certification (CC)	0	0	0	10.00	**90.00**

*Average error rate: 8.17 percent

activity performed, and the columns correspond to the values the HMM estimated. The values in the diagonal are instances in which the HMM correctly estimated the activity.

By analyzing the information per role, we perceived differences in the accuracy for estimating certain activities. For example, the HMM correctly predicted when nurses were executing the "coordinating" activity 100 percent of the time. For physicians, the HMM correctly predicted the activity of clinical case assessment 98.75 percent of the time and of patient care 96.67 percent of the

time; for interns, the HMM best predicted clinical case assessment.

We observed that such role differences in activity recognition accuracy occur because the task switching experienced by each role affects how the HMM predicts the activity. For example, as table 2 shows, the HMM confused information management with coordination. Medical interns and physicians constantly switch between these two activities, so the contextual information doesn't change significantly, and the HMM therefore doesn't have enough evidence to differentiate between them.

Comparing HMMs, NNs, and expert observers

We compared the results with those obtained using NNs[5] and with activity estimations by human (expert) observers familiar with hospital work.

We conducted a *human test* experiment to determine how the expert observers used contextual information to estimate activities. Our expert observers comprised one medical intern and five researchers involved in the original study's data collection (and thus familiar with hospital work and our coding scheme[5]). We gave these six observers the contextual variables that the NN used to infer activities (the person's role, the location, the time, the artifacts used, and the colleagues present). We then evaluated the precision with which they inferred those activities.

To form our test set, we took for each activity a proportional number of samples of the ones used to test the NN, using 120 samples (40 per role). We presented the set of samples in the testing report similarly to the observation report in figure 3a. We asked each expert to situate herself in the description shown by the testing report and,

on the basis of the evidence, infer the activity being performed. The expert observers correctly estimated the activities 48.82 percent of the time.

Figure 5 compares the activity-recognition accuracy of HMMs, NNs, and expert observers; the HMM achieved the best results.

Testing our model's robustness

We assumed our model inputs were accurate. However, in practice, inputs might be estimated through other, possibly error-prone methods. To address this and validate our results, we tested our model's robustness by introducing noise into the input matrix (that is, the contextual-variables matrix). We randomly substituted one value from each vector in the input matrix, per activity. Because we used a binary codification to create this matrix, the transformation consisted in changing a value from one to zero.

For example, suppose we want to predict Juan's activity while interacting with Rita. If we can't read Rita's RFID tag, the input vector representing the people with whom Juan is collaborating will be incorrect. To represent this error, we substituted the third value of the people's input vector from $1010100 \xrightarrow{s} 1010000$. (We used the same criteria to transform the input matrix representing the artifacts used.) We conducted several experiments, each time changing the percentage of noise included per activity. As figure 6 shows, the model's activity-estimation remained stable—and above 80 percent—even with the input noise at 40 percent.

The iHospital: From location-aware to activity-aware computing

Deploying activity-aware applications requires having the designers establish an adequate threshold of acceptable estimation errors. We used the existing error in hospital workers' everyday practices as the upper bound. If an application doesn't surpass this upper bound, then the worst-case scenario is that the application's errors are the same as the ones already in hospital workers' current practices. Figure 7 is a first approach at placing bounds on the estimation error for the services the iHospital provides.

Medication errors are common, with adverse drug events reported among 13 percent of hospitalized patients.[9] These errors occurred primarily because hospital staff forgot to record the medication dose admin-
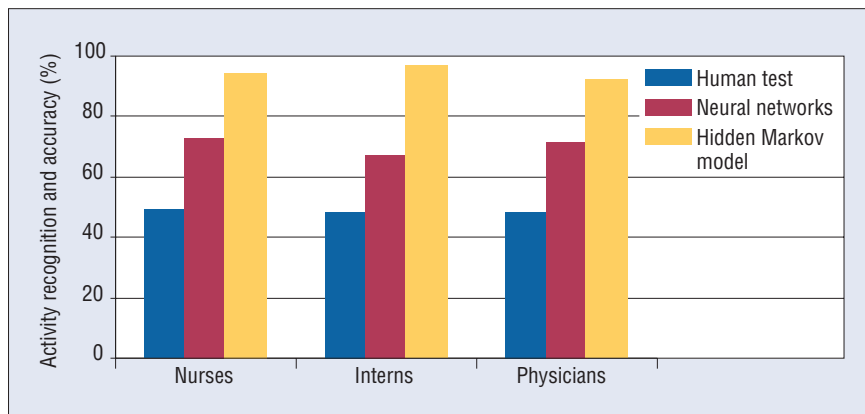

Figure 5. Comparing activity recognition by the HMMs, neural networks, and expert observers.
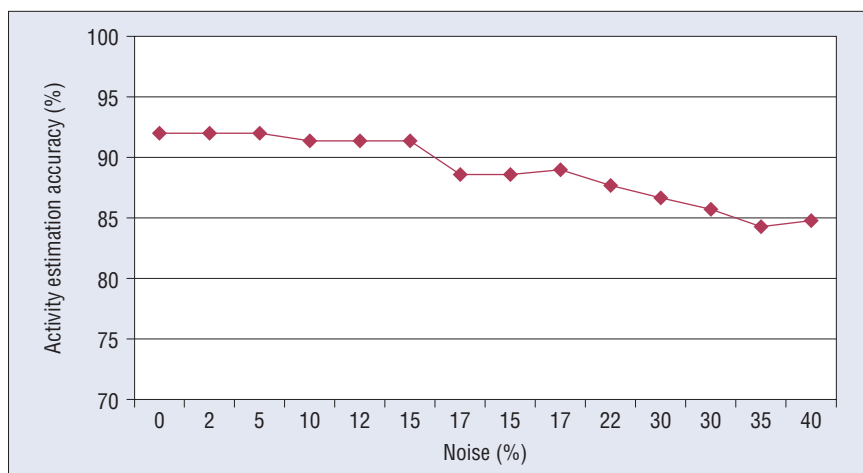

Figure 6. Activity-recognition accuracy as a function of the percentage of noise introduced to the input matrices.

istered to a patient. If a nurse doesn't record a dose, the nurse in the next shift might administer an additional dose. However, an activity-aware application could automatically register this information or remind the nurse to record such information. We argue that an activity-recognition error of up to 10 percent in this particular case would be sufficient for developing an activity-aware reminder or for adapting and personalizing the information presented to the hospital staff.

In addition, hospital work is highly fragmented owing to frequent interruptions and constant changes in the environment. According to one study, interruptions account for 37 percent of the time hospital staff spend collaborating and interacting with colleagues.[10] Despite this, context-aware applications in the iHospital are largely unaware of the social situations surrounding their usage and the consequences their actions might have. If an application were

aware of the user's availability, it could negotiate interruptions at appropriate times. Availability can be derived from knowing the activity a person is performing. So, a context-aware application that's aware of a person's level of interruptibility could decide whom to interrupt and when, or to automatically update the person's availability. Because current strategies for managing the hospital staff's level of interruptibility are either nonexistent or inaccurate, we feel an error of up to 30 percent would be acceptable and still make the application useful.

The staff also often uses contextual information such as the time of day and presence and location of colleagues to infer their activities. For example, if a physician is in a patient room with medical interns at 11 a.m., we can infer that he or she is performing a ward round. For this case, even a considerable error wouldn't affect hospital-staff practices, because the current strategies for
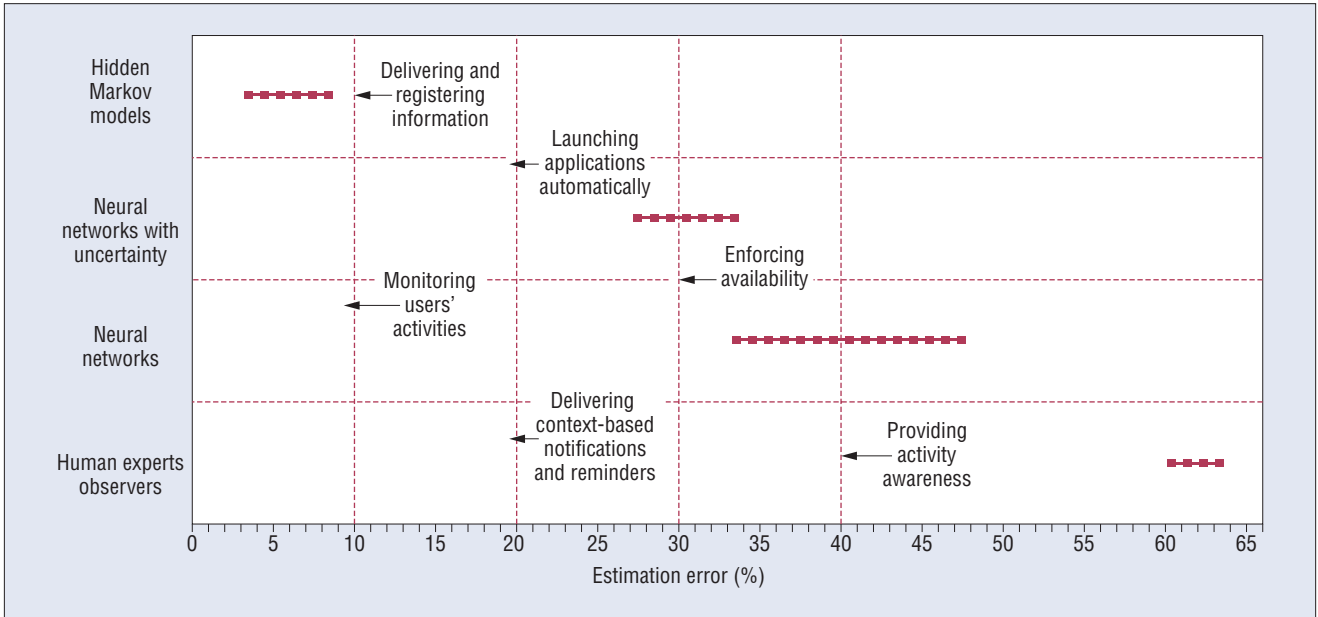
Figure 7. A range error comparison for activity-aware services that can be deployed or developed for a given activity error-estimation threshold. The arrows point to each activity's threshold, and the red boxes indicate activity-aware services that could be developed for a given activity error estimation.

Table 3. Nurses' availability.*

Actual availability	Estimated availability	
	Not available	Available
Not available	96.82	3.18
Available	1.6	98.4

*Average error rate: 2.39 percent

inferring colleagues' activities aren't accurate and are used only to get a sense of the rhythms of hospital work.

A current iHospital service that activity recognition could augment is the Vocera communication system. The application could infer user availability if it were aware of the activities its users might execute. For example, hospital workers performing clinical case assessment, patient care, classes and certification, and preparation generally don't want to be interrupted. In contrast, they're much less likely to mind being interrupted when they're engaged in information management or coordination activities.

Table 3 presents the activity estimation results obtained with our approach, grouped by those activities that can be associated with the user's availability. We grouped the activities based on the availability perceived by the hospital workers. For example, when a physician is involved in clinical case assessment, patient care, classes and certification, or even preparation, he or she gen-

erally doesn't want to be interrupted. We recurrently observed that medical workers, especially interns, wait until the discussion finishes before approaching a physician. In contrast, when physicians are engaged in other activities, such as information management or coordination, they're much more likely to be interrupted. Based on this, we classified activities directly related to the patient care as those requiring the hospital staff's full attention and other activities as ones with a higher level of interruptibility.

The table shows that the estimation accuracy for nurses is 97.61 percent. A system such as Vocera could use these results to decide whom to interrupt and when. For instance, a physician can request the help of any nearby nurse, and the system will decide, on the basis of the availability estimation, which nurse to call at that moment.

We could also potentially integrate activity recognition into a multiagent platform[2] in which autonomous agents represent hospital staff and resources to personalize in-

formation on the basis of the role or activity. For example, when a physician using a PDA performs clinical case assessment, a system could recognize the activity and offer to display the patient's clinical record, available lab or radiography results, and information related to the patient disease. The activity-aware system could also prioritize incoming information to prevent unnecessary interruptions.

In the future, we plan to conduct a field study to train and test the classifier for a new setting—in particular, nursing homes. We plan to measure and recognize activities of daily living to assist people with age-related cognitive decline.

References

1. J. Camacho et al., "MobileSJ: Managing Multiple Activities in Mobile Collaborative Working Environments," *Int'l J. E-collaboration*, vol. 4, no. 1, 2008, pp. 60–73.
2. M.D. Rodriguez et al., "Agent-Based Ambient Intelligence for Healthcare," *AI Comm.*, vol. 18, no. 3, 2005, pp. 201–216.

3. E.B. Moran et al., "Mobility in Hospital Work: Towards a Pervasive Computing Hospital Environment," *Int'l J. Electronic Healthcare*, vol. 3, no. 1, 2006, pp. 72–89.

4. V. Stanford, "Beam Me Up, Doctor Mc-Coy," *IEEE Pervasive Computing*, vol. 2, no. 3, 2003, pp. 13–18.

5. J. Favela et al., "Activity Recognition for Context-Aware Hospital Applications: Issues and Opportunities for the Deployment of Pervasive Networks," *Mobile Networks and Applications*, vol. 12, nos. 2–3, 2007, pp. 155–171.

6. D. Sánchez, M. Tentori, and J. Favela, "Hidden Markov Models for Activity Recognition in Ambient Intelligence Environments," *Proc. 8th Mexican Int'l Conf. Current Trends in Computer Science* (ENC 07), IEEE CS Press, 2007, pp. 33–40.

7. N. Oliver, A. Garg, and E. Horvitz, "Layered Representations for Learning and Inferring Office Activity from Multiple Sensory Channels," *Computer Vision and Image Understanding*, vol. 96, no. 2, 2004, pp. 163–180.

8. D. Zhang et al., "Modeling Individual and Group Actions in Meeting with Layered HMMs," *IEEE Trans. Multimedia*, vol. 8, no. 3, 2006, pp. 509–520.

9. A.L. Friedman, "Medication Errors Common, Difficult to Detect among Transplant Patients," *Archives of Surgery*, 22 Mar. 2007.

10. G. Alvarez and E. Coiera, "Interruptive Communication Patterns in the Intensive Care Unit Ward Round," *Int'l J. Medical Informatics*, vol. 74, no. 10, 2005, pp. 779–781.

T h e   A u t h o r s

Dairazalia Sánchez is a graduate student in computer science at CICESE. Her research interests include machine learning, mobile and ubiquitous computing, human-computer interaction, and human-activity modeling. She received her MSc in computer science from CICESE. Contact her the Computer Science Dept., Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Km. 107 Carretera Tijuana-Ensenada, Ensenada, B.C., 22860, Mexico; dcortes@cicese.mx.

Monica Tentori is a doctoral student in computer science at CICESE and a lecturer in computer science at the University of Baja California. Her research interests include ubiquitous computing, human-computer interaction, and medical informatics. She received her MSc in computer science from CICESE. She's a student member of ACM-SIGCHI. Contact her at the Computer Science Dept., Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Km. 107 Carretera Tijuana-Ensenada, Ensenada, B.C., 22860, Mexico; mtentori@cicese.mx.
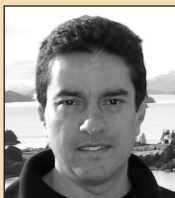
Jesús Favela is a professor of computer science at CICESE, where he leads the Mobile and Ubiquitous Healthcare Laboratory and heads the Department of Computer Science. His research interests include ubiquitous computing, medical informatics, and computer-supported cooperative work. He received his PhD from the Massachusetts Institute of Technology. He's a member of the ACM and the American Medical Informatics Association. Contact him at the Computer Science Dept., Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Km. 107 Carretera Tijuana-Ensenada, Ensenada, B.C., 22860, Mexico; favela@cicese.mx.