

Flexible WSNs Aims Easy Installation With Noise Reduce Method For Elderly People Care

Takanobu Otsuka

*Nagoya Institute of Technology, Gokiso Showa-ku,
Nagoya, Aichi 466-8555, Japan*

E-mail: otsuka.takanobu@nitech.ac.jp

Takayuki Ito

*Nagoya Institute of Technology
E-mail: ito.takayuki@nitech.ac.jp*

Received 19 June 2013

Accepted 25 June 2013

Much research has been conducted on anomaly detection by wireless sensor networks (WSNs). The existing WSNs require specialized knowledge and skills to install the sensors in environments such as houses and buildings. Therefore, we have developed a flexible WSN based on small sensor devices that can be easily installed. The users only need to place these sensors at the locations where they want to sense and to provide some information to the server through a web page. Then, these small sensor devices automatically create wireless networks, start communicating with the central server for logging continuous data, and show anomalies by using inference based on a basic Bayesian Network. However, a serious problem is that a large amount of noise data prevents correct inferences. Therefore, in this paper, we propose a method for reducing noise data based on location sampling of real human movements. Our experimental results show that our method is effective in increasing the inference accuracy for detecting anomaly data.

Keywords: Wireless Sensor Networks, M2M, Bayesian Network, Anomaly Detection

1. Introduction

In recent years, the fastest-growing information terminal devices are represented by the smartphone. Due to the dramatic progress made in compact communication modules and various sensors, smart phones are now sold at prices that make typical consumers feel free to buy them. In addition, a great deal of research has been conducted on anomaly detection by wireless sensor networks (WSNs). The existing WSNs require specialized knowledge and skills to install sensors in environments such as houses and buildings. This technology has also created services that allow users to view and share sensing data acquired by the sensor network, which may be installed around the world. Such services might include, for example, detection of sea

surface temperature and the status of tidal flows. Various sensing products are being steadily developed and implemented in systems to care for the health of senior citizens. However, these products are usually expensive, and their installation by the user is very difficult. In this research, we developed a wireless sensor network system that offers easy installation and easy operation [1]. Our preliminary experiments demonstrate that our system can find anomalous sensing information without any difficult installation procedures. Consequently, these small sensor devices can automatically create wireless networks, start communicating with the central server for logging continuous data, and show anomalies by using inference based on a basic Bayesian Network. A problem arises

when much noise data prevent correct inferences. Therefore, in this paper, we propose a method for reducing noise data based on location sampling of real human movements. Our experimental results show that our method is effective for raising the accuracy of inference in detecting anomaly data. The rest of the paper is organized as follows. Section 2 introduces previous studies and the relative position of our research. Section 3 presents configuration of test demonstration and knowledge. Section 4 describes the configuration of our test demonstration. Section 5 presents an anomaly detection method using support vector machine (SVM). Section 6 proposes a noise-reduction method using a Bayesian network. Section 7 outlines our experimental evaluations. Finally, Section 8 summarizes our paper and describes future work.

2. Related Works

2.1. General product features

Anomaly detection systems have been used widely in the security field. A representative application is a system for detecting an intruder. These systems [2][3] can alert a security company after detecting the intruder, using motion sensors and lead switches on window doors or entrance doors.

One system [4] simply notifies senior citizens two times a day when it is time for them to drink a cup of tea. Other systems [5][6] can analyze behavior patterns using motion sensors and send emergency reports. Another group of systems [7][8][9] applies RFID units and sensors to analyze behavior patterns in order to send notifications of abnormal events as a whistle-blower system. Yet another approach is based on systems [10][11][12][13] using wearable sensors that can monitor temperature with thermometers.

Video cameras have been used in some systems [14][15][16] for detecting abnormal behavior.

The above works raise a variety of problems:

- Generally they are expensive to install.
- They require separate communication cables.
- Their video cameras are viewed as too invasive of one's privacy.

- Wearable sensors increase the cognitive burden on users.
- They require huge computational power to process videos.
- Their main purpose is to alert rather than to watch/see.
- The visibility of their activity logs is not sophisticated.

2.2. Research on anomaly detection

Anomaly detection refers to the problem of finding patterns in data that do not conform to expected behavior [17]. These nonconforming patterns Flexible WSNs aims easy installation with noise reduce method for elderly people care are often referred to as anomalies, outliers, discordant observations, exceptions, aberrations, surprises, peculiarities or contaminants in different application domains [18]. Of these, anomalies and outliers are the two terms used most commonly, sometimes interchangeably, in the context of anomaly detection. Anomaly detection finds extensive use in a wide variety of applications such as fraud detection for credit cards, insurance or health care, intrusion detection for cyber-security, fault detection in safety- critical systems, and military surveillance of enemy activities. Anomaly detection is important because anomalies in data translate to significant and often critical actionable information in a wide variety of application domains. For example, an anomalous traffic pattern in a computer network could mean that a hacked computer is sending out sensitive data to an unauthorized destination [19]. An anomalous MRI image may indicate the presence of malignant tumors [20]. Anomalous readings from a spacecraft sensor could signify a fault in some component of the spacecraft [21]. Over time, a variety of anomaly detection techniques have been developed in several research communities. Many of these techniques have been specifically developed for certain application domains, while others are more generic.

In addition, the importance of data analysis has been increasing in the field of wireless sensor networks [22] [23]. Sensor data collected from various wireless sensors have special features.

3. Preliminary Experimentation about an Elderly Care Support System

3.1. The overview of our system

Our research aims to develop a system for easy operation and easy installation. Because we conducted a cooperative research with the Niihama medical Coop as a preliminary test in Japan. In this research, we could identify the problem in which system developer needs to know in the real world. We conducted a research with the cooperation of the Niihama Medical Co-op in Japan. Niihama Medical Co-op was consistently operates a large clinic, a group home, a day care, and a day service. It is a few medical institutions in Japan which gives home health care and medical care throughout. Generally medical institutions are divided medical services and home health care services. Therefore the burden on user is big the Niihama Medical Co-op. was integrating medical and home health care services. The Japanese government got the medical institutions to decrease the number of beds for senior citizens because of mainly the budget, and switch the policy toward to home health care.

- Usually care workers visit an elderly person once a week. It depends on the policy or situation. However, at least the sensors must work within one week without any on-site maintenances.
- The shower and bath times are most dangerous period where an senior person collapses because of the larger difference on temperatures. It is required to sense these situation carefully and correctly.
- Sensing before entering and after exiting the restroom is required because Japanese restroom usually do not have air conditioner.
- Detecting abnormal actions must be detected.
- It required to sense as many actions as possible by fewer sensors.

We developed a system to satisfy the above requirements as much as possible. We conducted a preliminary experiment with the medical institution.

We selected inexpensive parts and equipments for realize this experiment. They are available at the common retail stores.

- Motion sensor: AT Watch IR Mini
A presence sensor using infrared USB connectivity
- USB Extender: net.USB device server 1
This is extended by converting USB signals into IP packets up to 50 m product
- WiFi Router
A wireless LAN router with net.USB Device
- 3G Communicate device:
A portable 3G Wifi with the ethernet terminal
- Client PC: Asus EeePC X01H
This is one of the inexpensive PCs.

Laying power and LAN cables allows wireless communication between the sensor and the client PC. The connection to the client PC and the wireless connection must be hidden for dementia patients who cannot easily access the client PC. The maximum length of the USB cable is 5 m with the device net.USB server. That's a net of a sensor number of minutes who are using the USB server device. Our prototype system configuration diagram is shown in Figure 1.

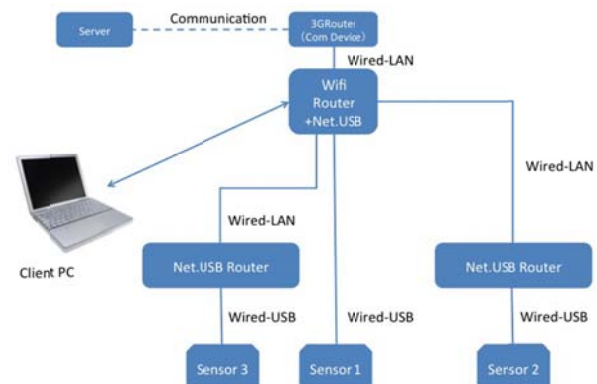


Figure 1: System configuration diagram

3.2. Actual installation example

This section indicates an installation example in the home of a senior citizen living alone. Except for special

circumstances, the sensors are installed in the following places:

- Entrance: this is for sensing people entering or leaving
- Bathroom: this is for accurate sensing of the temperature difference between the bath and the living room
- Lavatory: For sensing the number of bowel movements and safe return from using the toilet

An example of a standard installation is shown in Figure2.



Figure 2: Standard installation

We conducted field surveys to determine the length of the ethernet and power cables beforehand. Residents with disability without ethernet cable are laying the "kamoi" (under the ceiling) and at some high positions. We determined the sensor position in advance by measuring the cable length. For example, After surveying the cable lengths, we determined the cable length as 150 m and installed the cables into the four homes. Also because the net.USB device server that supplies voltage to the USB device server has DC12V output rather than an AC power adapter, we made a DC plug-in for power lines by soldering the cable length.

3.3. Experimental data by the test demonstration

The follows are the experimental data:

- Installation: four homes

- Data collection period: 2012/03/01 to 2012/03/31

The obtained data are shown in Fig. 3.

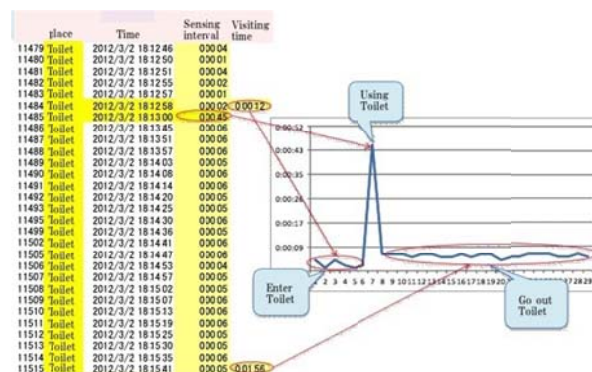


Figure 3: Data of User a

We recognized the activities from the data gathered from the toilet, a dressing room, and the door-sensing side. It became easy to figure out whether to go out by sending the sensor working information as data. When we export the behavioral data log in a .csv format, caregivers can view it on web pages.

We got good feedback about this method for nursing care.

3.4. Knowledge obtained through this preliminary experiment

In Japanese usual houses, laying the power and ethernet cables was very difficult. Because we had to ensure the supply, we focused on the sensor's mounting position. Many commercial products are saying this is very simple. However, we experienced many problems, including when the client application unexpectedly stopped writing or the settings were changed, and the settings of the PC power management programs were doubled. Therefore, we reduced the cost by combining commercial products but we also found some disadvantages. Despite operating under the same situations, one out of the four client PC installations often stopped. We also adopted a standard blue ethernet cable for the elderly. Such bright and flashy sights are too

prominent for the elderly and the partially visibly impaired. The same thing stands out for the installed sensor units at the low positions from a silver aluminum ground. We also need to consider environmental changes in seniors with dementia. For example, wearable sensors are stressful for them. We should also pointed out that they have only low PC operation skills for browsing the data.

4. Wireless Sensor Network for Easy Installation

4.1. Reason for selection and feature of Zigbee

We selected the most suitable in the actual transmission standards because it should be wireless. The follows is the actual transmission standards, it is shown the future.

- Wireless LAN
 - The data traffic is over specification for the sensor network.
 - Much power consumption
 - Peer to peer communication
 - The unit cost is low for the widely available.
- Bluetooth
 - On some level of power consumption
 - Peer to peer communication
 - The unit cost is low for the widely available.
- ANT+
 - The data traffic is suitable for the sensor network.
 - Low power consumption
 - The unit cost is expensive for the fewer available.
- Zigbee
 - Data traffic is suitable for the sensor network.
 - Ultra low power consumption
 - A mesh network can be built by setting needless
 - The unit cost is low for the widely available.

We select Zigbee in consideration of the above.

The most unique points of Zigbee is that it is ultra-low power consumption and it can built easily for a mesh type network. Specifically, the behavior of 6 months is possible with only the battery type CR2032. When a communication command is taken out from the unit A to the unit B, you can select a single course in the conventional standard. But you can be data communication

using other courses even if you cannot use a course, in the mesh type network. It is shown in figure4 for a mesh network of Zigbee.

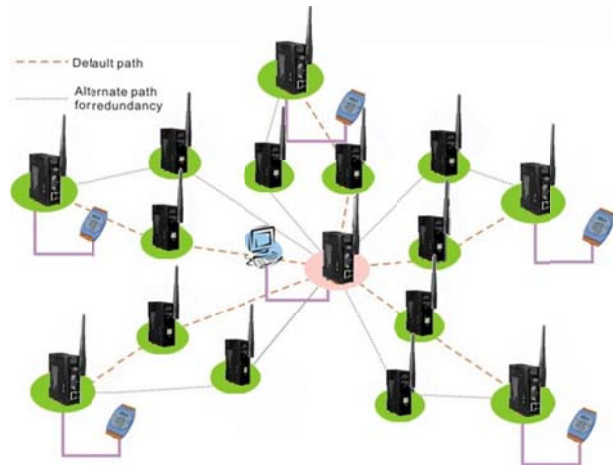


Figure 4: Mesh network of Zigbee

A mesh network of Zigbee. Installed Zigbee has a repeater function, so that you can operate as a repeater by setting Zigbee in the middle point in the case of short of communication distance. It is the optimal as a sensor network because you can use these features without special installation. Moreover some Zigbee have a program area alone, you can mount the sensor which easy calculation is attached although rich calculation can't be performed.

4.2. System outline

We have developed a sensor network that can be installed easily by anyone. We selected the most suitable transmission standards because it should be wireless. The most unique features of Zigbee are that it has ultra-low power consumption and it can be built easily for a mesh-type network. Specifically, Zigbee-based modules operate six months using only one type-CR2032 battery. When a communication command is sent from unit A to unit B, it can select a single route from A to B. Moreover, any route can be selected for data communication in the mesh network of Zigbee. Therefore, these strengths are leveraged to give the users the abilities to establish the

network by simply placing units appropriately and to view the sensing results on a web page.

This section gives an overview of the developed system (Figure 5) and describes how the web interface implements the following functions.

- Sensing time
- Device ID (Zigbee 64-bit address)
- Sensor battery voltage

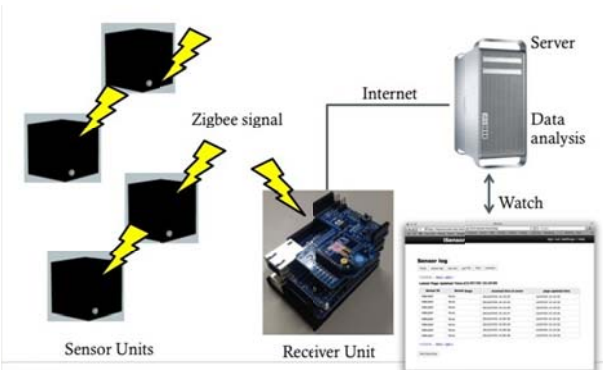


Figure 5: System configuration

We have also designed the system to have inconspicuous housing. Therefore, it was constructed as a small cube, with the sensor position offset from the cube's center to permit easy adjustment of the sensor after installation. The edge is 75 mm, depending on the size of the circuit board and battery, and we are targeting further miniaturization in the future. The housing and internal components are shown in Figure 6.

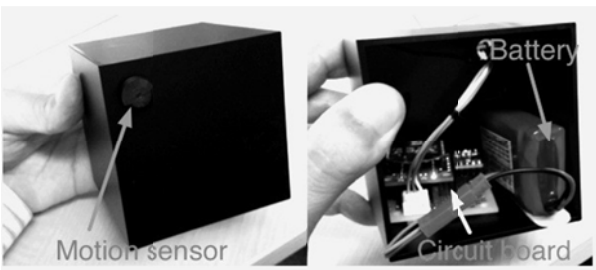


Figure 6: Housing and internal components

4.3. Web interface

Describes the web interface developed in this section. In a web page that implements the following functions.

- Sensing time
- Device ID(Zigbee 64 bit address)
- Sensor location
- Sensing information for each room

And that the server application development by Ruby, to update the information every 30 seconds. Currently have implemented the ability to monitor the battery voltage of the sensor unit.

We show web interface in figure 7.

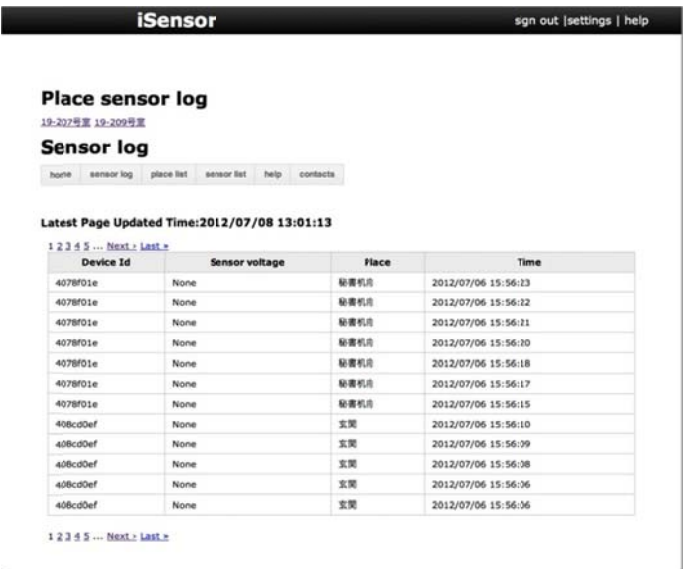


Figure 7: Web interface

4.4. Design of housing

We have also designed for inconspicuous housing. Therefore, it was small cube. Also offset to sensor position by cube center. Because can be adjust sensor by install surface. That angle and 75mm depending on the size of the circuit board and battery, and in the future has set a target of further miniaturization.

5. Anomaly Detection using SVM

Defining Anomalies

An important aspect of any anomaly detection technique is the nature of the anomaly under focus. Anomalies can be classified into three main categories[17]. First, if an individual data instance can be considered anomalous with respect to the rest of the data, then this instance is termed a *point anomaly*. This is the simplest type of anomaly, and it is the focus of a majority of the research on anomaly detection. Second, *Contextual Anomalies* exist when a data instance is anomalous in a specific context. The notion of context is determined by the structure of the dataset, and this has to be specified as a part of the problem formulation. Each data instance is defined using *Contextual attributes* and *Behavioral attributes*. The choice of applying a contextual anomaly detection technique is determined by the meaningfulness of the contextual anomalies in the target application domain. The third technique is based on *Collective Anomalies*. In this technique, a collection of related data instances is anomalous with respect to the entire data set. The individual data instances in a collective anomaly may not be anomalies in themselves, but their occurrence together in a collection is anomalous.

In this research we perform anomaly detection using a contextual anomaly technique that is common to the domain of elderly care and anomaly detection in a laboratory. This allows us to detect unusual activities in different times of a life. We evaluate anomaly detection techniques using real data obtained with a prototype sensor network, and the obtained sensing data are shown in Table 1.

Table 1: Example of actual sensing data

Device ID	Voltage	Sensing Time
40981d4f	3.8	2012-08-20 15:49:17 +0900
40981d4f	3.8	2012-08-20 15:49:31 +0900
40981d4f	3.8	2012-08-20 15:49:32 +0900
40981d4f	3.8	2012-08-20 15:49:36 +0900
40981d4f	3.8	2012-08-20 15:49:38 +0900
40981d4f	3.8	2012-08-20 15:49:45 +0900
40981d4f	3.8	2012-08-20 15:49:46 +0900
40981d4f	3.8	2012-08-20 15:49:47 +0900

Transforming data to SVM format

Support Vector Machines (SVMs) provide a useful technique for data classification. Although SVM is considered easier to use than Neural Networks, users not familiar with it often get unsatisfactory results at first[24]. SVM requires that each data instance be represented as a vector of real numbers. Hence, if there are categorical attributes, we first have to convert them into numeric data. Therefore, we converted the dataset as follows:

- Data label
- Hour of sensing data
- Minute of sensing data
- Second of sensing data

We converted each line in the above datasets, and did this, furthermore, at the time of the experiment. In addition, in the experiment performed on the training data, the training data were labeled "+1" if a positive example and "-1" if a negative example. Consequently, SVM recognized class as abnormal and normal classes.

Parameter Tuning for SVM

SVM is a useful technique for data classification, with successful applications in different fields such as bioinformatics, image segmentation, and data mining. A key problem with these methods is how to choose the optimal kernel and how to optimize its parameters in the learning process of SVM[25].

The objective of this study is to propose a genetic algorithm approach for parameter optimization to solve this problem. Using a grid search algorithm is a common technique for SVM parameter setting. However, this method suffers from various limitations such as being time-consuming and not performing well.

Dealing with imbalanced data

SVM has been extensively studied, and it has shown remarkable success in many applications. However, the success of SVM is very limited when it is applied to the

problem of learning from imbalanced datasets in which the negative instances heavily out-number the positive instances (e.g. in gene profiling and detecting credit card fraud) [26]. In this research, we investigate the factors behind this failure and explain why the common strategy of undersampling the training data may not be the best choice for SVM. We then propose an algorithm for overcoming this problem that is based on a variant of the SMOTE algorithm [27].

6. Bayesian Network based Method for Reducing Noise-data

6.1. Experimental environment

For the SVM used in anomaly detection proposed in this study, classification is performed to set the boundary-supervised learning. When a data instance is labeled so that it can be used with the SVM in Section 5, the sensor data actually obtained falls into only two classes: normal and abnormal classes. Included in the dataset are such factors as noise and faulty data due to sensing malfunctions of the sensor unit, which are caused

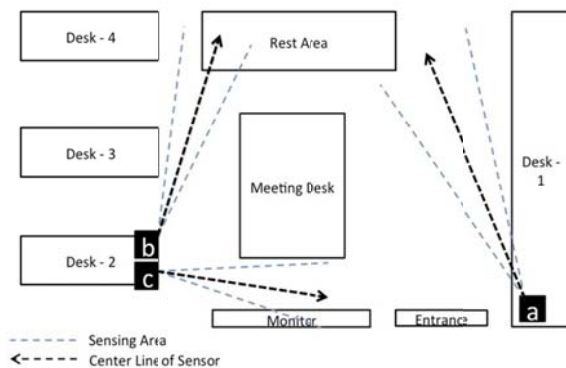


Figure 8: Room-A layout diagram of sensor units

by changes in the external environment. Here, noise is lowered but so is prediction accuracy, which in turn lowers measurement accuracy. Therefore, as a technique to improve prediction accuracy, we propose a method of noise removal using a Bayesian network.

While installing the sensor units in our room, we collected the sensed data for performing anomaly detection by sensing actual data. The sensor units have been positioned in accordance with the flow lines of students. Figure 8 shows the layout of Room-A.

Room-A is divided into two main areas, and the flow line is illustrated as the path toward the entrance from a desk. Sensors are arranged at right angles to the flow line, and since it is about 5 m, sensing of the entire laboratory is possible. For simplicity, the design of Room-A is similar to that of a typical business or household.

6.2. Node-causing settings

We set to nodes A, B and C the detection ranges of the sensors a, b and c in Room-A. As a result, we built a model of movement using four nodes (Outside, A, B and C). In considering moving and the outside, this model has made it possible to move to C from B and from A. However, one must go through the doorway to exit outside of A. When we set B to outside, it is impossible to move outside from C.

Figure 9 shows the flow line model using realistic data.

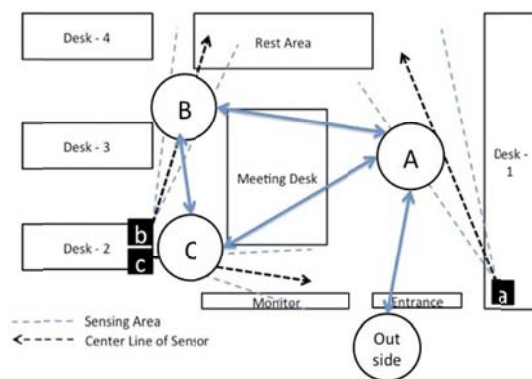


Figure 9: Flow model in Room-A

A Bayesian network is used to build a stochastic variable that causes the movement of the sensor in the detection range of sensor b from sensor a (e.g., AB, AC...).

Next, we describe how to set the probability of a random variable that will cause the node. We also set the

count probability of a sequence of sensor data acquired continuously within a certain time. In this research, this period of time is set to 1 minute. The transition probability is obtained by counting the sequence of sensor data. The results of counting the sensor sequence and the transition probability data using the prototype system are shown in Table2.

Table 2: Transition probability

Row of sensor data	Number of counts	Transition probability
aa	3836	0.4787818
ab	1421	0.177359
ac	2658	0.3317524
ad	97	0.0121068
ba	1393	0.0733312
bb	7500	0.39482
bc	9811	0.5164772
bd	292	0.0153717
ca	2696	0.0945169
cb	9774	0.3426588
cc	1568	0.5498878
cd	369	0.0129365
da	87	0.1147757
db	301	0.3970976
dc	370	0.4881266

We next convert from Table2 (Transition probability) to a matrix. Formula (1) shows the definition of the transition probability matrix.

$$P = \begin{pmatrix} p_{aa} & p_{ab} & p_{ac} & p_{ad} \\ p_{ba} & p_{bb} & p_{bc} & p_{bd} \\ p_{ca} & p_{cb} & p_{cc} & p_{cd} \\ 0 & p_{da} & p_{db} & p_{dc} \end{pmatrix} \quad (1) \quad \begin{matrix} \text{The} \\ \text{trans} \\ \text{ition} \\ \text{prob} \\ \text{abili} \end{matrix}$$

ty matrix calculated from sensor data is shown in Formula (2).

$$P = \begin{pmatrix} aa & ab & ac & ad \\ ba & bb & bc & bd \\ ca & cb & cc & cd \\ - & da & db & dc \end{pmatrix} \quad (2)$$

The stationary stochastic can be calculated by using its definition in Formula4.

$$\pi_1 = P(A), \pi_2 = P(B), \pi_3 = P(C), \quad (3)$$

$$\pi_4 = P(\Delta),$$

$$\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$$

$$\pi = \pi P, \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \quad (4)$$

The resulting values of stationary stochastic are shown Table 3.

Table 3: Stationary stochastic

π_1	0.138121243
π_2	0.331184645
π_3	0.504725888
π_4	0.025968225

We calculated the probability of each move by integrating the transition probabilities and where the steady-state probability was calculated to occur. Table 4 shows the calculated probability of occurrence of each movement. The probability of moving Bd, Cd using the generated noise model was set to an impossible move, as defined in Figure 9.

Table 4: Probabilities of cause node

P(DA)	0.0046527
P(AA)	0.0661299
P(AB)	0.0487832
P(AC)	0.0935272
P(BB)	0.1307583
P(BC)	0.3439981
P(CC)	0.2775426
P(NOISE)	0.0346079

6.2.1. Set of effect-nodes

We built a Bayesian network by a stochastic variable sequence of sensor data generated continuously within a certain time Next, we calculated the conditional probability by $P(X|Y), X \in \text{effect-node}, Y \in \text{cause-node}$.

We performed the sampling of a sensor sequenc by moving 100 times in the actual room. We the calculated the confidence factor of each node cause from the sampling results, and these confidence factors are shown in Table5.

A high degree of confidence was identified as the bold-type items in Table 5, where the variable was the probability of each resulting node being observed. The

sequence of sensor data, which is assumed to be caused by noise from the results, is bd ,cd ,d b,d . The SVM analysis is also assumed to be caused by noise from these results.

Next, the prediction accuracy was compared before and after removal of the sensor data.

Table 5: Confidence factors of caused nodes

	aa	ab	ac	ad	ba	bb	bc	bd	ca	cb	cc	cd	da	db	dc
P(NOISEjX)	0.021	0.021	0.02	0.021	0.021	0.02	0.02	0.554	0.02	0.02	0.02	0.448	0.021	0.665	0.643
P(dAjX)	0.008	0.003	0.003	0.427	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.427	0.003	0.003
P(AAjX)	0.742	0.04	0.039	0.305	0.04	0.039	0.038	0.039	0.039	0.038	0.038	0.039	0.305	0.039	0.039
P(ABjX)	0.116	1	0.029	0.127	1	0.066	0.028	0.059	0.029	0.028	0.028	0.029	0.127	0.065	0.029
P(ACjX)	0.221	0.057	1	0.243	0.057	0.055	0.054	0.056	1	0.054	0.092	0.101	0.243	0.056	0.122
P(BBjX)	0.08	0.08	0.079	0.08	0.08	0.677	0.076	0.237	0.079	0.076	0.077	0.078	0.08	0.079	0.079
P(BCjX)	0.225	0.226	0.224	0.226	0.226	0.465	1	0.42	0.224	1	0.349	0.38	0.226	0.461	0.453
P(CCjX)	0.177	0.178	0.177	0.178	0.178	0.174	0.17	0.176	0.177	0.17	0.829	0.434	0.178	0.176	0.176

7. Experimental Evaluations

7.1. Experimental setting

In this research, the evaluation index of the anomaly detection methods was obtained using the value of F and the precision and recall rates as well as the correct answers when classifying the data that have been labeled as class anomalies in the training data.

Calculated by Formula 5 as the correct rate, the value of F is then calculated by Formula (6) as the (Precision) compliance rate. After calculating the required Formula (7), the equation of Formula (8) is calculated using the values of each recall. The compliance rate represents the accuracy of the data instances that were classified as abnormal class, or that have been correctly classified as significantly abnormal class. Recall represents whether an instance can be completely classified as abnormal class based on the data anomalies that are included in the training data.

7.2. Evaluation of noise reduction by Bayesian network

A method of noise removal using a Bayesian network is proposed in this study. After labeling the data manually from the sensor system, the SMOTE algorithm was applied to data before noise reduction. This was done to learn the data for labeling. Then, after removing the sensor data that have been estimated as noise caused by the Bayesian

network, we manually labeled the data. The SMOTE algorithm [27] was then applied to the training data, followed by labeling.

$$\text{Accuracyrate} = \frac{\text{Judge as anomaly or normal data}}{\text{Classified data instances}} \quad (5)$$

$$\text{Precisionrate} = \frac{\text{Judge as anomaly data}}{\text{Classified anomaly data}} \quad (6)$$

$$\text{Recallrate} = \frac{\text{Judge as anomaly data}}{\text{Anomaly data in test instances}} \quad (7)$$

$$F\text{-measure} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

By implementing a five-fold cross-validation and parameter adjustment and then comparing the classification accuracy, we could show the effectiveness of the proposed method. Table 6 shows the class accuracy rate of abnormality due to the anomaly detection method used before reducing noise, as well as the precision rate, recall rate, and F-measure.

Table 6: Evaluation of each kernel function before noise reduction

Kernel	Accuracy rate	Precision rate	Recall rate	F-measure
Linear	0.983	0.983	0.984	0.983
Polynomial	0.991	0.998	0.984	0.991
RBF	0.992	0.993	0.992	0.992
Sigmoid	0.979	0.979	0.98	0.979

Table 7 shows the accuracy rate of the abnormal-class anomaly detection method by using data that have been subjected to noise reduction, compliance rate, recall rate, and F-measure.

Table 7: Evaluation of each kernel function after noise reduction

Kernel	Accuracy rate	Precision rate	Recall rate	F-measure
Linear	0.987	0.983	0.991	0.987
Polynomial	0.994	0.997	0.991	0.994
RBF	0.99	0.989	0.991	0.99
Sigmoid	0.984	0.978	0.99	0.984

Table 6 shows the evaluation results obtained before noise reduction and Table 7 shows those obtained after noise reduction.

As a result of noise removal, the kernel of prediction accuracy has been improved, aside from the RBF kernel. Prediction accuracy is also improved by noise removal, from these results. In addition, accuracy is worse for a learning model using the RBF kernel results from noise removal. This is because during noise removal, some accidentally erased data involved the support vector classification boundaries used for creating a high degree of accuracy. In order to prevent the erasure of data in the future, noise refinement will be increased from column to column, with two or three sensors of the random variable node result of the Bayesian network used when reducing noise. Moreover, to improve the accuracy of estimation, noise caused by making the node multi-level needs to be further reduced.

From the results of experiments carried out in this paper, we were able to evaluate anomaly detection methods by analyzing the sensor data using SVM. In this paper, data analysis applying the SMOTE abnormality-detection algorithm obtained better results in comparison with other kernel functions. In addition, prediction accuracy is improved by identifying the sequence of sensor

data that is estimated to be noise caused by the Bayesian network used to remove them. Our results showed that the proposed method is effective for noise removal.

8. Summary

In this paper, we proposed a method that could easily perform anomaly detection of a sensor network. We showed that the "abnormal activity in the context of a different time zone from the normal" could be detected by the proposed method. Next, we identified by experiment the sequence of datasets of sensor data that are estimated to be noise by the Bayesian network constructed, except for the sensor data before noise reduction. In evaluating classification using the port, the polynomial kernel was found to obtain the optimal result. This shows improvement in prediction accuracy by eliminating noise from the result. Further refinement can be achieved by increasing the sensor column by three sensors with two random variables of the node in future results of the Bayesian network. Furthermore, by using a multi-stage node to improve the accuracy of estimation noise, an even greater amount of noise can be removed.

Acknowledgements

This work is partially supported by the Funding Program for Next Generation World-Leading Researchers (NEXT Program) of the Japan Cabinet Office.

References

1. Takanobu Otsuka, Tatsunosuke Tsuboi, Takuya Yoshimura, and Takayuki Ito, "Prototyping and evaluation of a wireless sensor network system that aims at easy installation," JAWS (Joint Agent Workshop and Symposium) 2012.
2. SECOM CO.LTD. "SECOM HOME SECURITY," <http://www.secom.co.jp/homesecurity/plan/kodate/>
3. ALSOK CO. LTD. "HOME SECURITY SYSTEM," http://www.alsok.co.jp/person/hs_price.html
4. ZOUJIRUSHI CO. LTD "mimamori hot line" <http://www.mimamori.net/>

5. Shigeki Aoki, Masaki Onishi, Atsuhiko Kojima, Yasuhiro Sugahara, Kunio Fukunaga, "Recognition of a Solitude Senior's Behavioral Pattern Using Infrared Detector," The Institute of Electronics, Information and Communication Engineers, Technical Report No. 2001-50, 2002.
6. Katsunori Matsuoka, "Detection of unusual behavioral condition based on the long-term record of life behavior in a house," The Institute of Electronics, Information and Communication Engineers, Vol. 102, 2003.
7. Masahiro Furuya, Hajime Murakami, Wataru Miyamoto, "Detecting of an illness date based on usual activities of the single elderly person by using a few sensors," The Institute of Electronics, Information and Communication Engineers, Technical report No. 2002-125, 2003.
8. Hiroshi Akeyama, Takao Kawamura, Kazunori Sugahara, Takeshi Saitoh, Ryosuke Konishi, "Multifunctional Outlet with Web Functions," The Institute of Electronics, Information and Communication Engineers 7th of Forum of Information Technology, Vol. 25, 2008.
9. Shunji Yazaki, Toshio Matsunaga, Nobuhiro Tsukie, "Comfortable Life Supporting System for Elderly People Using Biosensing Technology and RFID," The Institute of Electrical Engineers of Japan, IIS, Industry System Group, 2009.
10. Hitoshi Tanaka, Yasushi Nakauchi, "Senior Citizen Monitoring System by Using Ubiquitous Sensors," Society of Mechanical Engineers, Journal No. 75-760, 2009.
11. Kiyoshi Itao, Hiroshi Hosaka, Ken Sasaki, Noriyoshi Yamauchi, Naoki Yahagi, Ryutaro Takahashi, Takashi Tajima, Satoshi Shimada, Yoshinori Shiote, Shiro Kanou, Hikaru Sato, Ikuko Urushibara, Naoki Asai, and Akio Sato, "A Health Information System using Wearable Sensors," Information Technology Promotion Agency, Next-generation Software Development Program, 2002.
12. Fumio Mizuno, Tomoaki Hayasaka, Yuichiro Yahagi, Hiromi Nakamura, Toshihiko Yoshida, Kenichi Tsuboi, Shigeo Wada, and Takami Yamaguchi, "Development of wearable computer for elderly home health care support" Welfare, Wellbeing, Life Support, Vol. 209, 2004.
13. Hiromichi Maki, Hidekuni Ogawa, Yoshiharu Yonezawa, Junichi Iwamoto, Ishio Ninomiya, Kouji Sada, "A Wearable Bioinformation Recording System," The Japanese Journal of Medical Instrumentation, Vol. 80, No. 4, 2010.
14. Hirokazu Seki and Yoichi Hori, "Detection of Abnormal Action Using Image Sequence for Monitoring System of Aged People," The Institute of Electronics, Information and Communication Engineers, Journal No. 122-2, 2000.
15. Shigeki Aoki, Yoshio Iwai, Masaaki Onishi, and Atsuhiko Kojima, "Recognition of Behavioral Pattern Using Position and Posture of Human Body," The Institute of Electronics, Information and Communication Engineers, Vol. 102, 2001.
16. Kenichi Nakagawa, Taro Sugihara, Hitoshi Koshiba, Ryoza Takatsuka, Naotaka Kato, and Susumu Kunifuji, "Development of Cooperative Care Support System for people with Dementia by Society Oriented Approach," Journal of Information Processing Society of Japan, Vol. 49, 2008.
17. Varun Chandola, Arindam Banerjee, and Vipin Kumar, "Anomaly Detection: A Survey," Technical Report, Department of Computer Science and Engineering, University of Minnesota, TR07-017, 2007.
18. D. Janakiram, V. Reddy, and A. Kumar, 2006. Outlier detection in wireless sensor networks using bayesian belief networks. In First International Conference on Communication System Software and Middleware. 16.
19. V. Kumar, 2005. Parallel and distributed computing for cybersecurity. Distributed Systems Online, IEEE 6, 10.
20. C. Spence, L. Parra, and P. Sajda, 2001. Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model. In Proceedings of the IEEE Workshop on Mathematical Methods in

- Biomedical Image Analysis. IEEE Computer Society, Washington, DC, USA, 3.
21. R. Fujimaki, T. Yairi, and K. Machida, 2005. An approach to spacecraft anomaly detection problem using kernel feature space. In Proceeding of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining. ACM Press, New York, NY, USA, 401410.
 22. W. Du, L. Fang, and N. Peng, 2006. Lad: localization anomaly detection for wireless sensor networks. J. Parallel Distrib. Comput. 66, 7, 874886.
 23. V. Chatzigiannakis, S. Papavassiliou, M. Grammatikou, and B. Maglaris, 2006. Hierarchical anomaly detection in distributed large-scale sensor networks. In ISCC 2006: Proceedings of the 11th IEEE Symposium on Computers and Communications. IEEE Computer Society, Washington, DC, USA, 761767.
 24. Chin-Wei Hsu, Chin-Chung Chang, and ChinJen Lin, "A Practical Guide to Support Vector Classification," <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf> (2010).
 25. Vladimir N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," 5th COLT, pp. 144-152 (1992).
 26. Rehan Akbani, Stephen Kwek, and Nathalie Japkowicz, "Applying Support Vector Machines to Imbalanced Datasets," Lecture Notes in Computer Science Volume 3201, 2004, pp. 39-50.
 27. Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," Journal of Artificial Intelligence Research 16 (2002) 321357.