Journal of Biomedical Informatics 44 (2011) 432–440

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Toward automated workflow analysis and visualization in clinical environments

Mithra Vankipuram [a,*], Kanav Kahol [a,b], Trevor Cohen [a,c], Vimla L. Patel [a,c]

[a] Center for Decision Making and Cognition, Department of Biomedical Informatics, Arizona State University, Phoenix, AZ, USA
[b] Simulation, Education and Training (SimET), Banner Good Samaritan Medical Center, Phoenix, AZ, USA
[c] School of Health Information Sciences at Houston, The University of Texas, Houston, TX, USA

ARTICLE INFO

Article history:
Received 2 May 2009
Available online 31 May 2010

Keywords:
Clinical workflow
Complex systems
Medical errors
Radio identification
Virtual reality
Visualization

ABSTRACT

Lapses in patient safety have been linked to unexpected perturbations in clinical workflow. The *effectiveness* of workflow analysis becomes critical to understanding the impact of these perturbations on patient outcome. The typical methods used for workflow analysis, such as ethnographic observations and interviewing, are limited in their ability to capture activities from different perspectives simultaneously. This limitation, coupled with the complexity and dynamic nature of clinical environments makes understanding the nuances of clinical workflow difficult. The methods proposed in this research aim to provide a quantitative means of capturing and analyzing workflow. The approach taken utilizes recordings of motion and location of clinical teams that are gathered using radio identification tags and observations. This data is used to model activities in critical care environments. The detected activities can then be replayed in 3D virtual reality environments for further analysis and training. Using this approach, the proposed system augments existing methods of workflow analysis, allowing for capture of workflow in complex and dynamic environments. The system was tested with a set of 15 simulated clinical activities that when combined represent workflow in trauma units. A mean recognition rate of 87.5% was obtained in automatically recognizing the activities.

1. Introduction

1.1. Workflow and medical error

The health care industry faces a number of challenges and arguably one of the most important ones lies in maintaining high levels of patient safety. A much-cited report released by the Institute of Medicine (IOM) estimates that as many as 98,000 people die each year due to medical errors [1]. Medical errors cause more deaths annually when compared to motor vehicle accidents, breast cancer or even HIV. Consequently, scientists across the world have actively researched the nature of medical errors and the activities that cause the errors. An integral component of this research effort is workflow analysis.

A workflow is a description of a sequence of operations or activities performed by various entities or agents in the system [2]. It provides a description of the context and conditions in which errors occur. Careful analysis of workflow can be employed to model the distribution of cognitive work and the information flow in complex environments. For example, Malhotra et al. [3] utilized ethnographic observations and interview data to analyze the workflow in an intensive care unit. The workflow served in the development of a cognitive model from which details of information flow could be extracted. This type of analysis could lead to the discovery of latent systemic flaws that potentially result in adverse events. In addition, monitoring and assessment of workflow in complex clinical environments can provide clues on the efficacy of patient management. For these reasons, workflows in clinical environments are an important aspect of patient care and safety research.

Recent research has approached the study of social systems such as clinical environments, using scientific theory based on *complex adaptive systems* [4]. A complex adaptive system is defined as a dynamic network of entities acting simultaneously while continuously reacting to each other's actions [5,6]. The control in such systems is often decentralized and highly dispersed. The overall behavior is the result of a large number of decisions made at every instant by individual agents or entities in the environment. In such environments, the inherent complexity of interactions makes capturing and analyzing workflow difficult and often unpredictable. These difficulties can be attributed to a disassociation between the complex nature of the environment and the tools available to analyze the workflow in such environments.

Critical care environments are complex and dynamic and change constantly with stress levels. The changes in staff continually alter team dynamics and the excess of technology and equipment creates unforeseen demands on clinicians. These

* Corresponding author. Address: Department of Biomedical Informatics, N 5th St., Phoenix, AZ 85004, USA. Fax: +1 602 827 2564.
E-mail address: mvankipu@asu.edu (M. Vankipuram).

1532-0464/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2010.05.015

characteristics allow for the categorization of critical care environments as complex adaptive systems. Consequently, the problems with workflow analysis in complex adaptive systems are applicable to critical care environments as well.

The tools currently used for workflow analysis in clinical environments include methods such as ethnographic observation, shadowing of individual clinicians, surveys and questionnaires [7]. These are *qualitative tools* of analysis. The data collected by these methods can be used to model segments of the workflow centered on a particular individual and their activities [3]. While the workflow generated in this manner captures many aspects of the overall workflow, more often than not, certain key pieces of information may never be captured. For example, observations are gathered from an individual's point of view and may not be adequate to capture multiple activities occurring at an instance of time. Theoretically, by increasing the number of observers in the clinical environment it is possible to capture information about the activities in the environment from several perspectives. However, based on informal interviews conducted with trauma clinicians, more than two observers are often considered to be disruptive to the clinical workflow. With such constraints imposed on data collection in complex environments, there is a need for an unobtrusive alternative that can augment existing methods of data collection and enable piecing together a more complete workflow, from individual and team perspectives.

1.2. Black-box approach to capturing clinical workflow

An example of complex social system that is similar to a clinical environment is aviation. A critical component of error analysis in aviation is the *black box*. The black box, as a tangible unit refers to devices installed on aircrafts that track both communication within the cockpit of the aircraft as well as performance parameters such as altitude, airspeed and heading. From a conceptual perspective, the black box is a constant monitoring tool that does not interfere with the procedures of aviation and simply monitors parameters pertaining to the flight.

In clinical environments, tools that are conceptually similar to a black box would be able to monitor workflow continuously. In addition, the tools would monitor the workflow without disrupting the activities of entities in the environment. In this paper, we propose a system that utilizes the black-box approach for continuous workflow monitoring and analysis in clinical environments. This system employs radio identification technology (RID) to offer continuous monitoring. RID-enabled tags are portable electronic devices used to uniquely identify entities in an environment. These tags can be used to continuously monitor agents and artifacts in the clinical environment. Basic information about interaction between the entities, such as duration of proximity and location can be recorded. This data, combined with qualitative measures provide an intermediate workflow that can be visualized in three-dimensional (3D), virtual reality (VR) environments. We can combine this system with conventional means of ethnographic observations for high-resolution capture of workflow. The end result is a system that enables the capture and visualization of workflow in complex environments for the purpose of enhanced workflow analysis.

2. Background and related work

Methods used to analyze workflow in clinical environments can be one for two types – *qualitative* methods or *quantitative* methods. While qualitative methods involve subjective observations gathered by researchers, quantitative methods typically involve the usage of sensor technology or video recordings to capture data about workflow. The main differences between the data captured using quantitative methods and qualitative methods are as follows:

(i) Using quantitative methods, accurate time-stamped data can be obtained. Human-intensive methods can only produce near-accurate time-stamped observations.
(ii) Qualitative methods of data collection produce relatively low volume, high quality data. On the other hand, quantitative methods produce a high volume of abstract data that in some way indicates underlying workflow.
(iii) Human-intensive qualitative methods best suited for low-intensity situations while automated quantitative methods are optimal for data gathering in high-intensity situations.

The following sections detail the existing qualitative and quantitative methods used to analyze workflow in critical care environments.

2.1. Qualitative methods for workflow analysis

Malhotra and his colleagues analyzed the workflow in the intensive care unit (ICU) in order to understand the process of evolution of error in a critical care setting [3]. Ethnographic observations and interviews are utilized to gather data to model workflow centered on the entities and activities in the environment. The process of gathering observations involves following a key member of the critical care team and recording all of their interactions with other members of the team, patients and equipment. These key players are then interviewed to corroborate the observations collected and to delineate their individual workflows. Using observations and interviews, a collective workflow is reconstructed by combining the individual workflows of each key player. The developed workflow summarizes how ICUs function and where errors are most likely to occur.

Laxmisan and her colleagues utilize ethnographic observations and interviews to analyze the workflow in an emergency department (ED) [8] as well. The workflow is analyzed to study the cognitive demands imposed by the workflow in the context of the work environment. Multi-tasking, interruptions, gaps in information flow and handovers during shift change are some of the aspects of the workflow that are studied in detail.

2.2. Quantitative methods for workflow analysis

Quantitative methods provide some means of gathering information about the activities and whereabouts of entities in an environment. An entity could be a person (nurse, physician, patient, etc.) or a machine (such as ultrasound machine). The tracked activities can then be pieced together (similar to integration of observations and interviews) to provide a complete picture of the workflow.

The sensors typically used for entity activity recognition include passive infrared sensors, radio identification tags and pressure sensors. The sensors, depending on their type are utilized to detect various activities that the entity is involved in. A number of systems have been developed for activity recognition and workflow monitoring using different types of sensors. These systems use the various types of sensors in some combination to model key activities of the entity being tracked. In general, these sensors are encased into a physical form representing a tag. These tags can sense different types of information like movement and location through the ensemble of sensors embedded in the physical form.

In the domain of healthcare, tags have been employed for tracking patients, equipment and staff to improve patient care and the efficiency of clinical workflow [9–15]. Fry and Lenert [15] developed a system for location tracking of patients, staff and equipment called MASCAL. The main aim of the system was to aid in streamlining patient care during mass casualty situations. RID tags are used by the system to track the location of key players (clinicians and equipment) in patient care during emergencies. This information is integrated with personnel databases, medical information systems and other applications (such as those that enable registration and triage) in order to centralize the management of resources during critical situations. In addition, MASCAL includes interfaces for centralized management of various entities in the system.

Chen et al. [14] studied the incorporation of RID technology in a clinical setting in non-psychiatric hospitals in Taipei, Taiwan. Tags were used to identify patients, and notify clinicians on the status of patients and patient related information (laboratory reports, radiology results, etc.). Preliminary studies showed that using the RID enabled framework, decreased the length of waiting for patients in intensive care units.

The other technique for activity monitoring is processing of video recordings. Hauptmann et al. [16] describe a system that recognizes activities from videos captured using video processing techniques. The system was developed to recognize activities of daily living (ADL) for patients. Examples of ADL activities include, visiting the washroom, eating, sleeping, etc. Cameras placed at key locations within the environment provide video feeds. These video feeds are processed to identify the patients and hence draw conclusions on the possible activities the patients were involved in.

2.3. Limitations of qualitative methods

Qualitative methods are human-intensive, i.e. they require significant amounts of human effort for data gathering and analysis. The dependency of qualitative methods on human effort has certain advantages and disadvantages. The main advantage is that human-intensive methods usually yield data that are of high quality. These data are detailed and descriptive, and potentially insightful inferences can be made using qualitative analyses of these descriptions. The disadvantage however, is that the dependence on people for data gathering and analysis limits the capabilities of these methods to capture important details of the collective workflow in a critical care environment.

Observation gathering is a classical qualitative method for workflow analysis that suffers from its dependence on human effort. It is difficult for individuals to monitor and document all activities that occur at every instant in a dynamic and complex environment. Interviews on the other hand suffer from the poor recall of events on the part of clinicians being interviewed. Facts about events may be altered as the memory of the event changes temporally (post hoc bias). Other real-time methods of data collection such as audio and video recording systems not only require consent from clinicians to be used to gather workflow data, but also require significant human efforts for processing data collected to retrieve meaningful information. Post-processing of real-time data involves manual analysis of audio and video data in order to detect various workflow events. The real-time data is then manually annotated with the key events that have been detected. This process requires time, effort and sufficient researcher expertise in order to be completed successfully. Such limitations make these methods more suited for workflow analysis in simple, low-activity environments.

2.4. Limitations of quantitative methods

In most quantitative methods, sensors for monitoring activities and locations are placed at pre-defined locations. The rigid infrastructure often makes installation costs prohibitive. In addition, maintenance can be complicated if spatial configurations are altered. Another issue lies with the modeling approaches employed to track workflow. In all the current systems, the sensor system is employed to determine the location of the entities from which activities are estimated. This system works well if the location identification is reasonably accurate. However, RID systems can often be highly erroneous, resulting in close to 200% errors in location estimates [17]. Location in these systems is determined by geometric triangulation methods that have limited performance in environments with electromagnetic fields. As clinical environments require large amounts of equipment, it is impossible to control for electromagnetic fields. To account for this high rate of error, activities that are covered by the current approaches are macro-movement based activities such as entering a room or going from one area of the hospital to another. Current systems are limited in documenting activities that occur in smaller area as the sensors cannot discriminate location in these environments with acceptable accuracy.

Video-based tracking suffers from similar issues. The locations of cameras are fixed. Areas need to be analyzed to ensure that camera cover all parts that need to be monitored. In addition, real-time analysis of videos for entity recognition can suffer from typical video processing problems, such as occlusion of entities by other entities, noise, motion blur, uneven lighting and so on. This, coupled with the requirements of privacy and security, often render video-based capture unusable.

As both qualitative and quantitative methods have advantages and disadvantages, an improved solution for workflow monitoring can be obtained by combining the two types of methods. In this work, workflow monitoring is performed using RID tags in conjunction with ethnographic observations. Unlike infra red tags, RID tags do not require a line of sight with other tags to record information. Hence, we utilize RID tags to gather the quantitative data. Observations gathered complement RID data by providing a detailed description of communication and interaction activities that cannot be captured using the tags. This is the approach taken for the development of the system described in this paper.

3. Methods

In general, workflow can be described by (a) the underlying cognitive processes that drive decision making, (b) physical movement, and (c) interaction and communication activities. In this work, we present the framework for a workflow analysis system that combines qualitative and quantitative methods of data collection to capture each of these three activities. Fig. 1 depicts the

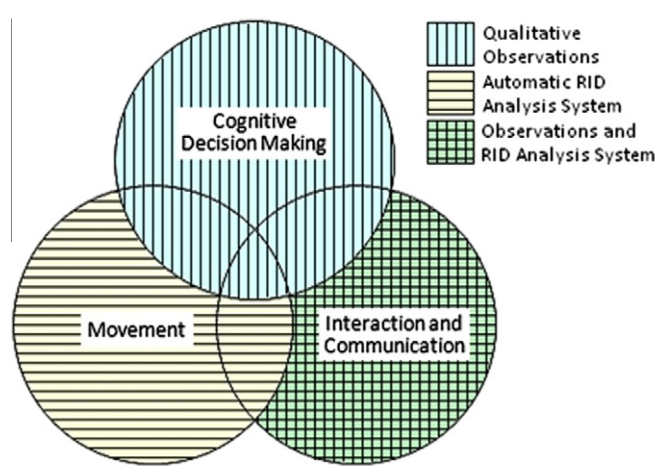

Fig. 1. Overview of activities captured and tools utilized.

types of activities that can be captured using the qualitative and quantitative data gathered. RID tags can provide quantitative information about movement activities, in addition to some basic interaction statistics such as proximity between two or more clinicians and time spent at particular locations. These statistics could be utilized to model the movement patterns of clinicians in the environment and the overall behavior of the clinical team. RID tags, however cannot gather information about specific details of communication between clinicians, or details about the thought processes of clinicians that result in a particular situation. We rely on qualitative observations to provide this information.

Traditionally qualitative data collection requires observers to focus on both cognitive activities and movement activities. This is in addition to collecting detailed information about the time of activity initiation and the sequence of activities in a workflow. Our system offers the means to offload the task of recognizing movement based activities and a subset of interaction and communication activities to the automatic algorithms that process incoming quantitative data from RID tags and estimate the activity being performed by the team (as depicted in Fig. 1). Our system can theoretically capture any movement activities that require team members to move at least 8 inches [17]. The communication and interaction activities that show movement patterns of a single entity or a group of entities can be suitably recognized by the developed system.

An example of such an interaction activity would be “patient arriving”. Typically, interactions with two or more entities can be monitored using RID information. When a patient arrives in a trauma room, the team members tend to converge at a trauma pod. This is an interaction and communication activity that can be captured by using movement as a proxy. As a general rule, any interaction and communication activity that is accompanied with measurable movement can be captured by this system and recognized. Following the same logic, any communication or interaction activity that is not accompanied by movement cannot be captured by the automatic analysis system.

In the future we aim to include additional sensors that can give us more detailed information on some of the activities. Incorporation of audio recording would facilitate automated tracking of communication between entities. Further, in order to increase the resolution of the movement activities that can be tracked through our system we aim to use accelerometers. Acceleration measurements could provide information on whether the tagged entity is moving. This would enable the refinement of movement tracking and activity recognition. The presented framework accounts for these future additions to allow for seamless integration.

3.1. Conceptual framework

In the given framework, we collect two different streams of data:

(a) qualitative data from observers, and
(b) quantitative data gathered from the RID tags.

Both the qualitative data and quantitative data are obtained from standardized sources. While time-stamped quantitative data is retrieved from the RID tags, observations are generated by observers using an activity tracking software tool depicted in Fig. 2.

The tool contains a list of commonly occurring activities for the nurse and physician. The activities chosen were based on an ontology developed by Zhang based on his prior work on analyzing the workflow in emergency departments [2]. Observers may select an appropriate activity from the list provided and add detailed comments a description text box. The observations are then automatically dated and timed and stored in the output observation file. In this way time-stamped data is obtained for both qualitative and quantitative data sources. This makes synchronization of the two data streams possible.

Quantitative data is obtained using *active* RID tags to gather data. Active RID tags have an inbuilt power source, hence the name active. In addition to being portable, active tags use low levels of energy ensuring that they do not interfere with other devices, such

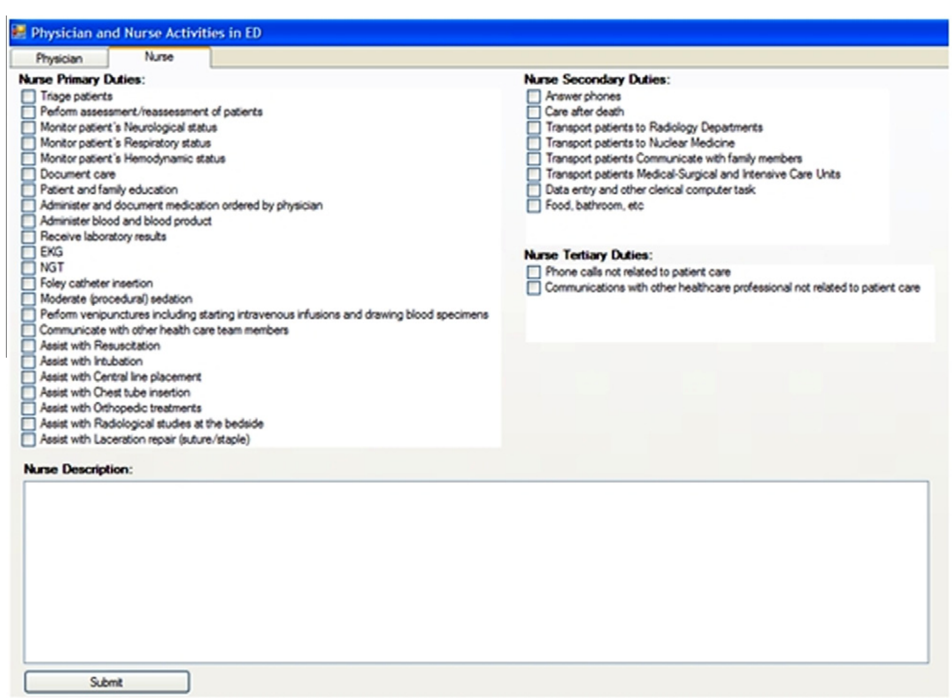

Fig. 2. Observation tracking tool.

Fig. 3. Active RID (SNiF®) tag and base station.

as telephones and other network connections found in a health care setting. This is a vital requirement for the system. We chose off the shelf tags available from SNiF®. These are shown in Fig. 3. The tags record encounters with other tags (tag–tag encounter) and base stations (tag–base encounter). For each encounter or interaction, the tags record:

(a) identification number of the tag or base station detected,
(b) time and date of encounter, and
(c) the received signal strength indication (RSSI) value.

The RSSI value provides proximity information. This value is inversely related to the distance between the interacting tags. Consequently, it can be used to measure approximate distances between tags and base stations involved in an interaction. The assessment of the temporally changing distance between entities (people and equipment alike), enables inferring the initiation of activities such as a resident leaving trauma, or a nurse documenting the case at the nurse's station. Such utilization of proximity information makes this a crucial metric for the functioning of the activity recognition system.

Base stations can be considered as fixed tags that provide tag location information. We place base stations at critical locations in the critical care environment. It was found that in a trauma unit, the trauma bays, nurse's station and entry and exit points were some of the key locations. These locations will vary from site to site, depending on variations in the workflow. Placing base stations in these locations enables the determination of activities that occur in specific regions. Combining tag–tag and tag–base encounter information we can obtain an abstraction of interaction and communication activities between various entities in the clinical environment.

In order to develop our quantitative analysis system there is a need for training and validating the models. The procedures for data collection to train our quantitative system are as follows. Firstly, we identify certain movement activities, communication activities and interaction activities that are of interest. Subsequently, we place the RID tags on entities and collect data on the movements of entities when performing those activities. This data is then employed to train a model for automatic recognition of these activities in the future observations.

3.1.1. An illustrative example

Let us see an illustrative example of how some activities in the clinical environments can be captured by appropriate placement of tags and base stations (which are similar to tags in functionality but have a larger form factor. See Fig. 3). Consider the scenario representing patient arrival is depicted in Fig. 4. Firstly, key members of the patient care team (resident, nurse and so on) gather by the bed of the patient. Following this, examination of the patient takes place. A resident may move to the telephone to consult or the nurse may move to the nurse's station to document details of the encounter. All these activities are linked to entities performing some type of movement in the environment.

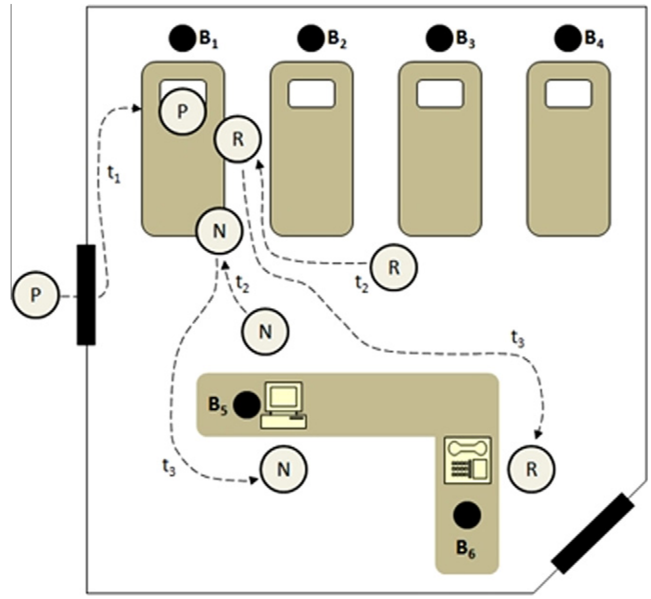

Fig. 4. Scenario: patient arrival at a trauma unit.

Formally we can express this sequence of activities in terms of time as:

(i) At time t_1: Patient arrives at the trauma unit and is sent to the trauma bay.
(ii) At time t_2: The nurse and a resident check in on the patient.
(iii) At time t_3: The resident seeks a phone consult while the nurse heads over to the station to continue with documentation.

In our diagram (Fig. 4), 'P' refers to the patient; 'N' refers to the nurse and 'R' to the resident on call. The black solid dots denote location of base stations (B_{1-6}). Base stations were placed at various key locations; one at each trauma bay, one near the phone and the other near the computer. For these given sequence of events, the following are the trends we see in the data derived from the tags.

(i) At time t_2: Tags R and N get close to B_1.
(ii) At time t_3: Tag N is very close to B_5 and Tag R is very close to B_6.

With the initial setup phase we know that B_1 is trauma bay 1, we can assume that the patient is being managed by the nurse and resident at time t_2 and that the patient arrived at the unit sometime before t_2. Therefore, at time t_3, the system can probabilistically estimate that the nurse was documenting the patient report, and the resident was seeking a phone consult. While the scenario presented is a simplification of the total process, it provides a conceptual view of how we can track activities through tags. In reality, activity models generated can be more complex. The models would be required to handle variations in activities performed while classifying them accurately. A pattern recognition approach is required to handle such variations. The approach taken in this work to classify activities is discussed in the following section.

3.2. Activity modeling and recognition using Hidden Markov Modeling

Hidden Markov Modeling (HMM) is a probabilistic modeling tool that is usually employed for temporal sequence analysis and have been effectively used in movement analysis, gesture and speech recognition applications [18,19]. An HMM models a temporal sequence of events (called an observation sequence) in terms of a state machine, in which the current state of the model is probabilistically dependent on the previous states. A well-trained HMM activity model can detect the temporal activities that the HMM has been trained for.

As with any method, HMM based activity recognition has certain advantages and disadvantages. The key disadvantage of HMMs lies in the fact that the amount of data that is required to train an HMM is very large. Another issue with HMMs is that they require positive data to train with, i.e. in order to effectively train an HMM to recognize a class of activities, we require a carefully constructed training set that best describes the activity. However, these disadvantages are outweighed by a trained HMM's capability to handle variations in the final style of execution of an activity. Activities may be performed in a different manner in critical care environments and it is important that the model of activities accounts for these variations. By training the HMM system in a robust manner, it is possible to recognize the motion and some communication activities regardless of the deviations for our application. In addition, HMMs scale well as they can be trained to learn activities incrementally. New activities can be trained for without affecting models of previously learned activities. For these reasons, we chose HMMs for the development of activity models and activity recognition.

Activity recognition using HMMs is a 2 step process. It involves (i) *training* HMMs for specific activity models and (ii) *testing* the HMMs for their recognition accuracy with annotated test samples. In order to develop robust activity HMMs, we first require data that describes the activity. This data is obtained from the RID tags. More specifically, the data utilized is the RSSI value of each tag–base encounter gathered during data collection. We collect this data for the activities of interest in multiple samples. We utilize half of the samples for training the HMMs and retain the rest for testing the developed models. A database of samples for each activity facilitates training the HMMs for each activity, thereby creating a library of HMM activity models for each activity. The training of HMM activity models is achieved using the Baum–Welch algorithm [20].

Once a library of HMMs is built with one HMM for each activity, the developed models can be tested. The testing of an activity sample proceeds by firstly, estimating the probability that the sample movement belongs to the library. This is achieved using the Forward–Backward [20] procedure for each of the HMM's in the library. The HMM that yields the highest probability for the test sequence is determined to be the type of activity that the movement sequence belongs to. The accuracy of recognition is measured as the ratio of the number of correctly identified test sequences to the total number of test sequences. In this manner, activity models are developed and tested for activity recognition. The recognized activities can then be visualized in 3D virtual environments.

3.3. Visualization of workflow in a virtual environment

The automated system for workflow analysis generates a series of activities that takes place in the clinical setting. Visualizing workflow in 3D enables researchers and clinicians alike to easily grasp the activities that make up the workflow. In addition to enable researchers review workflow in a novel way, the configurable VR visualizations can also be employed for educational purposes. For example, a resident would be able to go experience a trauma from the perspective of the attending or nurse. This kind of configurability would enable the cross-training of clinical teams. The visualizations can also be used to educate clinicians by illustrating cases of optimal workflow in relation to error-prone workflow.

In the domain of healthcare, virtual reality has been used to develop simulations for training of cognitive and psychomotor surgical skills and clinical decision making skills [21–23]. However, there is a lack of VR-based solutions for visualization of workflows and error scenarios even though such systems may have a major role to play in error prevention and mitigation. We can employ online VR environments such as Second Life® (http://secondlife.com/) and Active Worlds® (http://www.activeworlds.com/) for such visualizations. In this stage of the work, we have developed a standalone system that could be employed for such visualizations employing an open source gaming engine called Irrlicht (www.irrlicht.net).

A sample virtual trauma unit (see Fig. 5) was developed to mimic the trauma unit at Banner Good Samaritan Medical Center which is the site of development for the project. The virtual trauma room consists of four trauma pods or beds. The nurses' station faces the trauma pods. A computer and phone are key components that are included in the design of the nurses' station. Two exit doors are present in either side of the trauma room. These details are synchronous with the test and real-world setup. The current simulation contains three basic characters – the patient, resident and the nurse. The number and type of models to be utilized depend on the entities studied in the real-world. Models of the characters are built using modeling software (Maya and 3dMax; http://usa.autodesk.com/). Once the models are developed they can be controlled in the simulated world programmatically.

In order to obtain VR simulations of the workflow, the system generates a list of activities making up the workflow. These activities are then manually fed into the visualization engine to create the simulations. Currently, this stage of visualization process is completed offline. VR simulations created in this manner present a simulated view of real-world events. This is valuable to clinicians and researchers in highlighting the main events in the workflow within the context of the clinical environment.

4. System evaluation

The hypothesis being evaluated is that clinical activities involving movement patterns can be recognized by the HMM based

Fig. 5. Virtual trauma unit for workflow visualization.

Table 1
Activity list and corresponding clinical descriptions.

Activity	Movement	Clinical description
A1	1-to-2	Paged physician/nurse tends to patient on bed 1
A2	2-to-3	Physician/nurse moves to treat patient on bed 2
A3	3-to-4	Physician/nurse leaves trauma through entry/exit 1 after visiting patient on bed 2
A4	4-to-5	Physician/nurse enters trauma through entry/exit 1 and attends to the phone
A5	5-to-6	Physician/nurse after attending to a phone call move to use the computer at the nurse station
A6	6-to-1	Physician/nurse leaves trauma through entry/exit 2
A7	1-to-4	Physician/nurse enter and leave trauma
A8	4-to-6	Physician/nurse enter trauma through entry/exit 1 and move to use the computer at the nurse station
A9	6-to-2	After using the computer physician/nurse move to treat patient on bed 1
A10	2-to-4	After visiting patient on bed 1, physician/nurse leaves trauma through entry/exit 1
A11	5-to-1	After attending a phone call, physician/nurse leaves trauma through entry/exit 2
A12	1-to-3	Paged physician/nurse attends to patient on bed 2
A13	3-to-5	After visiting patient on bed 2 physician seeks a phone consult
A14	5-to-2	After completing a phone call physician/nurse moves to treat patient on bed 1
A15	3-to-6	After treating patient on bed 2 physician/nurse move to use the computer at the nurse station

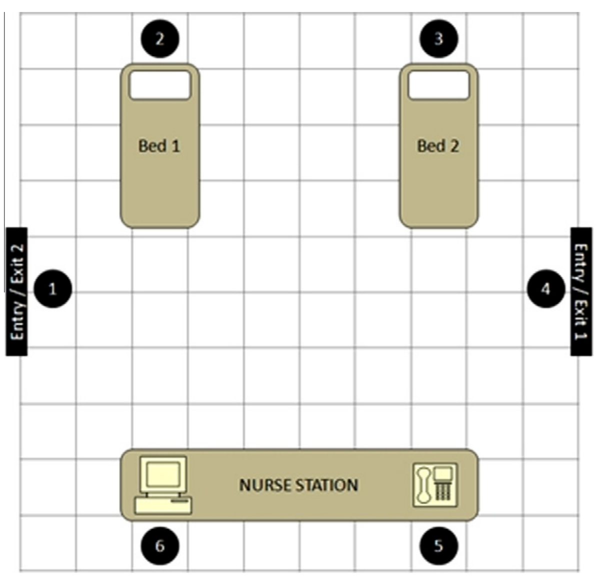

Fig. 6. Test setup for simulated clinical activities.

activity recognition system. In addition, the evaluation seeks to quantify the accuracy of activity recognition (the ratio of the number of correctly identified activities to the total number of test activities). All observations and data gathering were performed after obtaining approval from Institutional Review Boards of involved institutions.

4.1. System evaluation setup

In order to test the HMM based activity recognition system, we simulated 15 trauma activities (listed in Table 1) in a laboratory setting (depicted in Fig. 6), with 10 tags and 6 base stations. These activities were simulations of clinical activities. In order to simulate potential activities in a laboratory setting we observed commonly occurring movement tasks in the trauma unit, an example being "physician moving to phone for a consult" (Activity A13). Fig. 6 depicts the laboratory setup for testing and Table 1 summarizes the movement patterns and clinical descriptions of the 15 activities.

The setup for the testing involved the creation of a 20 ft by 20 ft grid in a laboratory setting (grid lines depicted in Fig. 6). Six base stations (depicted by black solid circles) we placed in pre-defined locations (base 1 and 4 at entry/exit points 2 and 1, respectively; bases 2 and 3 at beds 1 and 2; base 5 at the phone on nurse station; base 6 at the computer on the nurse station). This is congruous with base station setup in the real-world scenario.

We gathered movement data for the 15 sample activities listed in Table 1. For each RID tag–base pair or tag–tag pair an encounter is recorded every 3–4.5 s. This data is captured in a time modulated manner i.e. encounter information is communicated by detecting differences in the time of the encounter rather than the frequency. This results in a sparse matrix when considering the entire tag–base station configuration. Fig. 7 depicts a sample of the matrix generated. The encounters of a tag X with base stations A, B and C (gray filled boxes) are shown in a 60 s long timeline. We use linear interpolation to fill missing data in this sparse matrix. While this methodology provides an RSSI value for all base stations at all instances, it adds some noise to our system that may affect the overall activity recognition accuracy.

For each of these activities, we gathered 10 samples of data. Each sample involved a tagged entity (researcher) following the movement pattern prescribed for the activity. Each sample performed with 10 different tags, totaling 100 samples for each activity. This ensured sufficient randomization of activity movements, accounting for inter-tag variability as well. A total of 1500 samples (15 activities $\times$ 10 samples $\times$ 10 tags) were gathered for testing.

Out of the 100 samples gathered for each activity, 50 samples were used to train the HMM for activity recognition, and the other 50 were used as a testing set to evaluate the algorithms' accuracy.

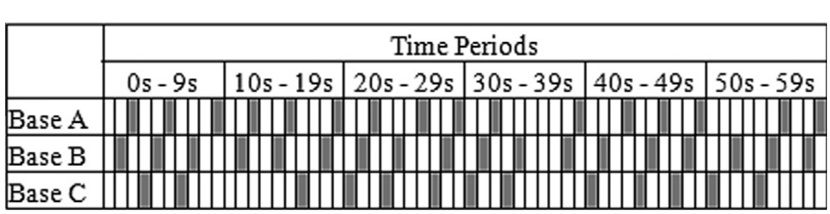

Fig. 7. Sparse matrix of tag–base encounters (gray fill indicating an encounter record with some tag).

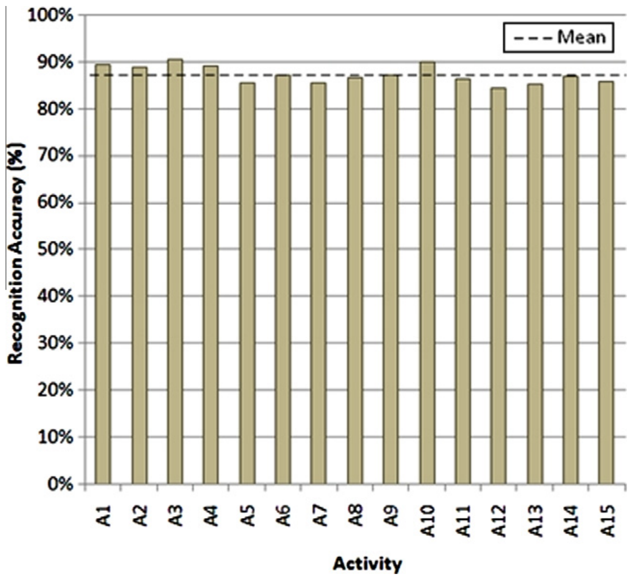

Fig. 8. Hidden Markov Modeling (HMM) based activity recognition results.

4.2. Results of HMM training

Fig. 8 summarizes the recognition accuracy for the 15 motion patterns (A1–A15) elucidated in Table 1. Recognition accuracy is the ratio of the number of activities correctly identified to the total number of activities used for testing. A mean recognition accuracy of 87.5% was obtained, with a maximum of 90.5% and minimum of 84.5%. The analysis of the incorrectly classified test samples revealed that misclassifications were a result of variations in the training set. As discussed previously, HMMs require to be trained on a well controlled sample that best represent the activity. Obtaining training data from real-world scenarios are abound to have variations that may compromise the quality of models generated. This is a limitation of utilizing HMM models with RSSI values alone for activity recognition. Additional sensors such as accelerometers could be utilized in conjunction with RID tags to improve the activity recognition rates. This is a part of our future work.

5. Discussion

We are currently developing models for various activities in the workflow in a trauma unit. The primary challenges to training HMMs for various activities lie in (i) developing a controlled set of samples that best represent the activity being modeled and (ii) the current limitations of RID tags. The linear interpolation adopted for dealing with missing data introduced further errors into the system. Our future work includes improving the recognition accuracy of the system by (i) increasing the sampling frequency of tags, (ii) using alternate methods of interpolation to fill the sparse matrix and (iii) incorporating accelerometers with existing tags to refine data describing the movement. In addition, as the evaluation of the system was conducted in a controlled environment with a limited number of tags, further evaluation and testing with multiple tags in critical care units would be required to complete the validation of this system.

Initial data on the usability of the current application gathered during informal testing revealed positive usability and utility results. The experts pointed to the ability of VR simulations to focus on process and how this could be employed for several applications in the healthcare system. This is a major contribution of this work as to date there is a lack of tools that allow simulation of real-world data in virtual worlds in an automatic manner. This toolkit provides an effective means to accomplish this goal and is specifically geared towards healthcare.

6. Conclusion and future work

The research and system described in this work lays the foundation for development of virtual worlds that are driven by real-world data and real-world needs. This would provide cognitive science researchers another dimension of data for clinical workflow analysis. Similar to a black box, the system provides a means to track all entities and a subset of their activities in the environment. In addition, the methods described can be applied to assess workflow not only in trauma, but to any environment in which multiple team members are sufficiently mobile and geographically dispersed such that their movements might meaningfully reflect their activities.

Currently the system is capable of augmenting conventional data collection mechanisms to offer multidimensional activity information that allows observers to focus on cognitive details rather than simply annotating movement activities. Another use for this system is for the retrospective analysis of data. While existing workflow analysis systems support this type of analysis, the system described in the work can provide additional information that can be used to direct researchers to key events that need to be analyzed further. Visualization of workflow is another benefit of using this tool. In addition to enhancing our understanding of clinical workflow, such visualizations can also be used as an educational tool for training residents and nurse trainees in identifying errors and preventing the potential errors from leading to adverse events.

Future work will include developing more complex activity models and improving the HMM recognition rates. Additional data gathering for HMM modeling of activities and inclusion of audio recordings and accelerometers will enable the enrichment of information gathered. This will provide more information about the workflow and improve the capability of the system to capture and analyze workflow automatically.

Acknowledgments

This research in the paper was supported by an award from the James S. McDonnell Foundation. (Grant #220020152 to Vimla L. Patel). We thank Justin Brown, Joanne Olsen and Gazi Islam for their assistance gathering observations and collecting data for this system.

References

[1] Leape L, Berwick D. Five years after to err is human: what have we learned? JAMA 1995;293:2385–90.
[2] Brixey JJ, Robinson DJ, Turley JP, Zhang J. The roles of MDs and RNs as initiators and recipients of interruptions in workflow. Int J Med Inform 2010;79(6):e109–15.
[3] Malhotra S, Jordan D, Shortliffe E, Patel VL. Workflow modeling in critical care: piecing together your own puzzle. J Biomed Inform 2007;40(2):81–92.
[4] Goldspink C. Modelling social systems as complex: towards a social simulation meta-model. J Artif Soc Soc Simul 2000;3(2). Available from: http://www.soc.surrey.ac.uk/JASSS/3/2/1.html.
[5] Waldrop WM. Complexity: the emerging science at the edge of order and chaos. New York: Simon and Schuster; 1992.
[6] Holland J. Hidden order: how adaptation builds complexity. NY: Addison-Wesley; 1995.
[7] Horsky J, Gutnik L, Patel VL. Technology for emergency care: cognitive and workflow considerations. In: AMIA annual fall symposium 2006; 2006. p. 344–8.
[8] Laxmisan A, Hakimzada F, Sayan OR, Green RA, Zhang J, Patel VL. The multitasking clinician: decision-making and cognitive demand during and after team handoffs in emergency care. Int J Med Inform 2007;76(11–12):801–11.

[9] Alwan M, Dalal S, Mack D, Kell SW, Turner B, Leachtenauer J, et al. Impact of monitoring technology in assisted living: outcome pilot. IEEE Trans Inf Technol Biomed 2006;10(1):192–8.
[10] Østbye T, Lobach DF, Cheesborough D, Lee AMM, Krause KM, Hasselblad V, et al. Evaluation of an infrared/radiofrequency equipment-tracking system in a tertiary care hospital. J Med Syst 2003;27(4):367–80.
[11] Stanford V. Using pervasive computing to deliver elder care. IEEE Pervasive Comput 2002;1(1):10–3.
[12] Matthai P, Kenneth PF, Mike P, Donald JP, Dieter F, Henry K, et al. Inferring activities from interactions with objects. IEEE Pervasive Comput 2004;3(4):50–7.
[13] Schrooyen F, Baert I, Truijen S, Pieters L, Denis T, Williame K, et al. Real time location system over WiFi in a healthcare environment. J Inf Technol Healthc 2006;4(6):401–16.
[14] Chen C, Liu C, Li Y, Chao C, Liu C, Chen C, et al. Pervasive observation medicine: the application of RFID to improve patient safety in observation unit of Hospital Emergency Department. IOS Press; 2005.
[15] Fry EA, Lenert LA. MASCAL: RFID tracking of patients, staff and equipment to enhance hospital response to mass casualty events. AMIA Ann Symp Proc 2005:261–5.
[16] Hauptmann AG, Gao J, Yan R, Qi Y, Yang J, Wactlar HD. Automated analysis of nursing home observations. IEEE Pervasive Comput 2004;3(2):15–21.
[17] Islam GS. An indoor location finding and fall detection system using sensor integration with RFID network and accelerometer data. Tempe: Arizona State University; 2009.
[18] Kahol K, Smith M, Ferrara J, Tripathi P, Leibowitz A, Panchanathan S. Gesture based hand movement analysis and haptic feedback for surgical training [selected as best poster]. In: Poster presentation at the 14th meeting of medicine meets virtual reality medicine meets virtual reality 14, 2006.
[19] Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 1989;77(2):257–86.
[20] Dugad R, Desai U. A tutorial on hidden Markov models. Technical report SPANN-961 signal processing and artificial neural networks laboratory, Bombay, India, 1996.
[21] Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance – results of a randomized, double-blinded study. Ann Surg 2002;236(4):458–64.
[22] Kahol K, Vankipuram M, Smith ML. Cognitive simulators for medical education and training. J Biomed Inform 2009;42(4):593–694.
[23] Ahlberg G, Heikkinen T, Iselius L, Leijonmarck CE, Rutqvist J, Arvidsson D. Does training in a virtual reality simulator improve surgical performance? Surg Endosc 2002;16(1):126–9.