Vol.2016-UBI-51 No.3
Vol.2016-ASD-5 No.3
2016/8/4

近接センサと医療データを用いた看護師と患者の近未来予測と効率化について

井上 創造[1,a)]　磯田 達也[1]　白水 麻子[3]　杉山 康彦[4]　野原 康伸[2]　中島 直樹[2]

概要：
本研究では，病院内の整形外科病棟フロア全体において，40 日間の昼夜にわたり，看護師の行動記録および近接センサによる位置情報，そして DPC および看護必要度という医療データを収集した．収集したデータに対して，1) ある日の患者の状態から次の日の看護業務の長短を予測出来るか，2) 一日の看護業務量から，患者の入院日数の長短，退院時 ADL の善し悪しを予測できるか，3) それらに影響を与えない看護業務量を減らすことが出来るか，をアンサンブル機械学習による予測および，変数重要度を調べることによって検証する．分析の結果，1) については 73.7%, 2) については入院日数の長短を 67.81%, ADL の改善を 74.77%で予測できた．3) については，アンサンブル学習の結果重要度の低い看護行動を除外しても精度が低下しないことを示し，それらの行動を削減できる可能性を示した．

1. Introduction

In the era of the Internet of Things, in the medical field, it is becoming possible to apply sensors to medical staff or locations to optimize hospital operations. Several technical trials of sensing and recognizing contexts or activities in hospitals exist[7], [8], [14]. However, further study is needed, because none of these trials have proven that these kinds of sensing lead to optimization of hospital operations and maintenance, or improvements to the quality of medical care for patients.

In this paper, we carried out an experiment to collect nurse activity and location data in cooperation with one floor of a hospital, which constitutes the orthopedic surgery department, for 40 days, 24 hours per day. In the experiment, we selected 35 nurses as subjects and asked them to carry out their daily duties while wearing a name-card-type sensor that identifies infrared beacons attached to each bed and to several areas of the floor to identify locations or nearby patients. Additionally, we asked the subjects to carry a mobile device and to input their activities, and collected nursing activity records. In addition, we collected medical data such as DPC data (for calculating medical payment), patients status data (such as nursing needs), and number of hospitalization days.

Based on the above-mentioned objectives, the goal of the analysis is to answer the following questions: 1) From which variables of patient status can we predict the next day's nursing load? 2) From which nursing activity times (of each day) can we predict the patient status such as hospitalization duration or activity in daily living (ADL) improvement at discharge?

To answer these questions, we applied RandomForest supervised machine learning to the dataset and investigated the importance of variables. As a result, for 1), the accuracy of predicting whether the next day's nursing time for a patient will be long based on the previous day's patient status was 73.7%, and we could obtain several important variables related to patient status with positive, negative, or more complex correlations. For 2), the accuracy of predicting whether the number of hospitalization days was high or not based on nursing activities was 67.81%. The prediction accuracy of whether the ADLs at discharge had improved or not was 74.77%. Moreover, we identified unimportant nursing activities which, when omitted, affected the reduction in the prediction accuracy very little. Using these result, we could demonstrate that we can make a daily forecast report of nursing workload for each patient and each activity type. We could also set an assumption for several types of activities to be simplified without decreas-

1 九州工業大学
〒 804-8580, 北九州市戸畑区仙水町 1-1
2 九州大学病院
〒 812-8552, 福岡市東区馬出 3-1-1
3 熊本県立大学
〒 862-8502 熊本県熊本市東区月出 3 丁目 1 － 100
4 株式会社シーイー・フォックス
a) sozo@mns.kyutech.ac.jp

ing the medical service quality, and thus, making the hospital operation more effective.

The contribution of our paper is two-fold:

(1) we demonstrate that we can predict and optimize the near future performance of patiences and nurses, and,

(2) we show the utilization of ensemble machine learning to determine the importance of variables and to select or omit several variables for understanding causes and results.

2. Experiment

In this chapter, we describe the method of collecting proximity sensor data and nurse activity data, and present an overview of the collected data.

2.1 Experiment to collect nursing data

In this section, we describe the data collection experiments performed in the hospital. We carried out an experiment to collect nurse activity data in cooperation with one floor of a hospital, which constitutes the orthopedic surgery department. The experiment was carried out for 40 days, from May 25th to July 3rd, 2015, 24 hours per day. The experiment was exclusive to nurses who agreed to the use of sensor data and to duties related to patients who consented to participate in the experiment. Table 1 shows an overview of the conditions of the experiment.

表 1 Conditions of the experiment

Experimental environment	Orthopedic surgery ward
Period of experiment	May 25th - July 3rd, 2015
Number of days of the experiment	40 days
Experiment hours per day	24 hours
Subjects	35 nurses (1 head nurse) and 118 patients
Infrared beacons:	
number	251
groups	61
frequency	1 sample per min.
Mobile terminal:	
activity types	115
grouped as	25

2.1.1 Sensor data

We divided one floor of the hospital ward into 61 blocks and installed information and communication equipment that uses infrared for sensing (hereafter, *infrared beacons*[*1]) in each block, including patient rooms, the staff station, and other rooms important for nursing care. In total, 251 infrared beacons were installed, in the locations shown in Figure 1. These infrared beacons were used for obtaining the location information of nurses, such as a given nurse being near a particular patient's bed or near a particular piece of equipment. For each bed, three infrared beacons are installed, as shown in the red-circled locations in the figure.

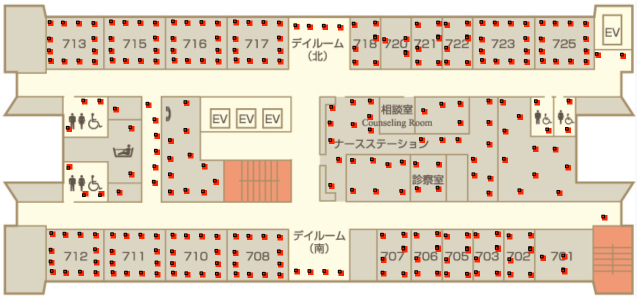

図 1 Floor map used in the experiment

We used 35 nurses as subjects of the experiment. We asked these nurses to attach a name-card-type sensor equipped with an infrared emitter (hereafter, *name-card sensor*[*1]), as shown in Figure 2. The name-card sensors were worn by each subject around the neck.

図 2 A nurse attaching a name card sensor

Using the infrared beacons and the name card sensors, location information from the subject can be collected by the communication of the becons' IDs to the name-card sensors.

These sensor data are stored by the name-card sensors, and collected by a stationary PC using a USB cable when the duties of each nurse are over and s/he puts the sensor in the charging stand.

We also collected activity information for nurses by asking them to input their activity records using smartphone devices[*2]. The activity types of the nurses were predefined, and were recorded by the subjects themselves by selecting from a list in an existing activity logging app[*3]. The duties of the nurses included 115 items, which were grouped into 25 categories.

*1 Hitachi High Technology's Business Microscope

*2 iPod touches from Apple, Inc.

*3 aTimeLogger

2.1.2 Medical payment data

In addition to the sensor data, we also collected data called diagnosis procedure combination data (hereafter, *DPC*)[11], which is a payment data and calculation system for acute hospital and applied for acute and chronic diseases that was introduced in Japan in 2003. The system is derived from the Diagnosis Related Groups and Prospective Payment System (DRG/PPS), which was developed by Yale University. DPC is refined as the 'Combination' of the payment for constant per-'Diagnosis' basis such as costs for supplies used in wards and examination rooms, and the variable per-'Procedure' basis such as costs in operation rooms or procedures.

From the viewpoint of hospital operations, by introducing DPC, it becomes more cost effective for the hospital to reduce the hospitalization duration of a patient (likewise for the patient), because the per-service component (which is often the major part of the payment), is the same, no matter how long the duration of the stay. Furthermore, the hospital can admit another patient for the vacant bed.

Moreover, as a secondary use of DPC, a uniform data analysis between different hospitals after anonymized becomes possible, which is fascinating for big data analytics.

Among the many fields in the DPC data, we utilized the information of per-procedure part (located in the EF file in the DPC files), which are summaries of the procedure and correlated to the actual payment, which will be covered by the self payment and by an insurance, as 1 *point* equals to 10 yen.

2.1.3 Nursing needs data

From 2006, a uniform nursing diagnosis system, called *Nursing needs*, was introduced in Japan. The diagnosis is composed of several categories such as vital status, applied nursing cares, or necessity for assistances. The system is required to calculate the medical payment appropriately in the in the designated hospitals increasing these years. Internationally, it is compatible with the standard of NANDA[6] to some extent.

Regarding the data introduced above, the data's structure is extracted, separated by data keyed by patient and data pairs keyed by patient and date.

Data keyed by patient include:

- Patient ID
- Birth year and date (age)
- Sex
- Date of hospitalization
- Date of operation
- Date of discharge
- Level of ADL at hospitalization
- Level of ADL at discharge
- Medical payment derived from DPC data

ADL is composed of several attributes; we aggregated and leveled from 0 to 20, where less is better.

Data keyed by pairs of patient and date:

- Patient ID
- Date
- Level of nursing needs

2.2 Results of data collection

The data we collected are structured as shown in Figure 3.

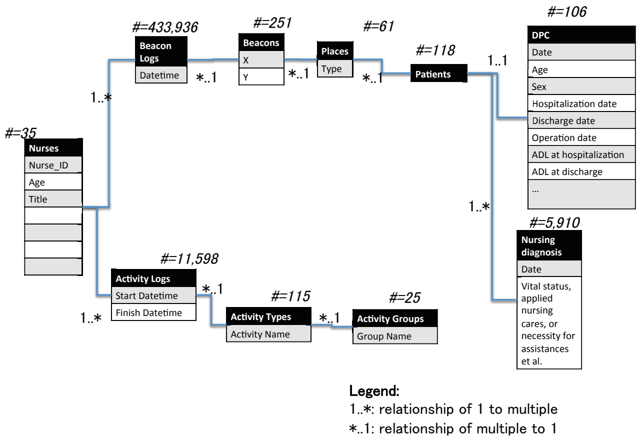

図 **3** Relational organization of the data.

As a result, we obtained 347 [days × people] of sensor data, 12,406 activity labels with 111 activity types, and the medical data of nursing needs from 118 patients, where 96 of them also provided the DPC data.

Then, we converted the daily location information of nurses into a graph. Figure 4 shows the positional information of each nurse over one day. The vertical axis shows the position of

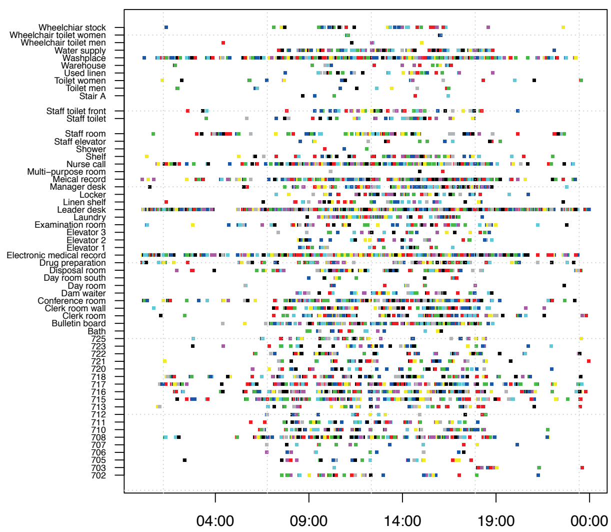

図 **4** Location information of nurses over one day.

nurses and the horizontal axis shows the time of day; the plots

Vol.2016-UBI-51 No.3
Vol.2016-ASD-5 No.3
2016/8/4

are colored by nurse. Figure 4 also shows that there is a difference in the number and position of nurses between day-work nurses and night-work nurses. Moreover, we can see that, in almost all time zones of day, nurses frequently visit the place for electronic medical record, the leader's desk, and the wash place. Especially for night-work, the probability of nurses staying at the nurse station is high. This result was a characteristic feature that was seen for all experimental days.

3. Analysis

In this section, we describe the data analysis using data obtained in the experiment for predicting the future state of patients and nurses, and discuss the improvements in nursing duties.

3.1 Objective

The primary goal of a hospital is to make a patient better at discharge than at hospitalization time. Moreover, it is better for the patient for recovery to take fewer days (it is also better for the hospital in terms of cost).

To make the above possible, optimization of the duty allocation of nurses to patients or specific tasks (and balancing workloads) is important. Optimizing the allocation of nurses leads to effective utilization of human resources; this will contribute to maintained hospital services in countries where nurse shortages are expected because of an increasingly elderly population in many advanced countries.

Based on the above-mentioned objectives, the goal of the analysis is to answer the following questions:

Question 1 : From which variables of patient status can we predict the next day's nursing load?

Question 2 : From which nursing activity times (of each day) can we predict the patient status such as hospitalization duration or ADL improvement at discharge?

3.2 Strategy

In the following analysis, we develop a common strategy to:

(1) set the target we want to predict as an objective variable y,

(2) set the explanatory variable vectors $\vec{x}$ using the data we want to use as input or see the correlations,

(3) apply the RandomForest algorithm to evaluate the prediction accuracy, and,

(4) at the same time, compare the importance of each explanatory variable, and identify effective or ignorable in the prediction.

For example, we set $\vec{x}$ as patient status and y as the next day's nursing time in Question 1, and $\vec{x}$ as the nursing activity time of each day and y as the number of hospitalization days in Question 2.

	Linear Regression	Decision Tree	Random Forest
Importance of each variable	△	△	✔
Statistical test	✔	△	△
Visualization	△	✔	✖
Generality against Overfitting	△	△	✔

図 5 Comparison of RandomForest regression, traditional linear regression, and decision tree.

When we use ensemble learning such as RandomForest, there are several advantages over traditional regression or machine-learning methods, as summarized in Figure 5. With ensemble learning, we can see the *importance* (the sum of accuracy improvement when using the variable) of variables after neutralizing interactions among variables (unlike a traditional regression, in which adopting one variable affects the importance of the remainder via a so-called *interaction*). By neutralizing interactions, we can see the effect of each variable on the response variable in *a partial dependent plot*.

We fully utilize these advantages to select explanatory variables that have a high impact on the objective variable, such as selecting the patient status for predicting the nursing load in the near future in Question 1. Conversely, we can also omit unimportant explanatory variables, which do not affect the prediction accuracy, such as detecting the nursing activities that do not affect the hospitalization durations of patients, and try to reduce tasks or alternate with non-professionals in Question 2.

We can take advantage of RandomForest, which automatically avoids overfitting and outputs general models. If we select a tree from the set of obtained trees, we can easily understand the partitioning conditions, unlike other algorithms such as support vector machines (SVM).

We believe that this strategy can be applied to many fields of ubicomp, such as skill assessment and human activity analysis (which has specific goals, such as in healthcare, education, and business activities).

3.3 Preprocessing

As a preprocessing step, we first inner joined the relational data illustrated in Figure 3 for each minute of each nurse's duties, in the following two ways:

- Nurses → Activity Logs → Activity Types, and
- Nurses → Beacon Logs → Beacons → Places → Patients → DPC

Again, we inner joined these process using the timestamps of activity logs and beacon logs.

Using this method, we can obtain the fully joined data, and we can determine

- what activity a nurse performed at which time, and for which patient,
- the nurse's profile,
- patient attributes such as hospitalization duration, and ADL at hospitalization or discharge, and
- each patient's status or nursing needs for a given day.

The beacon logs table is important, because otherwise, we do not know which patient is associated with the nurse activity. Using infrared beacons, we were able to join nursing activities and patient information and utilize these data in the overall analysis.

3.4 Procedure

Question 1: Nursing load prediction

To predict whether the next day's nursing time based on a patient's status on a given day, we set the variables as follows and applied supervised machine learning.

Explanatory variables:

- The attributes of the patient such as age, room.
- Nursing needs levels
- Number of days after hospitalization
- Number of days after an operation

Objective variables:

- Whether the total nursing time for the patient is longer or shorter than the median value.

We employed the RandomForest algorithm for machine learning and validated the result using one-patient-left-out cross validation to determine the predictability for a new patient.

Question 2: Prediction of patient status

Regarding patients, we focused on two important aspects for patients, i.e., the number of hospitalization days and whether the ADL is recovered at the discharge time relative to the hospitalization time, and set the following variables:

Explanatory variables:

- Nursing times for a patient for each activity type. Note that these are multivariate, each of which corresponds to an activity type.

Objective variables:

- Whether the hospitalization duration is longer than the median 29 days, or
- Whether the ADL score of the patient is less than the median 11 (which is poor) is improved by more than 1.

As in Question 1, we used RandomForest and applied one-patient-left-out cross validation.

We also investigated the importance of each explanatory variable; however, we focused on those with the lowest importance, because if we know that an activity type has low importance, this type will not affect (nor be correlated with) the patient's future statuses of objective values. Therefore, we can consider making the activity simpler, making the activity shorter, or eliminating it completely.

3.5 Results

Result 1: Nursing load prediction

The accuracy of predicting whether the next day's nursing time for a patient is long was 73.7%.

We show the result of the prediction and ground truths for each patient and day in Figure 6. In this, the horizontal axis is days after hospitalization, and the vertical axis is patients (sorted by age). Because we only have records for experimental days, and we aligned our days with the hospitalization day, the hospitalization days are different among patients; there are blanks for first or last days for certain patients.

In the graph, the light blue color (blue, red, orange) denotes true negatives (false negatives, false positives, true positives, respectively). The figure shows that the first days tend to be longer and latter days be shorter; however, in both days, the predictions are relatively successful for the most patients.

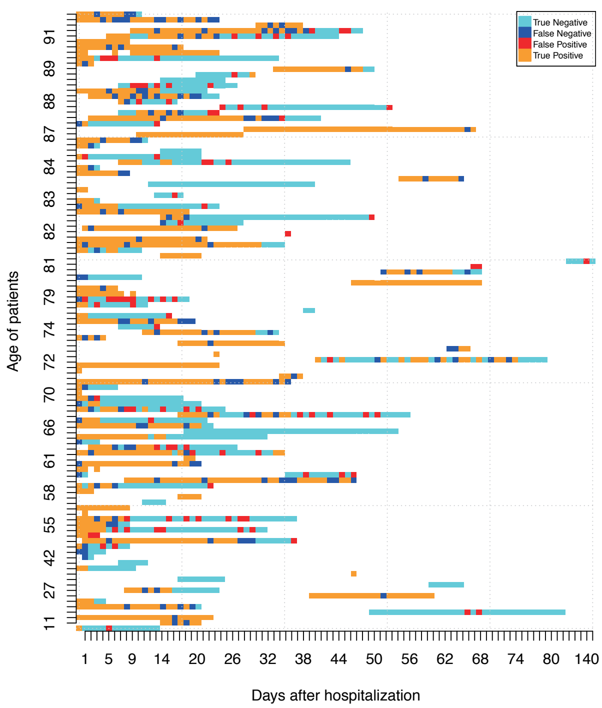

図 6　Prediction of whether the nursing activities are long based on the previous day's patient status

Vol.2016-UBI-51 No.3
Vol.2016-ASD-5 No.3
2016/8/4

Result 2: Prediction of patient status

The accuracy of predicting whether the number of hospitalization days was high or not was 67.81%.

We show the scatter plot between nursing time and rate of variable importance compared to the nursing time for each activity type in Figure 7. From the figure, activity such as “Change diaper”, “Measure blood pressure”, “Measure breath”, “Examine condition”, “Assist meal”, “Clean body”, “Record EMR”, “Assist urine excretion”, “Serve/clear meal”, “Prepare medicine”, and “Measure pulse rate” have low importance compared to the activity times.

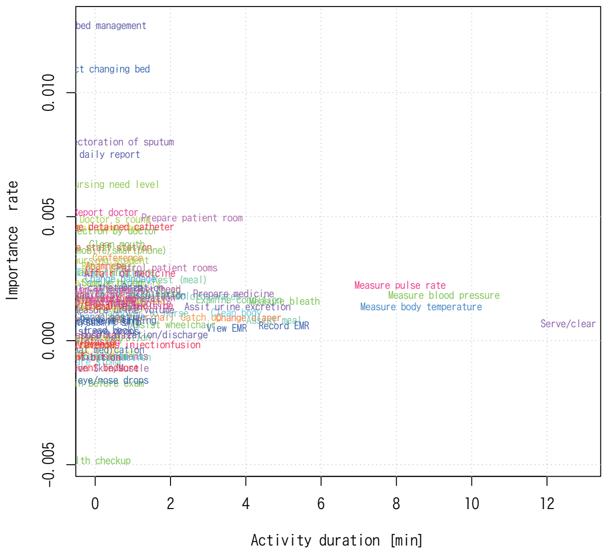

図 7 Scatter plot between nursing time and the variable importance for each activity type for predicting hospitalization days

When we omited these activities and trained and tested a new model, we obtained 67.55% accuracy, which was only 0.26% decline.

Moreover, the prediction accuracy of whether the ADLs at discharge are improved or not was 74.77%. We show the scatter plot between the nursing time and the rate of variable importance in Figure 8.

Also in this figure, the same activities as in Figure 7 have low importance.

When we omit these activities and trained and tested a new model, we obtained 72.31% accuracy, which was a 3.46% decline, which is larger than that presented above, but is still relatively small.

From the results above, we see that the nursing activities omitted here have little effect for predicting hospitalization days and ADL improvement. We can make an assumption that these activities should be simplified as much as possible.

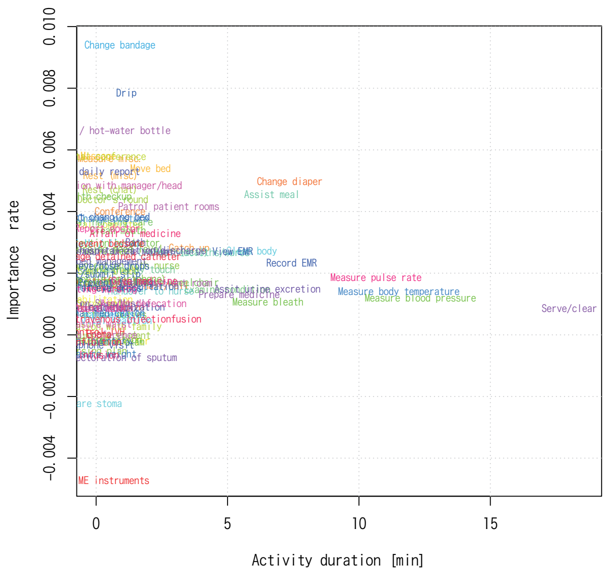

図 8 Scatter plot between nursing time and variable importance for each activity type for predicting ADL improvement

4. Discussion

4.1 Nurses load prediction

Even if the prediction accuracy is not perfect, it is valuable to be able to estimate the next day’s nurse workload for each patient, because we can produce a work load forecast (similar to that demonstrated in Figure 9) sent to the head nurse. The figure shows the forecast result for each patient and each nursing activity type based on whether they will be longer or shorter than average days.

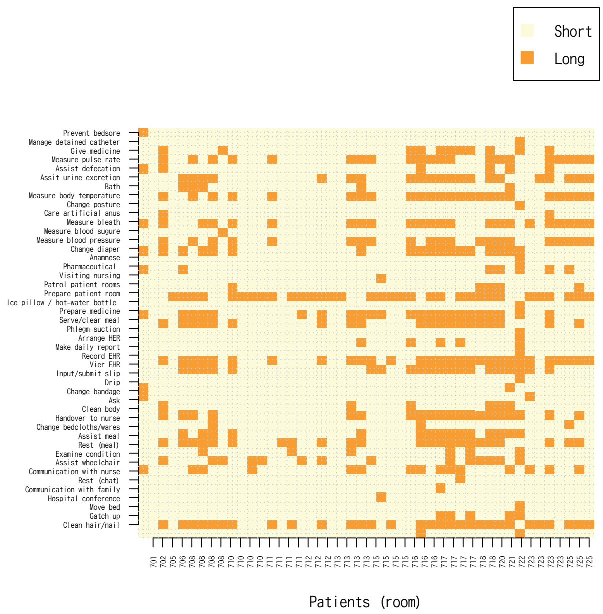

図 9 Example of daily workload forecast report sent to the head nurse

Because this forecast goes into detail for each patient and each activity, it will make it easier to consider the allocation of nurses and equipment than only coarse grained prediction such as the total duration for given patients.

4.2 Prediction of patient status

Based on a discussion with the subject nurses, most of the nursing activities that were found to be unimportant for predicting patient status were agreed to be less important to patient status by nurses. Some of the activities such as “Assist meal”, “Change posture”, and “Serve/clear meal” are not tasks that professional nurses are required to do. Therefore, after this experiment, the hospital started to delegate these activities to other staff members who are not professional nurses. This result implies that these analysis results can influence work reflection and assignment throughout the whole hospital.

In future work, we can analyze further assumptions using this dataset. If we include inquiry data from patients, we can analyze which instances of nursing care are satisfactory to patients. We only considered the efficiency of nursing and hospital processes, but the result may not be ideal if the patient feels uncomfortable. Therefore, this kind of user study is also required. We can also analyze which hospital, facility, or home the patients were transferred to after hospitalization. Predicting which functional facility a patient goes to afterwards is also interesting, because, nowadays, cooperation among hospitals and local communities is important. These analyses are left as future work; our open data can contribute to these research directions.

5. Related Work

Recently, in the medical field, many experiments have collected activity data from doctors, nurses, and patients; many studies make use of these collected (big) data for improving the efficiency of duties or for offering the appropriate medical services [1], [12], [15], [19], [21], [22], [23].As an example of a long term activity data collection experiment in the medical field, there is an experiment that was conducted for two years at the circulatory disease center of Saisei-kai Kumamoto Hospital[7], [8], [25]. In this experiment, the authors collected big data from both nurses and patients. They indirectly recorded sensor data from patients using a monitoring electrocardiogram; a wrist sensor as a 3D acceleration sensor; a bed sensor to measure heart rate, breathing, and acceleration; and an electronic clinical path. From the nurses, they collected acceleration data and entrance/exit information into/from patient rooms, and applied activity recognition on the accelerometer data. Unlike these studies, this paper focuses mainly on data analysis of nurse activity information input by mobile devices, location information collected by infrared beacons and name card sensors, and medical DPC collected from the medical information system that is operated during typical hospital operation.

Furthermore, a study of a method to make use of medical treatment data obtained from multiple hospitals in other hospitals was conducted[24]. This study utilized the medical data acquired from multiple hospitals using transfer learning at one hospital to make a model of many diseases and duties from a large amount of data. Our experiment was executed at one hospital, but comparing data with other hospitals is also valuable, not only because the staff can compare the data and mutually improve nursing activities, but also because we can apply transfer learning among data from different hospitals and train general prediction models.

In this paper, we used the RandomForest[2] algorithm for prediction and variable importance analysis. Using RandomForest (in addition to the fact that the algorithm automatically avoids overfitting for prediction, as described in Section 3), we can see the variable importance after neutralizing interactions among explanatory variables, unlike in a traditional regression. Of course, there are several models for dependency analysis among variables, such as covariance structure analysis[17] and Bayesian networks[16]. Unlike these models, our approach using RandomForest has a focus on the importance analysis of variances when we set an objective variable and train a model to predict the objective variable. This approach is valid and accurate in cases in which there are clear goals or objectives, such as optimizing the nursing workload and patient status in hospital operations.

6. Conclusion

In this paper, we carried out an experiment to collect the nurse activity data, location data, and medical data in cooperation with one floor of a hospital, which constitutes the orthopedic surgery department, for 40 days, 24 hours per day. We applied several analysis methods to patients and nurses aspects. Regarding the nursing aspect, we predicted the next day’s nursing time for a patient based on the previous day’s patient status. Regarding the patient aspect, we predicted whether the number of hospitalization days is high or not based on nursing activities, and predicted the ADL improvement at discharge. We found unimportant nursing activities, i.e., those that, if omitted, affected the decline of the prediction accuracy very little. Using these results, we can produce a daily forecast report of nursing workload for each patient and each activity type. Furthermore, we can make the assumption that several types of activities can be simplified without decreasing the medical service quality, and thus, make hospital operations more effective.

To obtain these results, we applied RandomForest supervised machine learning to the dataset and investigated the importance of variables. We demonstrated the utilization of ensemble machine learning to obtain the importance of variables and to select or omit several variables for understanding causes and cost reduction.

The dataset will be open, with the subjects' agreement; therefore, the contribution of research is expected to be extended and further explored in future research.

Acknowledgment

The part of this research are supported by Research and Development of NICT Social Big Data Utilization Method and Fundamental Technologies, and Grant-in-Aid for Scientific Research (B) "Sensor Context Estimation Techniques by Semantic and Physical Layer (Principal Investigator: Sozo Inoue)".

参考文献

[1] Bardram, J. E., and Christensen, H. B. Pervasive computing support for hospitals: An overview of the activity-based computing project. *IEEE Pervasive Computing 6*, 1 (2007), 44–51.

[2] Breiman, L. U. o. C. Random forest. *Machine Learning 45*, 5 (1999), 1–35.

[3] Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S. T., Tr??ster, G., Mill??n, J. D. R., and Roggen, D. The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters 34* (2013), 2033–2042.

[4] Damas, M., Amft, O., Toth, A. M., Baños, O., Pomares, H., and Tóth, M. A. A benchmark dataset to evaluate sensor displacement in activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, vol. 16 (2012), 1026.

[5] Hattori, Y., Inoue, S., and Hirakawa, G. A Large Scale Gathering System for Activity Data with Mobile Sensors. *2011 15th Annual International Symposium on Wearable Computers* (2011), 97–100.

[6] Herdman, T. H. E., and Kamitsuru, S. E. NANDA International nursing diagnoses: definitions and classification 2015-2017, 2014.

[7] Inoue, S., Ueda, N., Nohara, Y., and Nakashima, N. Mobile Activity Recognition for a Whole Day: Recognizing Real Nursing Activities with Big Dataset. In *ACM Int'l Conf. Pervasive and Ubiquitous Computing (Ubicomp)* (Osaka, 2015).

[8] Inoue, S., Ueda, N., Nohara, Y., and Nakashima, N. Understanding Nursing Activities with Long-term Mobile Activity Recognition with Big Dataset. In *The 47th ISCIE International Symposium on Stochastic Systems Theory and Its Applications (SSS)* (Hawaii, 2015), 10 pages.

[9] Kasteren, T. V., Noulas, A., Englebienne, G., Kr, B., van Kasteren, T., Noulas, A., Englebienne, G., and Kröse, B. Accurate Activity Recognition in a Home Setting. In *UbiComp '08 Proceedings of the 10th international conference on Ubiquitous computing* (2008), 1–9.

[10] Kawaguchi, N., Ogawa, N., and Iwasaki, Y. Hasc challenge: gathering large scale human activity corpus for the real-world activity understandings. In *Proceedings of the 2nd Augmented Human International Conference* (2011), 27.

[11] Matsuda, S. [Case mix based payment and DPC from an international perspective]. *Gan to kagaku ryoho. Cancer & chemotherapy 31*, 8 (2004), 1152–7.

[12] Morán, E. B., Tentori, M., González, V. M., Favela, J., and Martínez-Garcia, A. I. Mobility in hospital work: towards a pervasive computing hospital environment. *International journal of electronic healthcare 3*, 1 (2007), 72–89.

[13] Morik, K., Brockhausen, P., and Joachims, T. Combining statistical learning with a knowledge-based approach-a case study in intensive care monitoring. *ICML '99 Proceedings of the Sixteenth International Conference on Machine Learning* (1999), 268–277.

[14] Naya, F., Ohmura, R., Takayanagi, F., Noma, H., and Kogure, K. Workers' Routine Activity Recognition using Body Movements and Location Information. *2006 10th IEEE International Symposium on Wearable Computers* (2006), 105–108.

[15] Orwat, C., Graefe, A., and Faulwasser, T. Towards pervasive computing in health care - a literature review. *BMC medical informatics and decision making 8* (2008), 26.

[16] Pearl, J. Bayesian networks: A model of self-activated memory for evidential reasoning. In *Proceedings of the 7th Conference of the Cognitive Science Society, University of California, Irvine* (1985), 329–334.

[17] Reinartz, W., Haenlein, M., and Henseler, J. An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of research in Marketing 26*, 4 (2009), 332–344.

[18] Roggen, D., Calatroni, A., Rossi, M., Holleczek, T., Forster, K., Troster, G., Lukowicz, P., Bannach, D., Pirkl, G., Ferscha, A., Doppler, J., Holzmann, C., Kurz, M., Holl, G., Chavarriaga, R., Sagha, H., Bayati, H., Creatura, M., and Millan, J. d. R. Collecting complex activity datasets in highly rich networked sensor environments. *Networked Sensing Systems (INSS), 2010 Seventh International Conference on* (2010).

[19] Sánchez, D., Tentori, M., and Favela, J. Activity recognition for the smart hospital. *IEEE Intelligent Systems 23*, 2 (2008), 50–57.

[20] Sow, D., Biem, A., Blount, M., Ebling, M., and Verscheure, O. Body sensor data processing using stream computing. In *Proceedings of the international conference on Multimedia information retrieval* (2010), 449–458.

[21] Tentori, M., and Favela, J. Activity-aware computing for healthcare. *IEEE Pervasive Computing 7*, 2 (2008), 51–57.

[22] Tentori, M., and Favela, J. Monitoring behavioral patterns in hospitals through activity-aware computing. In *2008 Second International Conference on Pervasive Computing Technologies for Healthcare*, IEEE (jan 2008), 173–176.

[23] Tentori, M., and Favela, J. Monitoring behavioral patterns in hospitals through activity-aware computing. In *Proceedings of the 2nd International Conference on Pervasive Computing Technologies for Healthcare 2008, PervasiveHealth* (2008), 173–176.

[24] Wiens, J., Guttag, J., and Horvitz, E. A study in transfer learning: leveraging data from multiple hospitals to enhance hospital-specific predictions. *Journal of the American Medical Informatics Association 21*, 4 (2014), 699–706.

[25] Yasunobu Nohara, Sozo Inoue, Naoki Nakashima, Naonori Ueda, M. K., Nohara, Y., Sozo, I., Nakashima, N., Naonori, U., and Kitsuregawa, M. Large-scale Sensor Dataset in a Hospital. In *International Workshop on Pattern Recognition for Healthcare Analytics* (Tsukuba, Japan, 2012), 4 pages.