Title	Detection of Unusual Human Activities Based on Behavior Modeling
Author(s)	Hiraishi, Kunihiko; Kobayashi, Koichi
Citation	12th International Workshop on Discrete Event Systems: 182-187
Issue Date	2014-05-14
Type	Conference Paper
Text version	author
URL	http://hdl.handle.net/10119/12855
Rights	Copyright (C) 2014 International Federation of Automatic Control. Kunihiko Hiraishi and Koichi Kobayashi, 12th International Workshop on Discrete Event Systems, 2014, 182-187. http://dx.doi.org/10.3182/20140514-3-FR-4046.00029. This material is posted here with permission of International Federation of Automatic Control.
Description	

Detection of Unusual Human Activities Based on Behavior Modeling

Kunihiko Hiraishi * **Koichi Kobayashi** *

* *School of Information Science, Japan Advanced Institute of Science and Technology, Nomi, Ishikawa, 923-1292 Japan ({hira,k-kobaya}@jaist.ac.jp)*

Abstract: A type of services that require human physical actions and intelligent decision making exists in various real fields, such as nursing in hospitals and caregiving in nursing homes. In this paper, we propose new formalism for modeling human behavior in such services. Behavior models are estimated from event-logs, and can be used for analysis of human activities. We show two analysis methods: one is to detect unusual human activities that appear in event-logs, and the other is to find staffs that behave differently from others.

Keywords: behavior modeling, N-gram models, cooperative work, event-logs

1. INTRODUCTION

A type of services that require human physical actions and intelligent decision making exists in various real fields, such as nursing in hospitals and caregiving in nursing homes. The project group including authors calls such services "physical and adaptive intelligent services (PAI-services)," and is developing an IT-based system that aims to assist cooperation and knowledge sharing among staffs, and to reduce various kinds of stresses associated with their work (Uchihira (2013)).

One of the research questions that arise in the project is how to evaluate the effectiveness of such a newly introduced system. In other words, we want to know how the system contributes to improving human activities in the services. Traditional approaches are qualitative evaluation based on questionnaires and interviews, and quantitative evaluation based on statistics on the entire activities, such as the length of traffic line and efficiency in handling tasks. However, questionnaires and interviews cannot be done so frequently, and statistics on the entire activities is not suitable for finding unusual human activities that occasionally happen. In this paper, we propose a model-based approach to analysis of human activities in PAI-services.

Theoretical contribution of this paper is to propose new formalism for modeling adaptive and cooperative behavior among concurrently acting people. The formalism is based on discrete-event systems, and has sufficient expressiveness for analyzing human activities in PAI services. Technically, the proposed formalism is a collection of N-gram models with information sharing. We call it *communicating N-gram models*. In the formalism, multiple instances of N-gram models runs concurrently, and event occurrence in each instance of N-gram models may affect other instances of N-gram models.

The obtained behavior models are used for analysis of human activities, especially for detecting unusual human activities. However, the proposed formalism does not have sufficient information for executing discrete-event simulation. Such simpleness of formalism contributes to estimating models by easy calculation on statistics of event occurrence.

The paper is organized as follows. In Section 2, relater work is described. In Section 3, mathematical definitions and notations are presented. In Section 4, definition of communicating N-gram models is given. We also show an estimation method of the models from event-logs. In Section 5, two methods for detecting unusual activities are presented. In Section 6, the proposed detection methods are applied to analysis of event-logs in field experiments. Section 7 is the conclusion.

2. RELATED WORK

As a method for building behavior models from event-logs, process mining is well-known (van der Aalst (2011)). Process mining is a technique for extracting process models from large amount of event-logs output from IT systems. The obtained process models are used for improving processes in organizations. However, the processes we consider here is more complicated and unstructured. For example, tasks may be interrupted by nurse calls.

Behavior of staffs in nursing homes tends to be nondeterministic. The next action is determined by the current situation such as patient's response and availability of various facilities, where the patients' choices are not controllable by the staffs. As formalism that can deal with such nondeterministic behavior, Markov models is well known. In particular, factorial hidden Markov models (Gharramani (1997)) can represent concurrent processes, and interleaved mixture of hidden Markov models (Landwehr (2008)) can handle interruption of processes, both of which often appears in the field experiments. Moreover, there are several results on learning of Markov models (Angluin (1997); Sen (2004)). However, Markov models are not necessarily suitable for modeling unstructured and adaptive behavior, because it is hard to identify global states of Markov models. For modeling adaptive and concurrent behavior, rule-based description gives more simple and flexible way for the modeling. By this reason, we propose to use formalism based on conditional probabilities.

There are various researches on modeling of medical and nursing processes. In (Avrunin (2010)), medical processes are modeled by a process description language and analyzed by formal verification techniques. By observing actual nursing processes, process models of tasks, such as blood transfusion and dripping, are identified.

Application of discrete-event simulation in health care has significantly increased. Comprehensive survey is found in (Thorwarth (2009)). In many cases on the application of discrete-event simulation, performance issues such as analysis on patient queues and waiting time is the main concern, and the results are used for nurse scheduling, resource allocation, and also for change in admission policy and hospital extension.

In (Sundramoorthi (2007)), a stochastic simulation model for nursing activities is derived from real data in a hospital. Behavior models are obtained in the form of classification and regression trees. This approach is similar to that of this paper. Comparing with this modeling technique, the proposed modeling is not designed for the discrete-event simulation of human behavior, but for analysis of human behavior including collaboration of staffs. Behavior models proposed in this paper are microscopic models for limited situation such as activities on the dining time. Moreover, the models are obtained by simple calculation on event occurrence. This enable us to deal with large amount of data.

3. PRELIMINARIES

Let Σ be a finite set of symbols and let Σ^* denote the set of all finite sequences over Σ. For a positive integer N, a sequence of length N is called *an N-gram*. Let $\Sigma^N = \{s \in \Sigma^N \mid |s| = N\}$ be the set of all N-grams over Σ, where $|s|$ denotes the length of sequence s. The i-th symbol of sequence s is denoted by $s_{[i]}$ and the subsequence from the i-th position to the j-th position of s is denoted by $s_{[i,j]}$. In addition, we write $s_{[i,*]}$ to indicate $s_{[i,|s|]}$. Let s and v be sequences over Σ, where $|s| < |v|$. Then the number of occurrences of s as a subsequence of v is denoted by $O_s(v)$.

An N-gram model is a collection of conditional probabilities $Pr(\sigma|y)$, the probability that symbol σ occurs after $(N-1)$-gram y. N-gram models were originally proposed by Shannon (Shannon (1948)). Currently, N-gram models are widely used in text processing. Given a sequence v over Σ and a positive integer N, the maximum likelihood estimation of probabilities in the N-gram model is computed by

$$Pr(\sigma|y) = \frac{O_{y\sigma}(v)}{\sum_{\sigma' \in \Sigma} O_{y\sigma'}(v)} \quad (1)$$

When the length of v is not so large, we use smoothing techniques to estimate the probability for σ with low frequency (Chen (1996)).

A probabilistic automaton is a 6-tuple $G = (X, \Sigma, \delta, P, x_0, F)$, where $X = \{x_1, \cdots, x_n\}$ is the set of states, Σ is the set of symbols, $\delta \subseteq X \times \Sigma \to X$ is the state transition function, $P : X \times \Sigma \to [0, 1]$ is the function defining probability of each state transition, where $\sum_{\sigma \in \Sigma} P(x_i, \sigma) = 1$ holds for all $x_i \in X$, $x_0 \in X$ is the initial state, and $F \subseteq X$ is the set of final states. The underlying Markov chain of G consists of the set of states X and transition probabilities $P_{ij} = P(x_i, \sigma)$ for the $\sigma \in \Sigma$ such that $\delta(x_i, \sigma) = x_j$.

Given an N-gram model, we can obtain a probabilistic automaton $M = (X, \Sigma, \delta, P, x_0, F)$, where X is the set of all $(N-1)$-grams over Σ, δ is defined by $\delta(y, \sigma) := y_{[2,*]}\sigma$, $P(y, \sigma) := Pr(\sigma|y)$, and x_0 and F are arbitrary specified. In addition, we can define the probability $q_y(v) := O_y(v)/\sum_{y' \in \Sigma^{N-1}} O_{y'}(v)$ that each $(N-1)$-gram y occurs in v, where v is the event sequence used for estimating the N-gram model. The probability $q_y(v)$ indicates significance of sequence y in v.

On the other hand, there exists an N-gram model that approximates behavior of a given probabilistic automaton in the steady state. Suppose that a probabilistic automaton G has the steady state and the stationary probability is $\pi = (\pi_1, \cdots, \pi_n)$, where π_i is the probability that the system is in state x_i, then there exists the following N-gram model that approximates the behavior of G: for each $y \in \Sigma^{N-1}$ and $\sigma \in \Sigma$,

$$Pr(\sigma|y) = \sum_{x_i \in X_y} (\pi_i / \sum_{x_j \in X_y} \pi_j) \cdot P(x_i, \sigma) \quad (2)$$

where $X_y = \{x_j \mid \exists x_i \in X : \delta(x_i, y) = x_j\}$. If the underlying Markov chain is ergodic, then estimation by (1) converges to this probability. Moreover, the value of $q_y(w)$ approaches to $\sum_{x_i \in X_y} \pi_i$.

Fig. 1 is a probabilistic automaton whose underlying Markov chain is ergodic. This automaton has the unique stationary distribution $\pi = (35/107, 30/107, 42/107)$ as the solution of equations $\pi = \pi\mathbf{P}$, $\sum_i \pi_i = 1$, where $\mathbf{P} = [P_{ij}]$ is the transition probability matrix. After an occurrence of ab, possible states are 1 or 2. Therefore, the conditional probability $Pr(b|ab)$ is obtained by

$$Pr(b|ab) = \frac{\pi_1}{\pi_1 + \pi_2} \cdot 0.4 + \frac{\pi_2}{\pi_1 + \pi_2} \cdot 0.3 = 23/65.$$

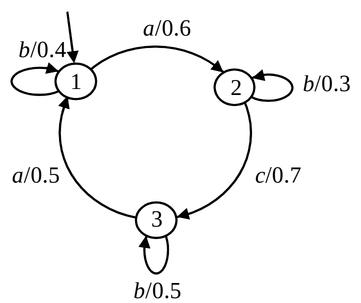

Fig. 1. A probabilistic automaton.

4. FORMALISM FOR BEHAVIOR MODELING

In this section, we describe formalism for modeling adaptive and cooperative human behavior.

4.1 Communicating N-gram Models

We first define an extended N-gram model such that each conditional probability is defined for given attribute values together with past history of events. We call it *an attributed N-gram model* (Hiraishi (2013)). From now on, we call each symbol $\sigma \in \Sigma$ *an event*.

Conditional probabilities in attributed N-gram models are defined in the form $Pr(\sigma \mid \mathbf{a}\!:\!y)$, where $\mathbf{a} = [a_1, \cdots, a_k]$ is a collection of attribute values and y is an $(N-1)$-gram. We require the domain of each a_i is a discrete finite set. We call such a pair $\mathbf{a}\!:\!y$ of attribute values $\mathbf{a}$ and k-gram $y \in \Sigma^k$ *an attributed k-gram*, and the set of all such attributed k-grams is denoted by Σ^k_{attr}. An attributed N-gram model can be seen as a collection of N-gram models each of which is identified by attribute values $\mathbf{a}$.

Let V be the set of all attribute values. *A communicating N-gram model* is a triple $\mathcal{C} = (\mathcal{A}, \Theta^{(0)}, \Delta)$, where $\mathcal{A}$ is an attributed N-gram model, $\Theta^{(0)} = \{\theta^{(0)}_1, \cdots, \theta^{(0)}_n\}$ is the set of initial process instances, and $\Delta : V \times \Sigma \to 2^V$ is the attribute update function. Each process instance at discrete time t is a pair $\theta^{(t)}_i = (\mathbf{a}^{(t)}_i, y^{(t)}_i)$, where $\mathbf{a}^{(t)}_i \in V$ and $y^{(t)}_i \in \Sigma^{N-1}$ [1].

The dynamics of $\mathcal{C}$ are given as follows. The set Σ is partitioned into the set of activity events Σ_{act} and the set of input events Σ_{in}.

Let $\Theta^{(t)} = \{\theta^{(t)}_1, \cdots, \theta^{(t)}_n\}$ be the set of process instances at time t. Then a process instance $\theta^{(t)}_k \in \Theta^{(t)}$ is nondeterministically selected, and an activity event $\sigma \in \Sigma_{act}$ occurs with probability $Pr(\sigma \mid \mathbf{a}^{(t)}_k\!:\!y^{(t)}_k)$. After the occurrence of σ, the discrete time is incremented by one, and process instances at time $t+1$ are determined by $\mathbf{a}^{(t+1)}_i \in \Delta(\mathbf{a}^{(t)}_i, \sigma)$ for all i, and $y^{(t+1)}_k := y^{(t)}_{k_{[2,*]}}\sigma$; $y^{(t+1)}_i := y^{(t)}_i$ $(i \neq k)$.

In addition to the autonomous behavior, attribute values may be changed by input events from outside. If an input event $\gamma \in \Sigma_{in}$ occurs at time t, then process instances at time $t+1$ are given by $\mathbf{a}^{(t+1)}_i \in \Delta(\mathbf{a}^{(t)}_i, \gamma)$ and $y^{(t+1)}_i := y^{(t)}_i$.

A sample path of a communication N-gram model $\mathcal{C}$ is a finite sequence of events that can occur according to the above rule.

Communicating N-gram models are used for modeling the following situation in PAI services:

- There are multiple staffs working in a field. Each staff has a role, and staffs having different roles behave differently. Each staff is modeled by a process instance, and the role is represented by an attribute.
- Communication between staffs exists. This is implemented by events for communication. Receiving messages from other staffs may affect the future behavior of the receiver.
- Change of tasks and roles may be instructed by the person in charge. Emergency calls may suddenly happen and some of the staffs must respond to them. Such calls are modeled by input events.

The idea behind proposing the attributed N-gram models id as follows. We do not have complete knowledge on the state of the target system. The known part of states is represented by attributes and the unknown part is approximated by $(N-1)$-grams, i.e., history of event occurrence.

4.2 Estimating Models from Event Logs

Given event logs, we show how to estimate models in the form of communicating N-gram models.

A pair of attribute values $\mathbf{a}$ and an event σ is called *an attributed event*, and is denoted by $\mathbf{a}\!:\!\sigma$. We assume that an event-log is a finite sequence $L = l_1 l_2 \cdots l_{|L|}$ of attributed events such that each l_j has the following form:

$$l_j = (id_j,\ \mathbf{a}_j\!:\!\sigma_j) \tag{3}$$

where $id_j \in \{1, \cdots, n\}$ corresponds to the index number of a process instance, $\mathbf{a}_j\!:\!\sigma_j$ is an attributed event. In addition to an event-log, we assume that the set of initial process instance $\Theta^{(0)} = \{\theta^{(0)}_1, \cdots, \theta^{(0)}_n\}$ is given.

From an event-log L, we extract the event sequence w^i for process instance i as follows. Let $\mathbf{a}^i_k : \sigma^i_k$ denote the k-th attributed event issued by the i-th process instance. Then we define $w^i := \mathbf{a}^i_1\!:\!\sigma^i_1, \mathbf{a}^i_2\!:\!\sigma^i_2, \cdots, \mathbf{a}^i_{|w^i|}\!:\!\sigma^i_{|w^i|}$. For the indices i and k in $\mathbf{a}^i_k : \sigma^i_k$, let $j_{(i,k)}$ denote the index j of the corresponding l_j in L, i.e., $l_{j(i,k)} = (i,\ \mathbf{a}^i_k : \sigma^i_k)$. where $\mathbf{a}^i_0$ be the attribute values of $\theta^{(0)}_i$.

Estimation of models consists of two parts: estimation of the conditional probabilities in the attributed N-gram model $\mathcal{A}$, and estimation of the attribute update function Δ.

Let $w = \mathbf{a}_1\!:\!\sigma_1, \cdots, \mathbf{a}_{|w|}\!:\!\sigma_{|w|}$ be a sequence of attributed events. The number of times that an attributed k-gram $\mathbf{a}\!:\!y$ appears in w, denoted by $O_{\mathbf{a}:y}(w)$, is defined as the number of elements in the following index set

$$\begin{aligned}\{1 \le j \le |w| - k + 1 \mid \\ \mathbf{a}_{j+k-1} = \mathbf{a},\ \sigma_j\sigma_{j+1}\cdots\sigma_{j+k-1} = y\}\end{aligned} \tag{4}$$

Then the maximal likelihood estimation of the conditional probabilities is given as follows:

$$Pr(\sigma|\mathbf{a}\!:\!y) = \frac{\sum_{i=1,n} O_{\mathbf{a}:y\sigma}(w^i)}{\sum_{\sigma'\in\Sigma}\left(\sum_{i=1,n} O_{\mathbf{a}:y\sigma'}(w^i)\right)} \tag{5}$$

Next we show how the attribute update function Δ is estimated. In fact, the event-log does not give sufficient information on the estimation of function Δ, because attribute values of a process instance is observed only when an event of the instance occurs. Therefore, it is necessary to use a priori knowledge on the attributes updation to identify Δ. For example, some of the attribute values may not change in the log, and some are given only from outside.

The domain of Δ is extended to $V \times \Sigma^*$ by $\Delta(\mathbf{a}, s\sigma) := \Delta(\mathbf{a}', \sigma)$ and $\mathbf{a}' \in \Delta(\mathbf{a}, s)$ for $s \in \Sigma^*$ and $\sigma \in \Sigma$. We aim

[1] Process instances at time less than $N-1$ cannot have $(N-1)$-gram as the event history. We may introduce a "null" event to represent such initial fragment of the event history.

to estimate Δ in the following senses: for any pair $\mathbf{a}_k^i$ and $\mathbf{a}_{k+1}^i$ of attribute values that appear in the event-log, find $\Sigma_{\mathbf{a}_k^i, \mathbf{a}_{k+1}^i} \subseteq \Sigma^*$ such that $\mathbf{a}_{k+1}^i \in \Delta(\mathbf{a}_k^i, s)$ if $s \in \Sigma_{\mathbf{a}_k^i, \mathbf{a}_{k+1}^i}$, where $\mathbf{a}_0^i$ is the attribute values of $\theta_i^{(0)}$.

For a sequence of events s, let $sub(s)$ denotes the set of all sequences other than the empty sequence ε obtained by removing any number of symbols from s, e.g., $sub(abc) = \{a, b, c, ab, ac, b, abc\}$. We define the following:

- $s_k^i := \sigma_{j_{(i,k)}} \sigma_{j_{(i,k)}+1} \cdots \sigma_{j_{(i,k+1)}-1}$ is the sequence of events between $l_{j_{(i,k)}}$ and $l_{j_{(i,k+1)}-1}$ in the event-log L,
- $\underline{\Sigma}_{\mathbf{a},\mathbf{a}'} := \bigcup_{i=1,n} \bigcup_{k:\mathbf{a}_k^i=\mathbf{a}, \mathbf{a}_{k+1}^i=\mathbf{a}'} \{s_k^i\}$,
- $\overline{\Sigma}_{\mathbf{a},\mathbf{a}'} := \bigcup_{i=1,n} \bigcup_{k:\mathbf{a}_k^i=\mathbf{a}, \mathbf{a}_{k+1}^i=\mathbf{a}'} sub(s_k^i)$.

Then the set $\Sigma_{\mathbf{a},\mathbf{a}'}$ satisfies

$$\underline{\Sigma}_{\mathbf{a},\mathbf{a}'} \subseteq \Sigma_{\mathbf{a},\mathbf{a}'} \subseteq \overline{\Sigma}_{\mathbf{a},\mathbf{a}'} \tag{6}$$

Given an event-log L and a set of initial process instance $\Theta^{(0)}$, let $\mathcal{A}_L$ be the attributed N-gram model obtained by the method described above, and let Δ_L be any attribute update function that is consistent with $\Sigma_{\mathbf{a},\mathbf{a}'}$'s satisfying (6). By the construction of $\mathcal{A}_L$ and Δ_L, we have the following theorem.

Theorem 1. The event sequence $\sigma_1\sigma_2 \cdots \sigma_{|L|}$ extracted from L is a sample path of the communicating N-gram models $\mathcal{C}_L = (\mathcal{A}_L, \Theta^{(0)}, \Delta_L)$.

5. DETECTION OF UNUSUAL ACTIVITIES

In this section, two analysis methods for event-logs are presented. The first method focuses on the entire behavior of each process instance, the second method focuses on fragments of activities.

5.1 Analysis of Entire Behavior

Let $w = \mathbf{a}_1 : \sigma_1, \cdots, \mathbf{a}_{|w|} : \sigma_{|w|}$ be a sequence of attributed symbols. The symbol part of w is denoted by $s\langle w \rangle := \sigma_1\sigma_2 \cdot \sigma_{|w|}$, and the last attribute value is denoted by $a\langle w \rangle := \mathbf{a}_{|w|}$. We can evaluate the difference between w and a given attributed N-gram model by *the cross entropy* defined by

$$\mathcal{H}(w) := -\sum_{j=1,|w|} \frac{1}{|w|} log_2 Pr\left(s\langle w_{[j]}\rangle \mid a\langle w_{[j-N+1,j-1]}\rangle : s\langle w_{[j-N+1,j-1]}\rangle\right) \tag{7}$$

$\mathcal{H}(w)$ becomes smaller if the probability distribution of w is closer to that of the attributed N-gram model.

By comparing $\mathcal{H}(w^i)$ for event-logs $w^i, i = 1, \cdots, n$, we can identify processes that behave differently from other processes.

5.2 Detection of Unusual Activity Patterns

The second method is for detecting unusual activity patterns in event-logs. This problem is classified as anomaly detection on time series (Chandola (2009)).

Let s be a short event sequence of a fixed length. If s corresponds to unusual activities and occurs in an event-log, then the number of times s occurs in the event-log is different from its expected value computed by the behavior model. This is the main idea of the proposed method.

Based on the attributed N-gram model, the conditional probability $Pr(s|\mathbf{a}:y)$ that r-gram s appears after $\mathbf{a}:y$ is given by

$$Pr(s|\mathbf{a}:y) := Pr(s|\mathbf{a}:y_{[j,|y|]}) \tag{8}$$

where $j = max\{|y| - N + 2, 1\}$, i.e., we take only the last $(N-1)$-gram into account. Moreover, the right hand side is recursively computed by

$$Pr(s|\mathbf{a}:y) := Pr(s_{[1]}|\mathbf{a}:y) \cdot Pr(s_{[2,*]}|\mathbf{a}:ys_{[1]}) \tag{9}$$

Let $w = \mathbf{a}_1 : \sigma_1, \mathbf{a}_2 : \sigma_2, \cdots, \mathbf{a}_{|w|} : \sigma_{|w|}$ be a sequence of attributed symbols. The expected number of times r-gram s occurs in w is approximated by:

$$E_s(w) := \sum_{\mathbf{a}:y \in \Sigma_{attr}^{N-1}} O_{\mathbf{a}:y}(w) \cdot Pr(s|\mathbf{a}:y) \tag{10}$$

Now we define *the specificity* of r-gram s in w by the following log ratio:

$$d_s(w) := log \frac{O_s(s\langle w \rangle)}{E_s(w)} \tag{11}$$

When $E_s(w) = 0$, $O_s(s\langle w \rangle)$ is also 0 and $d_s(w)$ is defined to be 0. This quantity was originally introduced by the authors for diagnosis of discrete-event systems (Hiraishi (2013a)), but its usage was different. If $d_s(w)$ is larger than 0 (smaller than 0), then s may correspond to some unusual activities. We note that $d_s(w)$ is a dimensionless quantity and is independent of the length of w.

6. APPLICATIONS TO REAL DATA

6.1 The SVM System and Field Experiments

The system developed in the project is called the smart voice messaging system (SVM system). The SVM system consists of smartphones with an application software (SVM terminals), server PC's on cloud, and Bluetooth markers located in fields. Once each staff speaks a short sentence to the terminal, then the message is sent to the server as a voice message, and is distributed to other staffs. Simultaneously, the voice message is recognized and transformed into text data. Important keywords representing the situation are also extracted (Fig. 2). The messages should be sent to only staffs who need the information. In order to realize such smart message distribution, various kinds of information are used, such as location data measured by Bluetooth markers, acceleration sensor data that is used for estimating activity, and the keywords extracted from the voice messages.

The SVM system is tested in a nursing home several times. The situation in the experiments is described as follows:

- Field: In a nursing home with three floors, there are patients' rooms, living salons and other rooms such as a staff station and treatment rooms.

- Roles of staffs: In each period of a day, there are around 8 staffs in the field. Each staff has his/her own role, e.g., the in-charge nurse (commander), staffs responsible for 1F/2F/3F, staffs capable of nursing, etc. In the experiments, all staffs carry SVM terminals together with standard equipment.
- Workflow: The experiments was done at lunch time and dinner time. At first each staff takes a patient from his/her room to a salon, assists the patient to have a meal, cares for several things after the meal (brushing teeth, toilet, give medicine, entertainment events, etc), and finally takes the patient back.

Staffs behave independently, but on some occasions one staff may help other staffs. In such collaboration, awareness of other staffs is important and we expect the SVM system contributes to knowing other staffs' situation.

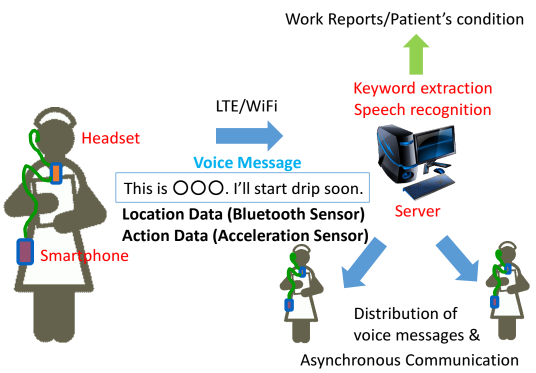

Fig. 2. The smart voice messaging system.

6.2 Situation Mode

N-gram models is used for representing behavior of individual staff. Since there are multiple staffs in the field, behavior of each staff may be influenced by other staffs. To represent the current status of staffs, we introduce an attribute, called *the situation mode*, that represents arrangement of staffs in the field.

Assume that the working time is partitioned into several periods, and let p_i denote the i-th period. For the j-th staff and the k-th location loc_k, let $g_{j,k}^{(i)}$ denote the accumulated length of time that the j-th staff spent in location loc_k during period p_i, and let $G_k^{(i)} := \sum_j g_{j,k}^{(i)}$. If there exists a location loc_h such that $G_h^{(i)} / \sum_k G_k^{(i)} > \theta$, where $\theta \geq 0.5$ is a threshold value, then the situation mode $a_{sm}^{(i)}$ in period p_i is defined as $a_{sm}^{(i)} := loc_h$, and $a_{sm}^{(i)} := \emptyset$ otherwise.

The motivation to introduce the situation mode is as follows. In the lunch/dinner time, most of the staffs are in salons and the situation mode is '1F salon' or '2F salon'. In these situations, if some staff is in a patient's room, then the staff may do some unusual tasks. However, working in a patient room is a normal task before/after the meal.

6.3 Event-logs

Each SVM terminal receives signals from the Bluetooth markers and sends raw data to the server. After processing the data, the result is recorded in the following form:

$$(date, \mathit{Staff\text{-}ID}, type, \mathit{in\text{-}time}, \mathit{out\text{-}time}, duration) \quad (12)$$

where *type* is either the location or "moving". Events with type "moving" means that no location data is obtained during the period and its duration is calculated.

To estimate models in the form of communicating N-gram models, we first replace each event by a symbol. We call this step *event abstraction*. On one hand, the same symbol may be assigned to different events if we do not have to distinguish them. For example, if we focus on movement between floors, the same symbol is assigned to movement on the same floor. On the other hand, in order to distinguish duration of tasks, we may assign different symbols to the same task according to its duration.

Field experiments were done at a nursing home on May 20-24, 2013. The total number of recorded events is 5,420. Event-logs are analyzed by the proposed approach under the following conditions:

- Event symbols are a (1F patients' rooms), b (2F patients' rooms), c (3F patients' rooms), x (1F salon), s (2F staff-station), y (2F salon), z (3F care-station), and e (stairs/ landing/elevator). Capital letters are used for long stays, e.g., we use 'A' instead of 'a'. In addition, M is assigned to long movements. The threshold for determining long stays/movements is set to 120 seconds.
- We use three attributes: the situation mode, the role (mainly work on 1F or 2F), and whether the staff is the in-charge nurse or not. Roles and the in-charge nurse do not change during each experiment. The working time is partitioned into periods of 10 minutes, and the situation mode is determined for each period. The threshold for computing the situation mode is 0.5. Change of the situation mode is realized by an input event from outside.
- We choose $N = r = 4$.

6.4 Analysis Results

We first describe separation of working time by the situation mode. Fig. 3 is the situation mode for event-logs at dinner time on May 25. It is observed that caring at the dinner time is roughly separated into three periods: dining time (most of the staffs are in salons), caring after dinner (staffs take patients to treatment rooms sequentially), caring in the patient's room (staffs take patients back to rooms).

Next, we show how cross entropy is used for finding staffs whose behavior is different from the average one. Fig. 4 is the resulting histogram for the cross entropy of each staff. We can identify two staffs with higher values than others. These staffs spent most of time in 1F salon. Such behavior is different form the average one.

Finally, we try to identify unusual activities in event-logs. Fig. 5 indicates specificity of the 4-gram that begins at each point in time. There are several points in time at which the specificity is high. We pick up four points indicated in the graph. The actual behavior at those points is as follows:

A. Frequent movements between different floors: 1F salon, 2F salon, 3F room, and 2F room.

B. Long movement and long task at the same location.
C. The following voice message was sent just before the point: "Ms. XXX has returned to her room by herself. I will go to see her now." (Usually Ms. XXX needs assistance on her movement.)
D. Long stay at the elevator hall.

We do not expect automatic detection of unusual behavior. Such points in time with high specificity values should be looked back in conferences by the staffs.

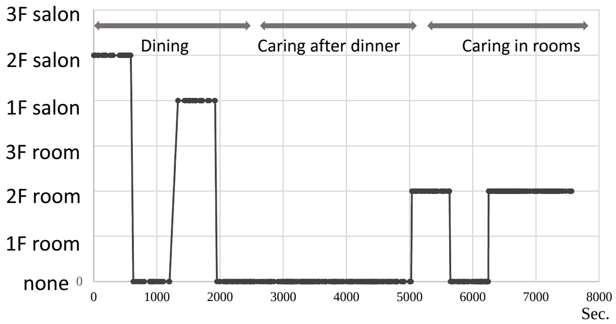

Fig. 3. Change of the situation mode.

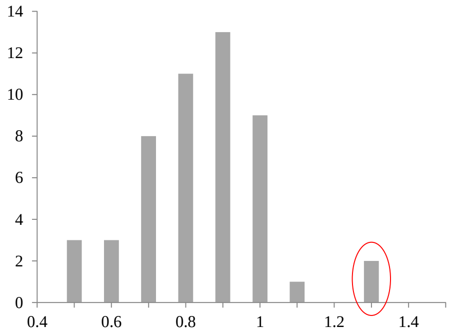

Fig. 4. Distribution of cross entropy.

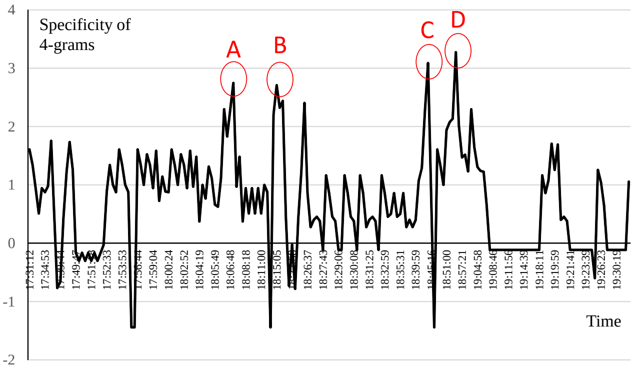

Fig. 5. Detection of unusual activities.

7. CONCLUSION

We have proposed formalism, called communicating N-gram models, for modeling human activities in physical and adaptive intelligent services. We have applies the proposing formalism to building behavior models of staffs in a nursing home. The proposed formalism represents both independent actions of each staff and mutual communication between staffs.

Communicating N-gram models cannot be used for discrete-event simulation, since timing information on event occurrence, such as probabilistic distribution of inter-event time, is not included. However, event-logs in (12) has such timing information. Incorporating such timing information with the models, we will have models for simulation. This remains as future work.

ACKNOWLEDGEMENTS

This research is supported by S3FIRE program at JST/RISTEX, and is partly supported by Grant-in-Aid for Scientific Research of MEXT under Grant No. 25330011.

REFERENCES

Uchihira, N., Choe, S., Hiraishi, K., Torii, K., Chino, T., Hirabayashi, Y., Sugihara, T., Collaboration Management by Smart Voice Messaging for Physical and Adaptive Intelligent Services, PICMET 2013, 2013.

van der Aalst, W. M. P., Process Mining, Springer, 2011.

Ghahramani, Z., Factorial hidden Markov models, Machine Learning, Vol. 29, No. 2-3, pp. 245–273, 1997.

Landwehr, N., Modeling interleaved hidden processes, 25th International Conference on Machine Learning, pp. 520–527, 2008.

Angluin, D., M. Csürős, M., Learning Markov chains with variable memory length from noisy output, COLT 97, pp. 298–308, 1997.

Sen, K., Viswanathan, M., Agh, G., Learning continuous time Markov chains from sample executions, First International Conference on the Quantitative Evaluation of Systems, pp. 146–155, 2004.

Avrunin, G. S., et al., Experience modeling and analyzing medical processes: UMass/Baystate medical safety project overview, 1st ACM International Health Informatics Symposium, pp. 316–325, Arlington, VA, November 2010.

Thorwarth, M.,Arisha, A., Application of discrete-event simulation in health care: a review, Reports. Paper 3. http://arrow.dit.ie/buschmanrep/3

Sundramoorthi, D., et al., A data-integrated simulation model to evaluate nurse-patient assignments, Health care Management Science, Vol. 12, No. 3, pp. 252–268, 2007.

Chandola, V., Banerjee, A., Kumar, V., Anomaly detection : A survey, ACM Computing Surveys, Vol. 14, Issue 3, Article No. 15, 2009.

Shannon, C. E., A mathematical theory of communication, Bell System Tech. J., Vol. 27, pp. 379–423, 623–656, 1948.

Chen, S. F., Goodman, J., An empirical study of smoothing techniques for language modeling, Proceedings of the 34th annual meeting on Association for Computational Linguistics, pp. 310–318, 1996.

Hiraishi, K., Kobayashi, K., Choe, S., Uchihira, N., Behavior modeling in physical and adaptive intelligent services, Proc. CogSIMA2014, pp. 221–226, 2014.

Hiraishi, K., Yoshimoto, M., Kobayashi, K., Diagnosis of stochastic discrete event systems based on N-gram models with wildcard characters, Proc. the 6th IFIP/IEEE Distributed Autonomous Network Management Systems (DANMS 2013), pp. 1383–1389, 2013.