Proactive Workflow Modeling by Stochastic Processes with Application to Healthcare Operation and Management

Chuanren Liu
Rutgers, The State University
of New Jersey
New Jersey, USA
chuanren.liu@rutgers.edu

Yong Ge
University of North Carolina at
Charlotte
North Carolina, USA
yong.ge@uncc.edu

Hui Xiong*
Rutgers, The State University
of New Jersey
New Jersey, USA
hxiong@rutgers.edu

Keli Xiao
Stony Brook University
New York, USA
keli.xiao@stonybrook.edu

Wei Geng
Awarepoint Corporation
California, USA
wgeng@awarepoint.com

Matt Perkins
Awarepoint Corporation
California, USA
mperkins@awarepoint.com

ABSTRACT

Advances in real-time location system (RTLS) solutions have enabled us to collect massive amounts of fine-grained semantically rich location traces, which provide unparalleled opportunities for understanding human activities and discovering useful knowledge. This, in turn, delivers intelligence for real-time decision making in various fields, such as workflow management. Indeed, it is a new paradigm for workflow modeling by the knowledge discovery in location traces. To that end, in this paper, we provide a focused study of workflow modeling by the integrated analysis of indoor location traces in the hospital environment. In comparison with conventional workflow modeling based on passive workflow logs, one salient feature of our approach is that it can proactively unravel the workflow patterns hidden in the location traces, by automatically constructing the workflow states and estimating parameters describing the transition patterns of moving objects. Specifically, to determine a meaningful granularity for the model, the workflow states are first constructed as regions associated with specific healthcare activities. Then, we transform the original indoor location traces to the sequences of workflow states and model the workflow transition patterns by finite state machines. Furthermore, we leverage the correlations in the location traces between related types of medical devices to reinforce the modeling performance and enable more applications. The results show that the proposed framework can not only model the workflow patterns effectively, but also have managerial applications in workflow monitoring, auditing, and inspection of workflow compliance, which are critical in the healthcare industry.

---

*Contact Author.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; H.2.8 [**Database Management**]: Database Applications— *Data Mining*

General Terms

Algorithms, Application

Keywords

Indoor Location Traces, Workflow Modeling, Healthcare Operation and Management

1. INTRODUCTION

Recent years have witnessed the increasing deployment of real-time location systems (RTLS) in hospitals. These RTLS solutions, as shown in Figure 1, allow people to track and monitor the movement of moving objects (e.g., medical devices, doctors and patients), and the interactions between people and medical devices. However, currently such systems are only used for basic tasks, such as simply locating a wheel chair and checking the availability of an inpatient bed. Meanwhile, hospital managers are still facing many managerial challenges. Particularly, there are three important tasks of operation and management in hospital: *workflow monitoring*, *workflow auditing*, and inspection of *workflow compliance*. For example, many healthcare providers have their own work protocols to ensure that the healthcare practices are executed in a controlled manner. Non-compliance practices may be costly and expose the healthcare providers to severe risks, such as litigation, prosecution and damage to brand reputation. However, it is not an easy task to systematically evaluate the degree to which the ongoing healthcare processes are compliant with the predefined protocols. Thus, effective inspection of workflow compliance is very needed by healthcare providers to maintain the performance and reputation of their services.

Conventionally, to accomplish these tasks, hospital managers mainly rely on inspecting detailed workflow logs, which are managerially daunting. In particular, the logs can be heterogeneous in format and stored in different media. For example, many medical records are still documented in papers, which are not easy to be systematically processed and

analyzed in real time. More importantly, these workflow logs are provided passively by the personnel, which might be biased. In contrast, a proactive approach requiring minimum human intervention would be a promising complement for the workflow management tasks. Indeed, the existing RTLS and collected location traces provide unparalleled opportunities for us to develop new ways to help with the workflow management tasks. To this end, this paper provides a focused study of workflow modeling by the integrated analysis of indoor location traces in the hospital environments. One major goal is to proactively unravel the workflow patterns hidden in the location traces, by automatically constructing the workflow states and estimating parameters describing the transition patterns of medical devices. The discovered knowledge from the indoor location traces are then transformed to actionable intelligence for healthcare operation and management. Indeed, the learned workflow model is valuable in that a wide range of practical problems in hospital environments can be relieved with the modeling results.

Although modeling the workflow patterns in the hospital environments is of significant managerial value, it is a non-trivial task to systematically construct and estimate the models with the massive indoor location traces. First, there are some inherent characteristics in the indoor scenario, which pose new challenges to the workflow analysis.

- The granularity and quality of indoor location traces captured by wireless location systems might vary a lot from place to place due to several factors: the density of sensor receivers, the underlying localization techniques, and the sensor device itself (e.g., the battery power percentage).

- The topologies of indoor space are often much more complex than outdoor space. Therefore, some fundamental assumptions of outdoor scenarios may not hold for indoor scenarios. For example, the widely-used similarity measurements for outdoor trajectories based on geometry distance [4] or overlapping degree [6] are not much meaningful for indoor location traces.

- The structure of indoor space is very dynamic and the modeling framework should be automatic and adaptive accordingly. For instance, the structure and utility of many modern buildings may be frequently changed per the ongoing needs. Such dynamic changes will alter the semantic of indoor location traces.

As a result, even there are intensive works on the analysis of outdoor location traces [6, 12, 5], most of them are not suitable or practical for modeling the indoor location traces of medical devices. Moreover, the results of those methods are not very helpful towards the workflow analysis in hospital environments. For example, the method developed by Giannotti et al. [6] aims for the discovery of frequent trajectory patterns based on given thresholds, i.e., minimum support and time tolerance, from outdoor location traces. However, such frequent patterns cannot provide a parsimonious description for healthcare activities in hospitals. For example, to help with workflow compliance inspection, we need to consider all activities rather than only a subset of moving patterns, i.e., frequent patterns. Similarly, the periodic patterns mined from outdoor spatio-temporal data by Mamoulis et al. [12] are also a subset of moving activities and cannot fully meet the needs of workflow monitoring, auditing, or compliance inspection in a hospital. While Yin

Figure 1: A demonstration of the real-time location system in hospitals. At the bottom layer, the moving objects (medical devices, patients, doctors, etc.) are attached with sensor tags, which send signals to the sensor receivers. The sensor receivers at the middle layer transmit the signal data to the network bridges. At the top layer, network bridges connected with data/application servers will collect the signal data and the data/application servers will calculate locations of the tracked objects.

et al. [17] proposed stochastic models for indoor activities, the purpose of these models is to predict the goals of activities, rather than summarize overall activity patterns in descriptive ways. Besides, their methods are mostly supervised and need sufficient training data.

To accommodate the unique characteristics of indoor location traces of medical devices in a hospital, in this paper, we propose a stochastic process-based framework to model the healthcare workflow. First, to extract each workflow state which is related to a particular type of healthcare activity, we develop a density-based clustering algorithm to partition the indoor space. Particularly, to incorporate the complex indoor topologies, we compute the neighborhood of a location record and its density in the indoor space based on the transition history rather than the geometry distance. With the clustering results, we further transform the original location traces into the sequences of workflow states. Then we analyze the workflow sequences by modeling the transitions among states with finite state machines.

Second, due to the dynamic indoor structure, we need an automatic and adaptive modeling framework to support the critical applications in real time. However, during some specific periods of observation, the estimation of parameters in the finite state machines may be unstable due to the data scarcity issue of the location traces. Meanwhile, there is some natural correlation in the location traces of a group of medical devices, which are often used together for a particular task. By leveraging such correlation, we may overcome the data scarcity issue and improve the robustness of the parameter estimation for the finite state machines. Therefore, we explore MAP (Maximum A Posteriori) estimation in the Hidden Markov Random Fields (HMRFs) to estimate the transition probabilities, which can effectively leverage the correlation in the location traces of related medical devices.

Moreover, in addition to the correlation relationship, we will be able to integrate more prior or domain information about the workflow of medical devices with the HMRFs, which will further improve the effectiveness of our workflow modeling.

Finally, we demonstrate the effectiveness of our models with the real-world data collected from multiple hospitals in US. We have also implemented and deployed a management information system, *HISflow*, based on our methods to show how the discovered knowledge can help with the three important managerial tasks in hospitals: workflow monitoring, auditing, and inspection of workflow compliance.

2. PRELIMINARIES AND PROBLEM FORMULATION

In this section, we first introduce the data format of the indoor location traces. Then, we formally formulate the problem of workflow modeling for medical devices.

2.1 Data Description and Transformation

Our location traces (trajectories) of medical devices are collected from indoor environments of several US hospitals. The location trace of an object O is represented as a sequence: $O_t = (L_1, L_2, ...)$, where L_i represents the ith record in the sequence. $L_i = (start_i, end_i; x_i, y_i, z_i)$ contains the specific coordinate and the corresponding timestamp when that record is recorded, where $start_i$ and end_i are the start time and the end time of L_i. In other words, during the time frame from $start_i$ to end_i, the object O stays at the coordinate (x_i, y_i, z_i) in a three-dimensional indoor space.

However, the indoor wireless communication may be interrupted by environmental factors and lead to some errors or noises for the localization of moving objects. Therefore, a coordinate localized by the RTLS might not indicate the exact position, but a small area surrounding the coordinate. In addition, it may not be meaningful to directly use these coordinates for workflow modeling in the indoor space. For example, for two recorded coordinates which are close to each other on the same floor, although the geometry distance of them is very small, the actual moving distance from one coordinate to another one may be very long when there is a wall between these two coordinates.

To cope with these challenges, we normalize the original location traces for workflow modeling. Specifically, we project each raw coordinate to a semantic location of the building, such as a room in the hospital, based on the floor maps of the building. For the data and maps in this study, each hallway is also treated as a room and some long hallways have been segmented into several small rooms. Then, we map $L_i = (start_i, end_i; x_i, y_i, z_i)$ to $L_i^* = (start_i, end_i, r_i)$, where r_i is the room containing the coordinate (x_i, y_i, z_i). After this projection, two neighboring coordinates of the raw location traces may be mapped into the same room. In practice, we merge these consecutive records within the same room to one union record. Specifically, if $i < j$ and $r_i = r_{i+1} = \cdots = r_j$, we replace the subsequence $(L_i, L_{i+1}, \cdots, L_j)$ with only one record $L_i^* = (start_i, end_j, r_i)$.

Now, each raw location record L_i is mapped to a symbolic room (graph node), and each location trace $Tr = (L_1^*, L_2^*, \cdots)$ is transformed to a symbolic sequence (traveling path). This data preprocessing drastically reduces the computational cost, since we significantly reduce the number of records in the data after the projection. Also, this pre-

processing step greatly smooths out the noise and alleviates the impact of errors on the workflow modeling tasks.

2.2 Concepts for Workflow Modeling

The ultimate goal of our workflow pattern modeling is to automatically summarize the healthcare activities in a systematic manner. To this end, the concept of workflow patterns is actually a hierarchy with several levels. At the lowest level, the location trace itself can be seen as an instance of workflow patterns. For example, in Figure 2, we show location traces of an infusion pump with red line segments. However, it is difficult to comprehensively understand the pattern hidden in the raw location traces.
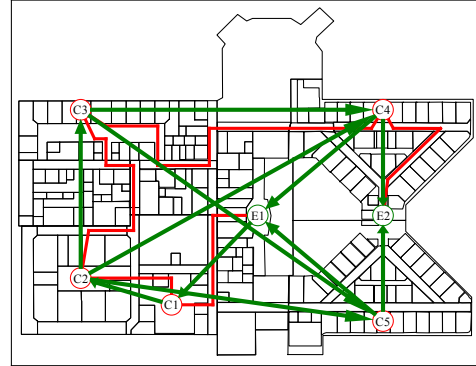

Figure 2: The workflow instances of infusion pump.

Locations	Functions
C_1	Post-anesthesia care unit (PACU)
C_2	Operating room (OR)
C_3	Intensive care unit (ICU)
C_4, C_5	Patient care unit (PCU)
E_1, E_2	Elevator

Table 1: The functions of key locations in Figure 2.

On the other hand, three workflow stages at the highest level are widely used in the healthcare industry to describe the workflow logistics: *preprocessing maintenance stage*, *in-use stage*, *postprocessing maintenance stage*. The *in-use stage* of a medical device corresponds to the period when it is used for any healthcare purpose. Before the *in-use stage*, a device is in the *preprocessing maintenance stage*, e.g., held in the storage room. After the *in-use stage*, the device must go through the *postprocessing maintenance stage* before the next circle of use. These maintenance processes include cleaning up, sterilization, disinfection, etc. However, modeling the workflow patterns based on these three stages is too coarse to get useful results to help with tasks of operation and management in a hospital. In particular, one important factor missed by the workflow stages is the spatial information, and the modeling results based on these high level stages cannot facilitate the workflow evaluation from the spatial perspectives, such as the evaluation for the utilization efficiency of the indoor space. Therefore, we need to model the workflow patterns using a middle-level representation of location traces of medical devices.

2.2.1 Workflow States

To that end, we introduce an important concept, *workflow state*, which will serve as a basis for our workflow modeling at

the proper level of granularity. Indeed, to better understand the location traces, we need to annotate a few key locations in these trajectories. For example, as shown in Figure 2, we annotate location traces with important location spots: C_i $(i = 1, \cdots, 5)$. The medical functions of these spots are summarized in Table 1. With these annotations, the location trace in red in Figure 2 can be represented as $C_1 \mapsto C_2 \mapsto C_3 \mapsto C_4$. Such representation with annotated spots makes it easy to understand the workflow behind the location trace.

In fact, Figure 2 is the map of the second floor of a hospital building, which is centered around E_1, the elevator connected to the basement. The red location trace of one infusion pump starts from the storage room in the basement to the elevator E_1. After E_1, the first spot C_1 is Post-Anesthesia Care Unit (PACU) where the patient is ready for medical procedures such as surgery and the medical devices are attached with the patients. Spot C_2 is the area of operating rooms (OR) where medical procedures happen. After the medical procedure, the patients and the medical devices are moved to C_3, Intensive Care Unit (ICU), based on the medical needs. When the situation of the patient becomes stable, the patient and the medical devices being used will be further moved to C_4, Patient Care Unit (PCU), before the patient is discharged. After the patient is discharged, the medical devices will be moved through elevator E_2 to the basement for cleaning up in a disinfection room and depositing into the storage rooms.

From the above example, we know that each spot with a particular healthcare purpose corresponds to a workflow state, and the location traces represented with such workflow states are very meaningful for understanding and modeling the workflow patterns. Therefore, we formally define the workflow states in the context of hospitals as follows:

DEFINITION 1 (WORKFLOW STATE). *A workflow state is a location spot in the indoor space where specific healthcare activities frequently happen.*

According to this definition, a workflow state might be a cluster of location records with specific activities. In the above example, we have workflow states C_1, C_2, C_3, C_4. By representing the location traces with workflow states, the workflow patterns are more clear to be understood and inferred than that with symbolic rooms or high level stages.

2.3 The Problem Statement of Workflow Pattern Modeling

With workflow states, a further task is to describe the transition patterns of the moving objects among the workflow states. As shown in Figure 2, the transition from one state may go to several different states. For example, when the situation of a patient is stable after the medical operation at C_2, the medical devices and the patient might be moved directly to PCU without passing through ICU (C_3). Also, there are more than one PCU spots in the hospital. For example, in addition to C_4, we have C_5 as another PCU spot in Figure 2. Note that, Figure 2 is only one floor in one building. In fact, we may have many workflow states since there might be several buildings and many floors in one hospital. Thus, the overall workflow patterns are indeed much more complex than the examples discussed above.

Furthermore, we have multiple types of medical devices moving around in the indoor space. In hospitals, many different medical devices are often used together for a particular task. Therefore, although different types of medical

devices have different workflow patterns, there is some natural correlation among their location traces. Modeling such correlation not only helps reinforce the robustness of the workflow modeling, but also provides better understanding of the overall workflow patterns.

Now, we can formally state the problem of workflow pattern modeling in the indoor healthcare environments.

PROBLEM 1 (WORKFLOW PATTERN MODELING). *Given the collected location traces ($\{Tr\}$) of moving objects, such as medical devices, workflow pattern modeling is to discover the workflow knowledge including workflow states (C), and parameters (μ) describing the transition patterns of the moving objects among the workflow states. In addition, we would also like to learn the correlation measurement (s) among the transition patterns for different types of moving objects.*

To accomplish these tasks, in Section 3, we first propose a clustering algorithm to construct the workflow states. Next, in Section 4, we model the transition patterns among the workflow states with finite state machines. In Section 5, we further show how to learn the correlation measurement and build the transition models in a reinforcement way.

3. WORKFLOW STATE CONSTRUCTION

Inspired by Liu et al. [10], we develop a DBSCAN-style clustering algorithm to construct the workflow states from the location traces. This density-based method is adopted to meet the requirements for the workflow analysis. Particularly, the clustering method should be able to automatically determine the number of clusters, and detect clusters of different densities and arbitrary shapes. The detected clusters should also be spatially contiguous, since we want to identify areas of semantically meaningful locations, such as '2nd floor northeast patient rooms' and 'basement central storage rooms'. However, even if we adopt the framework in [10], we have to develop some new methods to address the challenges caused by the unique characteristics of the workflow traces. First, we need a meaningful neighborhood definition for the location records in the indoor space. Second, to cope with the definition of workflow states, we need a density definition associated with workflow activities to detect clusters of different densities.

To address these challenges, let us first introduce how to define the neighborhood of a location record by exploiting the topologies of the indoor space. Since the location traces are represented symbolically, the indoor space map can be modeled with a corresponding symbolic graph. With all the symbolic nodes V as vertices of the graph $G = (V, E)$, we can use edge E to represent neighboring relationship among the nodes. An edge is added between two vertices if a direct transition path between the corresponding pair of symbolic location nodes exists. Formally, we define the neighboring relationship based on the transition connectivity:

DEFINITION 2 (TRANSITION NEIGHBORHOOD). *For room r_0, we denote*

$$TN(r_0) = \{r \mid \exists\, Tr, (rr_0) \in Tr\}$$

as the transition neighboring rooms of r_0. One room r is said to be transition neighbor of another room r_0 if one can transit from r to r_0 without passing through other rooms.

With Definition 2, we define the neighborhood of a location record in the symbolic space:

DEFINITION 3. *The neighborhood of a location record $L_0 = (start_0, end_0, r_0)$ is a set of location records observed in the rooms that are involved in transitions to r_0, denoted by*

$$N(L_0) = \{L = (start, end, r) | r \in TN(r_0)\}.$$

Thus, instead of identifying the neighborhood of a location record based on a distance threshold as used in DBSCAN, we define the neighborhood of a location record by directly querying the transition history from the data. In this way, we avoid the use of parameters, such as the distance threshold, which DBSCAN is very sensitive to.

To detect the clusters of different densities, we propose a *weighted density* definition to measure the density of individual neighborhood and identify core location records. If we know the distribution of waiting time T in room r_0, then a natural choice to weight the location record $L = (start, end, r_0)$ is the p-value $w(L) = \Pr(T \leq d)$, where $d = end - start$ is the waiting time of L. Intuitively, a longer waiting time d will lead to a bigger weight, as it takes time to carry out medical procedures and we are identifying the spot where specific medical activities frequently happen. If we do not know the distribution of T, which often happens in reality, we propose to estimate the weight as

$$w(L) = \frac{|\{(start, end, r) | r = r_0, end - start \leq d\}|}{|\{(start, end, r | r = r_0\}|}.$$

With this formulation, we have the following definition of the weighted density for a neighborhood:

DEFINITION 4. *The weighted density of a neighborhood $N(L_0)$ is defined as:*

$$w(N(L_0)) = \frac{\sum_{L=(start, end, r) \in N(L_0)} w(L)(end - start)}{|N(L_0)|}.$$

EXAMPLE. By the definition above, suppose $N(L_0) = \{L = (start, end, r_0)\}$, let us calculate the weighted density for the following distributions of waiting times:

- $D_1 = \{1, 2, \cdots, 10\}$, 10 observations.
- $D_2 = \{5.5, 5.5, \cdots, 5.5\}$, 10 observations.
- $D_3 = \{1, 1, \cdots, 1\}$, 55 observations.

Although the total waiting time of D_1, D_2 and D_3 is the same as 55, our weighted densities are 3.85, 5.5, 1 respectively. For D_3, the neighborhood $N(L_0)$ is more likely in hallway or in front of elevator. Although the total waiting time in these places is long due to a lot of trajectories passing around, there is no long period of waiting time for carrying out medical procedures. Our weighted density definition can underrate this kind of situations. Similarly, observations with shorter waiting time in D_1 are also weighted lower. □

Finally, we compare this estimated weighted density with a user-specified threshold θ to decide if a location record is a core point or not. Although there is similar parameter in DBSCAN, i.e., minimum number of points required to form a cluster, the parameter θ is not the count of points but a weight measurement based on historical statistics. Due to the space limit, readers are referred to Liu et al. [10] for details of the algorithm, except that the neighborhood and density definitions should be replaced by Definition 3 and 4.

4. WORKFLOW TRANSITION ESTIMATION

With the constructed workflow states C, we can transform the original location trace into a workflow sequence $Tr = (s_1, s_2, \cdots, s_l)$, where $s_i = (c_i, d_i)$, $c_i \in C$ is the ith workflow state and d_i is the waiting duration in c_i. This transformation makes it feasible to model the workflow sequences of the objects moving among the indoor states with a continuous time Markov chain (CTMC). Specifically, with the constructed workflow states C as stochastic states in the workflow processes, a CTMC(μ) is defined with two sets of parameters $\mu = (P, q)$: waiting time parameters q_c for each state $c \in C$, and state transition probabilities $P_{cc'}$ from state c to c'. Let $S(t) \in C$ denote the state of the process at time $t \in [0, \infty)$. CTMC stays in $S(t) = c$ at time instant t for a random amount of duration $D_c \sim \text{Exp}(q_c)$. When CTMC finally leaves c, the next state of the system is c' with probability $P_{cc'}$ where $c' \neq c$. For $c = c'$, we let $P_{cc} = 0$. With these parameters, for the process Tr starting from s_1 and ending at time $d = \sum_{i=1}^{l} d_i$, the probability of Tr is

$$\Pr(Tr|P, q) = \exp(-q_{c_l} d_l) \prod_{i=1}^{l-1} q_{c_i} \exp(-q_{c_i} d_i) P_{c_i c_{i+1}}. \quad (1)$$

Without causing confusion, in the following, we denote states in C by integers $1, 2, \cdots, |C|$ for simplicity. For a set of processes $\{Tr\}$, the probability calculated above can also be expressed in terms of N and D, where N_{ij} is the transition count from state i to state j, and $D_i = \int [S(t) = i] dt$ is the total waiting time in state i. Specifically,

$$\prod_{Tr} \Pr(Tr|P, q) = \Pr(N, D|P, q)$$

$$= \left(\prod_{ij} P_{ij}^{N_{ij}}\right) \left(\prod_i q_i^{N_i} \exp(-q_i D_i)\right),$$

Here, $N_i = \sum_j N_{ij}$ is the transition counts from state i. Note that by definition we have $N_{ii} = 0$ and we compute $P_{ii}^{N_{ii}} = 1$.

One important fact in the likelihood of parameters P, q is that, we can estimate P_{ij} and q_i separately. That is, $\Pr(N, D|P, q) = \Pr(N|P) \Pr(N, D|q)$, where:

$$\Pr(N|P) = \prod_{i,j} P_{ij}^{N_{ij}} \quad (2)$$

$$\Pr(N, D|q) = \prod_i q_i^{N_i} \exp(-q_i D_i) \quad (3)$$

The solution to maximize $\Pr(N, D|q)$ is $q_i = \frac{N_i}{D_i}$. In other words, q_i is estimated by the reciprocal of the mean waiting time at state i. In the following, we will focus on optimizing $\Pr(N|P)$ to estimate the transition probability matrix P.

Under the normalization constraints $\sum_j P_{ij} = 1$ and $P_{ij} \geq 0$, it is straightforward to estimate the parameters P using a maximum likelihood estimation approach. However, the result estimated in this way will be unstable if the training data is not sufficient (e.g., when analyzing the daily data instead of monthly or yearly data). Fortunately, the data collected from different types of moving objects may help mutually reinforce each other to estimate the moving patterns in a robust way. To leverage this potential, we introduce the Maximum A Posteriori (MAP) estimation approach which simultaneously estimates transition models for multiple types of moving objects. More importantly, with
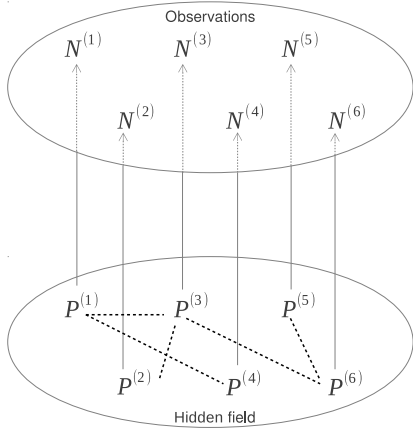
Figure 3: A Hidden Markov Random Field

this approach, we will also be able to understand the collaborating relationship among different types of medical devices.

5. MAP ESTIMATION IN HMRFS

For the healthcare procedures of one patient, multiple types of medical devices are often needed. Thus, these different types of moving objects transit together and produce the correlated location traces. Therefore, although different medical devices follow different patterns, their workflow models can be built jointly with the consideration of these correlations. To this end, we first piece together the workflow states for all the K types of moving objects:

$$C = \cup_{k=1}^{K} C^{(k)}$$

where $C^{(k)}$ is the workflow states of the kth type of moving objects. Note that, when $C_1 \in C^{(1)}$ and $C_2 \in C^{(2)}$ overlap, i.e., $C_1 \cap C_2 \neq \emptyset$, we merge them together as $C_1 \cup C_2 \in C$. Now, all the K types of transition sequences can be expressed with the common transition states C, while each type has one transition count matrix $N^{(k)}$ and one transition probability matrix $P^{(k)}$. To estimate the K transition probability matrices $\{P^{(k)} | 1 \leq k \leq K\}$ together based on $\{N^{(k)} | 1 \leq k \leq K\}$, we use Hidden Markov Random Fields (HMRFs) to leverage the correlation among all types of objects. Specifically, an HMRF has the following components:

Hidden field: In our model, the hidden field includes the transition matrices $P^{(k)}$ for $1 \leq k \leq K$ for all types of objects, whose values are hidden or unobservable.

Observations: The observed data includes the transition counts $N^{(k)}$ for $1 \leq k \leq K$.

As shown in Figure 3, there is also a connecting structure in the hidden field. We would like to reinforce the estimation of the connected parameters by leveraging the prior distribution of the random field. For our problem, we consider pairwise correlation among different types of objects, so the prior depends on $\sum_{k_1, k_2} V(k_1, k_2)$, where $V(k_1, k_2) \geq 0$ is the potential of the link between $P^{(k_1)}$ and $P^{(k_2)}$. Specifically, we have $\Pr(P) \propto \exp(-\sum_{k_1, k_2} V(k_1, k_2))$. By applying this prior, the objective function to be maximized is:

$$\Pr(P|N) \propto \exp(-\sum_{k_1, k_2} V(k_1, k_2)) \prod_{k,i,j} (P_{ij}^{(k)})^{N_{ij}^{(k)}}$$

or equivalently

$$\log \Pr(P|N) = -\sum_{k_1, k_2} V(k_1, k_2) + \sum_{k,i,j} N_{ij}^{(k)} \log P_{ij}^{(k)}. \quad (4)$$

Note that a normalization constant is omitted in the log-likelihood equation.

5.1 Adaptive Parameterized Correlation

HMRFs can integrate the knowledge about the neighborhood structure by formulating the potential $V(k_1, k_2)$. In the following, we demonstrate this framework by an unsupervised approach, where the neighborhood structure itself is also learned from the data. Specifically, we define the potential between them as

$$V(k_1, k_2) = w \times s(k_1, k_2) \times \text{Tr}(P^{(k_1)} P^{(k_2)}), \quad (5)$$

where $w \geq 0$ is a scaling constant. Here, the correlation $s(k_1, k_2)$ measures the similarity between the transition patterns of types k_1 and k_2. The matrix trace $\text{Tr}(P^{(k_1)} P^{(k_2)}) = \sum_{ij} P_{ij}^{(k_1)} P_{ji}^{(k_2)}$ is actually the probability that we can observe about the contradicting moves. For example, in the extreme case when $\text{Tr}(P^{(k_1)} P^{(k_2)}) = 0$, we have $P_{ij}^{(k_1)} P_{ji}^{(k_2)} = 0$ for any given states i, j. In this case, we have $P_{ji}^{(k_2)} = 0$ only if $P_{ij}^{(k_1)} > 0$. In other words, for two types of objects that have no contradicting workflows, the probability of transition from state j to i is 0 for the second type of objects, only if we have a positive transition probability from state i to j for the first type of objects.

To measure the correlation adaptively with respect to the observed data, we propose to parameterize the Frobenius inner product using a non-negative matrix M, which leads to the following formula:

$$s(k_1, k_2) = \frac{< N^{(k_1)}, N^{(k_2)} >_M}{\|N^{(k_1)}\|_M \|N^{(k_2)}\|_M},$$

where the weighted Frobenius inner product is $< X, Y >_M = \sum_{ij} X_{ij} M_{ij} Y_{ij}$ and weighted Frobenius matrix norm is $\|X\|_M = \sqrt{< X, X >_M}$. Indeed, N is directional observation and $s(k_1, k_2)$ corresponds to a von-Mises Fisher (vMF) distribution [2] as the underlying generative model.

5.2 Iterative Optimization Algorithm

We need an iterative updating algorithm to learn the correlation from the data and optimize the objective function in Equation 4 simultaneously. Specifically, there are two sets of parameters: Frobenius weights M and transition probabilities P. To update the Frobenius weights by increasing Equation 4, a closed-form solution is unattainable. In this case, gradient ascent update can be used alternatively. Particularly, the gradients are

$$\frac{\partial \log \Pr(P|N)}{\partial M_{ij}} = -w \sum_{k_1, k_2} \frac{\partial s(k_1, k_2)}{\partial M_{ij}} \text{Tr}(P^{(k_1)} P^{(k_2)}) \quad (6)$$

and

$$\frac{\partial s(k_1, k_2)}{\partial M_{ij}} = \frac{N_{ij}^{(k_1)} N_{ij}^{(k_2)}}{\|N^{(k_1)}\|_M \|N^{(k_2)}\|_M}$$

$$- \frac{1}{2} \frac{< N^{(k_1)}, N^{(k_2)} >_M}{\|N^{(k_1)}\|_M \|N^{(k_2)}\|_M} \left(\left(\frac{N_{ij}^{(k_1)}}{\|N_{k_1}\|_M} \right)^2 + \left(\frac{N_{ij}^{(k_2)}}{\|N_{k_2}\|_M} \right)^2 \right).$$

With the given M, the optimal P can be obtained by solving the following KKT system:

$$-2\sum_{k'} s(k,k')P_{ji}^{(k')} + N_{ij}^{(k)}/P_{ij}^{(k)} - \lambda_{ki} = 0 \qquad (7)$$

for $1 \leq k \leq K$ and $1 \leq i,j \leq |C|$. Here λ_{ki} is a KKT multiplier. With the normalization constraints $\sum_j P_{ij}^{(k)} = 1$, this quadratic system can be numerically solved. Note that, by iteratively conducting above updates, Equation 4 will be monotonically increased and we can stop the process when a sufficiently accurate solution is reached.

6. EMPIRICAL EVALUATION

Here, we evaluate the performances of our workflow models in indoor healthcare environments. First, we show the identified workflow states and analyze the goodness-of-fit of our transition models with respect to important statistics. Then, we show that the learned correlation can reinforce the modeling performance, especially for real-time applications. The proposed models will also be compared with some baseline models. Finally, we showcase a few intelligent applications in healthcare operation and management that can benefit from the learned knowledge.

6.1 The Experimental Data

We use real-world data sets for validation. These data sets are collected from several hospitals in US. Medical devices operated in these hospitals have been attached with sensor tags and tracked by RTLS developed by Awarepoint Corporation. In Table 2, we show basic statistics of the data collected for various types of medical devices in Hospital 1. Specifically, we will build workflow models for 7 types of medical devices. The 2nd and 3rd columns show the number of medical devices of each type operated in this hospital, and the number of location records collected during the period from January 2011 to August 2011.

Type	#Objects	#Locations	#States	θ seconds
Wheelchair	121	524415	121	1200
PCA II Pump	66	4431	2	1200
Venodyne	403	1588045	74	1200
Feeding Pump	83	157370	58	3600
Heating Pad	1	211	1	1200
PCA Pump	136	231057	74	1800
ETCO2	137	220380	79	1200

Table 2: Basic data statistics in Hospital 1

6.2 The Workflow States

In Hospital 1, there are totally 123 different workflow states identified. The last two columns in Table 2 give the number of workflow states identified for each device type and the corresponding threshold for density-based clustering. We can see that different types of medical devices are used in different ranges of region. For instance, wheelchairs move among 121 workflow states, while feeding pumps transit only in 58 workflow states. We also automatically identified that PCA II Pump and Heating Pad only move in a small range, which means the workflows of these two types of devices are quite simple. Thus, in the following we focus on analyzing the remaining 5 types of devices which have more complex workflow patterns. For the threshold of clustering, we empirically specify the parameter based on the mobility

of medical devices and the classical work time for treatment in order to achieve well-separated workflow states.

The identified workflow states are verified to be semantically meaningful by the domain experts. For example, in Figure 4, we show the constructed workflow states for Venodyne, which moves most frequently, on 4 building floors. A summary of these workflow states is reported in Table 3. Particularly, with the developed clustering algorithm based on the the weighted density (demonstrated as transparency), we can effectively focus on the most important location spots. For example, as shown in Figure 4a, although most of the right part on this floor is planned to be PCU (Patient Care Unit), we find that only a small subset of rooms were frequently used. This is especially surprising in the top right corner, where only one room was used efficiently. These observations lead to an important application, workflow auditing, which we will further discuss in Section 6.5.2.
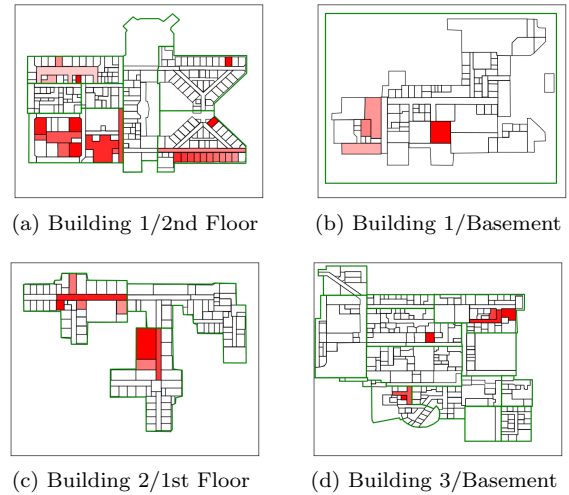

(a) Building 1/2nd Floor   (b) Building 1/Basement

(c) Building 2/1st Floor   (d) Building 3/Basement

Figure 4: Examples of workflow states.

Building/Floor	State	Function
1/2nd	middle bottom	Post-anesthesia care unit
1/2nd	left bottom	Operating room
1/2nd	left top	Intensive care unit
1/2nd	right top	Patient care unit
1/2nd	right middle	Patient care unit
1/2nd	right bottom	Patient care unit
1/Basement	middle	Clinical Engineering
1/Basement	left middle	Storage
1/Basement	left bottom	Decontamination Service
2/1st	left top	Behavioral Health
2/1st	middle	Equipment Services
3/Basement	right top	Storage
3/Basement	middle	Nuclear Medicine
3/Basement	middle bottom	Hemodialysis

Table 3: Semantic summary of workflow states in Figure 4.

6.3 The Goodness-of-Fit

We experimented with different values of the scaling constant w in Equation 5 to build the transition models. If $w = 0$, our framework gives the maximum likelihood estimation. When $w > 0$, the scaling constant controls the tradeoff between maximizing $\Pr(N|P)$ and deconstructing the potentials in the hidden field. Users can chose a value according to the correlation degree of the workflow patterns. In this paper, we perform an exhaustive search in $[\frac{r}{2}, \frac{1+r}{2}]$

with step size 0.1 where r is the average non-weighted Frobenius inner products of the transition matrices.

To measure the goodness-of-fit of the learned transition models, we compute the *average log-loss* with test location traces. For a test location trace $Tr = (L_1, L_2, \cdots, L_t)$ and its corresponding workflow sequence $Tr = (s_1, s_2, \cdots, s_\ell)$ transformed with workflow states C, where $s_i = (c_i, d_i)$, the average log-loss is calculated as

$$L(Tr) = -\frac{1}{\ell} \log \Pr(Tr|P, q), \qquad (8)$$

where $\Pr(Tr|P, q)$ is defined in Equation 1 and $\mu = (P, q)$ is learned from training data. According to Rissanen and Langdon [14], smaller average log-loss implies better compression of the data when using the estimated parameters.

In the following, after constructing the workflow states C, we randomly partition the data into 10 subsets and compute the average log-loss in 10 rounds. In each round, we use 9 of these subsets as training data to estimate the transition parameters $\mu = (P, q)$ and compute $L(Tr)$ for each Tr in the remaining test data. The solid green lines in Figure 5 show the average log-loss along the sequence length ℓ for different types of medical devices. We also show the performances with dashed green lines for the baseline models. The baseline model of each type straightforwardly estimates the transition parameters without considering the correlation among different types of moving objects. In the figure, we can see that our models based on HMRFs achieve lower information loss consistently for all types of moving objects.

In practice, as we will discuss in Section 6.5, sufficient observations might not be available to monitor the ongoing workflow patterns. In this case, our models can perform much better than baselines by holistic estimation. To demonstrate this, we repeat the above comparison by using only 1 subset as the training data. In Figure 5, as shown by the red lines, our models achieve much better performance. Our models not only achieve lower information losses, but also produce robust results without rigid jumps.

6.4   The Simulation Performance

In addition to the goodness-of-fit in terms of information loss, a realistic model should also be able to simulate the modeled process without breaking important characteristics. In the research of time series data, one important characteristic is the autocorrelation function, ACF, which describes the serial dependence structure in the sequence models. In our workflow models, the state space C is categorical, thus we propose to investigate the χ^2 statistic to test the homogeneity. Specifically, for a set of categorical sequences $\{Tr = (c_1, \cdots, c_l)\}$ where $c_i \in C$ and $C = \{1, 2, \cdots, |C|\}$, the *auto χ^2 function*, AXF, is calculated as

$$\text{AXF}(h) = \sum_{i=1}^{|C|} \sum_{j=1}^{|C|} \frac{(M_{ij} - E_{ij})^2}{E_{ij}}$$

where $M_{ij} = \sum_{Tr} \sum_{k=1}^{l-h} [c_k = i, c_{k+h} = j]$, $M_{i*} = \sum_{j=1}^{|C|} M_{ij}$, $M_{*j} = \sum_{i=1}^{|C|} M_{ij}$, $M = \sum_{Tr}(l - h)$ and $E_{ij} = \frac{M_{i*} M_{*j}}{M}$. The degree of freedom of the corresponding χ^2 distribution for $\text{AXF}(h)$ is $(|C| - 1)^2$. We can then calculate the p-value of the test of homogeneity with a null hypothesis that $(c_1, \cdots, c_{l-h})$ and $(c_h, \cdots, c_l)$ are independent. In Figure 6, we show the p-values of $\text{AXF}(h)$ for the raw observations in solid lines. As can be seen, the serial dependence structures

of the simulated processes shown in dashed lines are quite close to that of raw observations. Furthermore, we prefer the p-values that increase rapidly along h. This means the long-term dependence is little in the sequences and thus CTMC is a feasible model.

6.5   The Prototype Applications

The learned workflow model is valuable in that a range of practical problems can benefit from the results of modeling. In this paper, we have implemented a management information system, HISflow, to exploit the discovered knowledge for healthcare operation and management. In the following we elaborate on the techniques used in our implementation.
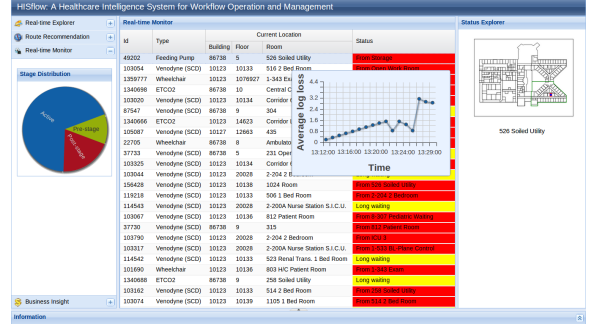

Figure 7: The screenshot of HISflow.

6.5.1   Workflow Monitoring

When the mined workflow patterns are legislated as procedure codes, we can identify abnormal behavior from daily healthcare activities in a real-time manner. When such anomaly occurs, warnings and alerts can be activated by the management system. Such a system helps reduce the risks of faults or accidents of healthcare services. To develop this system, a simple approach is to rank the ongoing trajectories of all monitored medical devices based on average log-loss in Equation 8. In this way, the devices ranked in the top are worth more scrutiny. However, this ranking results might not be intuitive from the management perspective. In fact, it is vital to provide more insights into the cause for the higher log-loss trajectories. To this end, we further identify the recent transition $s_i = (c_i, d_i)$ in $Tr = (s_1, s_2, \cdots, s_l)$ which causes $L(Tr)$ to increase. With the identified s_i, we can then clarify whether the increase is caused by c_i or d_i via comparing $-\log P_{c_{i-1} c_i}$ and $-\log q_{c_i} + q_{c_i} d_i$. If $-\log P_{c_{i-1} c_i}$ is larger, the increase is caused by c_i because the transition coming from c_{i-1} is unlikely to end at c_i. Otherwise, it indicates that the device stays at c_i too long.

A screenshot of the real-time monitoring is shown in Figure 7. As can be seen, the average log-loss of the selected device ($Id = 49202$) increases linearly at the beginning, which means the device stayed at a workflow state c and $-\log q_c + q_c d$ increases linearly with waiting time d. As shown in red, when this device is moved from Storage Room (left middle spot in Figure 4b) to 526 Soiled Utility Room (blue spot in the floor map in Figure 7), it is ranked into the top due to the unlikely transition. We also show the related information in the grid and the location status on the floor map. Such an intuitive and interpretable real-time monitor system is valuable for the hospital managers to improve the quality of healthcare services.
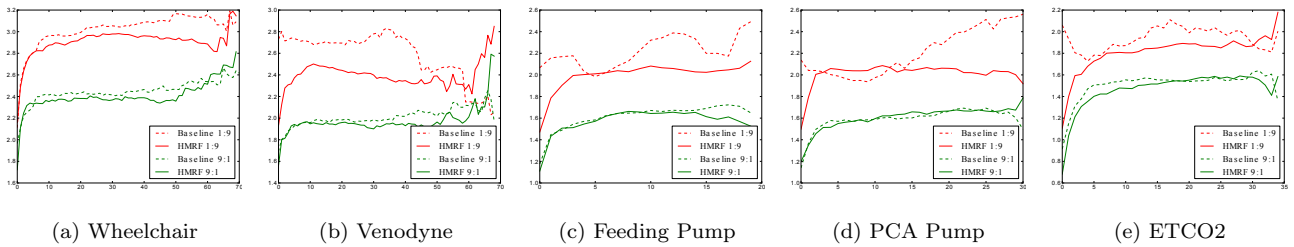
(a) Wheelchair (b) Venodyne (c) Feeding Pump (d) PCA Pump (e) ETCO2

Figure 5: A comparison of the average log-loss. X-axis: the sequence length ℓ. Y-axis: the average log-loss.

(a) Wheelchair (b) Venodyne (c) Feeding Pump (d) PCA Pump (e) ETCO2
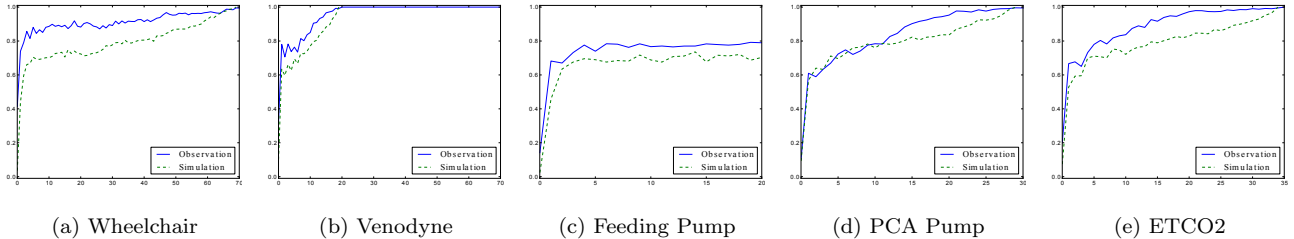
Figure 6: A comparison of the AXF structure. X-axis: the time-lag h in AXF. Y-axis: the p-value of AXF.

6.5.2 Workflow Auditing

Our workflow models can also provide insights into auditing the efficiency of the healthcare services. Particularly, hospital managers would like to know how and how much the indoor spaces and the medical devices are being used. With our models, the utilization efficiency of medical devices and indoor space can be measured with sound statistics.

First, we can use waiting time parameter q in each state c to measure the utilization efficiency of the medical devices. Specifically, a small q indicates that the healthcare procedures taken in these states need a lot of time. For the workflow states corresponding to preprocessing maintenance stage (e.g., Storage rooms) and postprocessing stage (e.g., Soiled Utility Rooms, Decontamination Service Rooms), this small q means the processing procedures are not efficient. Second, to investigate the utilization efficiency of the indoor space, we would like to know the proportion of time the stochastic process spend at each state. In our CTMC model, this proportion at the ith state can be estimated by the stationary probability $r_i = \lim_{t\to\infty} Pr(S(t) = i) = \frac{\pi_i/q_i}{\sum_j \pi_j/q_j}$ where $\pi_j = \sum_i \pi_i P_{ij}$ and $\sum_j \pi_j = 1$.

Our system periodically calculates these statistics which allow hospital managers to audit the utilization efficiency of the medical devices and indoor space. By showing the trends over time and comparisons between different hospitals, one can get both intuitive understanding and quantitative measurement of the ongoing operation performance.

6.5.3 Workflow Compliance Inspection

The procedure of healthcare services is an operation process controlled by the science of medical treatment as well as industry legislation. Non-compliance practices lead to adverse affects such as litigation, prosecution and damage to brand reputation. To inspect the workflow compliance, we can calibrate the classical workflow sequences with respect to given workflow policies. In our stochastic model, the classical workflow sequences can be constructed by finding the most likely trajectories. Note that, in addition to providing statistic evidence on workflow compliance, the constructed classic workflow sequences are also helpful for understanding the ongoing workflow practices. Technically, given the initial state c, the final state c' and a time duration d, find-

ing the most likely trajectory is an optimization problem to $\max_{Tr} Pr(Tr|P, q)$, where $Tr = (s_1, \cdots, s_l)$, $s_i = (c_i, d_i)$, with constraints $c_1 = c, c_l = c'$ and $\sum_{i=1}^{l} d_i = d$. With parameters $\mu = (P, q)$ learned by our framework, this boundary value problem [13] can be solved by dynamic programming algorithms.

7. RELATED WORK

Workflow analysis conventionally relies on detailed workflow logs [1, 15, 7]. The workflow processes are typically represented by activity graphs. Given the execution logs, which are lists of activity records, the workflow mining can be formalized as a graph mining problem by deeming the execution logs as walks on the activity graphs [1]. In practice, there might be discrepancies between the actual workflow processes and the processes perceived by the management. In this case, to discover a completely specified workflow design model, Van der Aalst et al. [15] presented an algorithm to extract a process model from the workflow logs and represent it in terms of a Petri net. Instead of discovering the complete model, Greco et al. [7] later formalized the problem of discovering the most frequent patterns of executions, i.e., the workflow substructures that have been scheduled more frequently and had lead to a desired final configuration. However, these methods rely on the workflow logs which are often recorded by people in the healthcare industry. Thus, the results may be distorted due to the missed activities and bias in the logs. These distorted results can be misleading for many operation and management tasks in hospitals, such as the inspection of workflow compliance. In comparison, as we discussed in Section 1, in this paper we propose a proactive approach to workflow modeling by mining the digital location traces of moving objects automatically recorded by RTLS in the hospitals, which requires minimum amount of human intervention. The statistical modeling results provided by our approach are helpful for a range of operation and management tasks in hospital environments.

In terms of methodology, another category of related work is the modeling and prediction of human activities. For instance, Yin et al. [17] proposed stochastic process models to predict the goals of indoor human activities. Furthermore, for multiple-goal recognition, Chai and Yang [3] proposed

a two-level architecture for behavior modeling and Hu and Yang [8] developed a dynamic Bayesian model where skip-chain conditional random fields were used for modeling interleaving goals. However, these approaches are supervised and require sufficient training data. Also, the activities considered in the above papers are not within the hospital environments, where specialized workflow and activities happen. In the hospital environments, the desired knowledge to be discovered is different from and more complicated than that within other environments, such as school or company.

In terms of analytics of location traces, trajectory pattern mining is also related to this work. For instance, Giannotti et al. [6] introduced trajectory patterns as frequent behaviors in terms of both time and space, where the frequent trajectory patterns are computed based on the given thresholds. In [9], methods were proposed to discover the periodic patterns from spatio-temporal data, where a periodic pattern is defined as a regular activity which periodically happens at certain locations. Also, Yang and Hu [16] proposed methods to discover sequential patterns from imprecise trajectories of moving objects. However, these methods developed for outdoor space are not designed for the purpose of workflow modeling in the indoor hospital environments, and more importantly, the mined frequent patterns cannot provide a parsimonious description of healthcare activities in hospitals and support the applications we have considered.

Finally, the last category of related work is the detection of area-of-interest with trajectory data. For instance, Liu et al. [11] proposed a non-density-based approach, called mobility-based clustering, to identify the hotspots of moving vehicles in an urban area. The key idea is that sample objects are used as "sensors" to perceive the vehicle crowdedness in nearby areas using their instant mobility, rather than the "object representatives". Moreover, Zheng et al. [18] proposed a stay point concept and identified hotspots from human moving trajectory. One location was considered as a hotspot if a lot of moving objects stay nearby over a thresholded time period. In addition, Giannotti et al. [6] used the neighborhood function to model Regions-of-Interest. Basically, they partitioned the spatial space into grids and quantified the interest of each grid with the density and the direction information of each grid. As we have discussed, even some methods mentioned above are very successful for analyzing outdoor location traces, most of them are not applicable to the proposed indoor hospital environments because of the unique characteristics of the indoor space.

8.   CONCLUDING REMARKS

In this paper, we exploited the location traces of medical devices for modeling the healthcare workflow patterns in the hospital environment. Specifically, we developed a stochastic process-based framework, which provides parsimonious descriptions for long location traces. This framework provides new opportunities to understand the logistics of a large hospital in a concise manner. From the application perspective, the discovered knowledge, such as workflow states, transition patterns, and co-transiting relationships, can be integrated for use in the management information system we developed. With this system, we showed that valuable intelligent applications for healthcare operation and management can be enabled to manage, evaluate and optimize the healthcare services.

References

[1] Rakesh Agrawal, Dimitrios Gunopulos, and Frank Leymann. Mining process models from workflow logs. In *EDBT*, 1998.

[2] A. Banerjee, I. Dhillon, J. Ghosh, and S. Sra. Generative model-based clustering of directional data. In *SIGKDD*, 2003.

[3] X. Chai and Q. Yang. Multiple-goal recognition from low-level signals. In *AAAI*, 2005.

[4] Y. Ge, H. Xiong, Z.H. Zhou, H. Ozdemir, J. Yu, and K.C Lee. Top-eye: top-k evolving trajectory outlier detection. In *CIKM*, 2010.

[5] Y. Ge, H. Xiong, C. Liu, and Z.H. Zhou. A taxi driving fraud detection system. In *ICDM*, 2011.

[6] F. Giannotti, M. Nanni, D. Pedreschi, and F. Pinelli. Trajectory pattern mining. In *SIGKDD*, 2007.

[7] Gianluigi Greco, Antonella Guzzo, Giuseppe Manco, and Domenico Sacca. Mining and reasoning on workflows. *TKDE*, 17(4):519–534, 2005.

[8] D.H. Hu and Q. Yang. Cigar: concurrent and interleaving goal and activity recognition. In *AAAI*, 2008.

[9] Z.H. Li, B.L. Ding, J.W. Han, R. Kays, and P. Nye. Mining periodic behaviors for moving objects. In *SIGKDD*, 2010.

[10] C. Liu, Y. Ge, and H. Xiong. A stochastic model for context-aware anomaly detection in indoor location traces. In *ICDM*, 2012.

[11] S.Y. Liu, Y.H. Liu, L.M. Ni, J.P. Fan, and M.L. Li. Towards mobility-based clustering. In *SIGKDD*, 2010.

[13] T.J. Perkins. Maximum likelihood trajectories for continuous-time markov chains. *NIPS*, 2009.

[14] J. Rissanen and G.G. Langdon. Arithmetic coding. *IBM Journal of research and development*, 1979.

[15] Wil Van der Aalst, Ton Weijters, and Laura Maruster. Workflow mining: Discovering process models from event logs. *TKDE*, 16(9):1128–1142, 2004.

[16] J. Yang and M. Hu. Trajpattern: Mining sequential patterns from imprecise trajectories of mobile objects. In *EDBT*, 2006.

[17] J. Yin, D. Shen, Q. Yang, and Z.N. Li. Activity recognition through goal-based segmentation. In *AAAI*, 2005.

[18] Y. Zheng, L. Zhang, X. Xie, and W.Y. Ma. Mining interesting locations and travel sequences from gps trajectories. In *WWW*, 2009.