Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing

RFID-Based Personalized Behavior Modeling

Hui-Huang Hsu[1], Zixue Cheng[2], Timothy K. Shih[3], and Chien-Chen Chen[1]
[1]*Department of Computer Science and Information Engineering, Tamkang University, Taiwan*
[2]*Department of Computer Science and Engineering, University of Aizu, Japan*
[3]*Department of Computer Science, National Taipei University of Education, Taiwan*
[1]*huihuanghsu@gmail.com*, [2]*z-cheng@u-aizu.ac.jp*, [3]*timothykshih@gmail.com*

Abstract

In this research, we aim at building an intelligent system that can detect abnormal behavior for the elderly at home. Deployment of RFID tags at home helps us collect the daily movement data of the elderly. The clustering technique is then used to build a personalized model of normal behavior based on these RFID data. After the model is built, any incoming datum outside the model can be seen as abnormal. In this paper, we present the design of the system architecture and show the preliminary results for data collection and preprocessing.

Keywords: *RFID, elderly care, ambient intelligence, machine learning, clustering analysis*

1. Introduction

Elderly care is an important issue in an aging society. There are many elderly people who live alone. Even though some elderly people live with their children, they are at home alone most of the time since young people need to work or go to school. This research would like to utilize information technology to unobtrusively detect abnormal behavior of the home-alone elderly according to their movement at home at day time. Abnormal behavior includes not only emergencies like pass-out and heart attack, but also others like "not eating," "not going to toilet regularly," and "lack of movement." Since everyone can have very different behavior at home, it is not possible to predefine the so-called abnormal behavior. Thus machine learning technology is needed to build a behavior model for each individual.

The RFID (Radio Frequency Identification) technology uses a reader to detect tags. When the reader detects the signals from the tags, RSSI values (0-255) that represent signal strength can be obtained. The RSSI value has a negative relationship with the distance of the tag from the reader [1]. In this research, we want to build a behavior model (viewed as normal) for an elderly person by a set of long-term collected RSSI data. We then can use this model to detect subsequent "abnormal" behavior of the person. The system architecture includes deploying active tags in the living environment, e.g., the living room, the dining room, the kitchen, the rest room, and the bed rooms. The reader is to be carried by an elderly person. The detected RSSI values are recorded following the movement of the person. The behavior model built by machine learning can be used to determine if the subsequent behavior is normal or not. Since only data of "normal" behavior are collected, the model (a classifier) can be built by only positive examples. This is an essential research issue in machine learning. We propose to use clustering techniques to deal with it.

This approach, different from computer-vision-based approaches, not just detects predefined events like if the elderly person falls, but finds the living patterns of the person. The machine learning technique can be used to learn all the patterns without defining all of them in advance. Also, it is not necessary to install cameras at home. This can relieve the concern of personal privacy issues [2]. Researches related to using RFID or other sensors for elderly care can also be found in [3], [4], and [5]. But all of them do not focus on behavior modeling as we do.

In the next section, we give an overview on needed technologies. The system architecture for behavior modeling by RFID is then presented in Section 3. Section 4 shows some preliminary results of data collection and preprocessing. Section 5 discusses future development and draws a brief conclusion.

978-0-7695-3737-5/09 $25.00 © 2009 IEEE
DOI 10.1109/UIC-ATC.2009.29

2. Technology Overview

An RFID system includes two parts: tags and readers. Digital information of an object, usually a serial number, is stored in a tag attached to the object. When a reader sends an electromagnetic signal to a tag, the tag responds the reader with the stored digital information. This information is then used to identify the object. There are basically two kinds of RFID tags: passive and active. Passive tags simply reflect the signal sent from the reader for the transmission of the stored information. They draw power from the reader and need no battery. On the other hand, active tags have their own transmitter and power source (battery). In-between the two, semi-passive tags operate their circuits on their own power, but communicate with the reader in the same way as passive tags. RFID tags are designed to operate in different frequencies. General speaking, low-frequency tags (usually passive tags) are used in a detecting distance shorter than 0.5 meter; high-frequency tags (usually active tags) are used for a wider range (several meters to one hundred meters). The RFID system has been widely used in many applications, e.g., tracking goods in a supply chain, tracking cows and pets, access control systems, and payment systems [6]. The RFID system has advantages over other automatic identification techniques, like bar codes and smart cards. To name a few: it does not need close contact; it does not need light; it can read many tags at the same time; and it can store additional information.

In our research, we chose the RFID system to record the movement of the elderly at home. Quite a many researches can be found in the literature for RFID tracking/positioning [7][8][9][10]. However, we do not need such complicated systems since exact location of the elderly person is not needed. Time sequences of RSSI values collected from active tags deploying in various corners at home are a sufficient representation of the movement of the elderly person who carries the reader. Other indoor positioning systems used wireless LAN [11], floor pressure/load sensors [12], and infrared-based small motion detectors (SMDs) [13]. Furthermore, video tracking systems can also be used to monitor indoor human behavior [14] or outdoor traffic [15]. But object tracking in video is a much more complex task. It can draw concern on personal privacy.

Another major issue of this research is the behavior modeling by only positive examples. The goal for such modeling is to detect abnormal behavior. In the literature, we have found applications of anomaly detection in information security [16][17], credit fraud detection [18], and document selection [19]. All of them used a machine learning approach. But in [17], one-class SVM (support vector machines) that finds the hyperplane between the positive examples and the origin is used. Such a classifier is biased and not suitable in our case. In [19], artificial negative examples were used. This is also impractical since negative examples are difficult to define and generate.

3. Behavior Modeling with RFID data

In our setting, a notebook PC, a wireless LAN access point (AP), a CF-card RFID reader (Fig. 1 (a)), a PDA with a CF-card slot (Fig. 1 (b)), and 10 active RFID tags with various sizes (Fig. 2) are used. The system architecture for data collection can be seen in Fig. 3. In the system, active tags are deployed in the environment. The person carries a PDA with the CF-card reader and wears a bracelet tag. The PDA transmits the received RSSI values from the tags to the notebook PC every second through the wireless AP. The collected data are then stored and preprocessed in the notebook PC for further analysis. We can of course collect the movement data in a reverse way, meaning, the person carries a tag and readers are deployed in the environment. But the cost would be much higher.

(a) **(b)**

Fig. 1. RFID Reader with PDA (a) CF-card RFID Reader (b) PDA with a CF-card slot

Fig. 2. Active RFID Tags

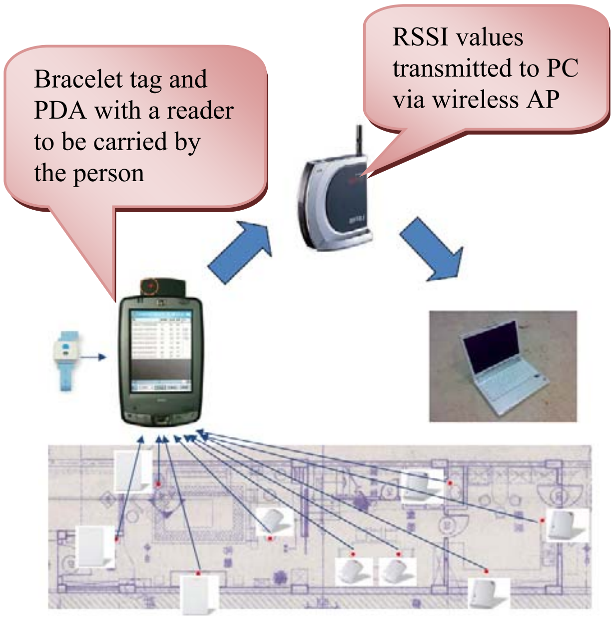


Fig. 3. System Architecture for Data Collection

Our aim is to collect the movement data of an elderly person alone at home. It will not be appropriate to use the same setting with more than one person in the same environment. If a person is not alone, such behavior analysis might not be necessary. Another important reason is that the RFID signal strength can be affected by other people who move in the same space. From our tests, even when nothing is moving in the environment, the RSSI values could be unstable sometimes.

In Fig. 4, we show the idea of behavior modeling through clustering [20]. After the RFID data sequences are properly smoothed and sampled as to be discussed in Section 4, the preprocessed data sequences are *windowed* by a sliding window. Two behavior (movement) models will be built. They are models for short-term behavior and long-term behavior. The window size is 5 minutes for short-term behavior and 5 hours for long-term behavior. The short-term model is used to detect anomalies like "falling down" (meaning unusually staying in a place for a few minutes). On the other hand, the long-term model is used to detect anomalies like "staying in the toilet for a period longer than usual" and "not opening the refrigerator for a few hours."

The windowed data sets for short-term modeling and long-term modeling are then clustered respectively through standard clustering technologies like K-Means. As shown in Fig. 4, we can build a hyper-elliptic boundary for each cluster. The number of clusters would be the number of different kinds of behavior for the person. In this example, we have four clusters in the two-dimensional space. In real cases we can have 10 to15 clusters in the 10-dimensional space (due to 10 RFID tags). The hyper-elliptic boundary can be found using singular value decomposition (SVD). The principal axes and their respective variance can be calculated and used to determine the hyper-elliptic boundary. Once the boundaries are determined, we can have a model to detect anomalies. In this example, point A is outside the boundary of either one of the four clusters. So it is viewed as an anomaly.

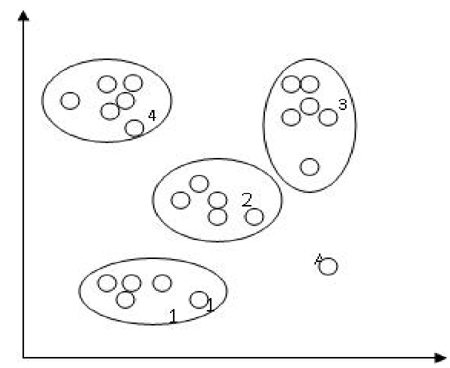


Fig. 4. Clustering for Behavior Modeling

4. Data Collection and Processing

The first floor (18.5m × 4.7m) of a house was chosen to deploy the tags. Nine of the tags are deployed (Fig. 5). The tags were placed at the entrance (#1), at two side tables in the living room (#2, #4), on top of the TV (#3) in the living room, at the two end of the dining table in the dining room (#5, #6), in the toilet (#7), in the refrigerator (#8), and near the sink in the kitchen (#9). The 10th tag on a bracelet was worn by the person who carries the reader. This tag can also measure and transmit body temperature of the person. This can ensure that the reader is with the person all the time. The following data are transmitted: the tag ID number, the RSSI value, the connectivity quality (LQI), the battery indicator (DI) and detected temperatures (only for Tag #10 that has temperature sensors). In our research, only the first two items are used to build the behavior models. A set of sample data collected and transmitted in one second is shown in Fig. 6.

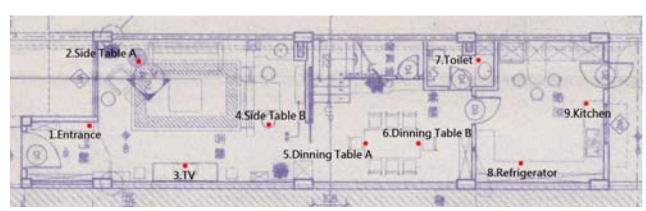


Fig. 5. Deployment of Tags in a Testing Environment

ID:0001000107520133,RSSI:133,LQI:219,DI:255,T1:-,T2:-
ID:0001000107520125,RSSI:132,LQI:73,DI:255,T1:-,T2:-
ID:0001000107520124,RSSI:111,LQI:107,DI:255,T1:-,T2:-
ID:0001000107462018,RSSI:0,LQI:0,DI:0,T1:-,T2:-
ID:0001000107462016,RSSI:103,LQI:213,DI:255,T1:-,T2:-
ID:0001000107462011,RSSI:0,LQI:0,DI:0,T1:-,T2:-
ID:0001000107291249,RSSI:0,LQI:0,DI:0,T1:-,T2:-
ID:0001000107291246,RSSI:0,LQI:0,DI:0,T1:-,T2:-
ID:0001000107291241,RSSI:0,LQI:0,DI:0,T1:-,T2:-
ID:0001000107503013,RSSI:150,LQI:227,DI:255,T1:29.25,T2:29.875

Fig. 6. Samples of Collected Data

To do a primary test, we collected the RSSI values for half an hour. During the test, the person stayed in the couch for 10 minutes, walked to the toilet in one minute, stayed in the toilet for 90 seconds, moved to the refrigerator in 15 seconds, stayed in front of the refrigerator for another 15 seconds, walked to the kitchen tag, stayed in front of the kitchen tag for 60 seconds, walked back to the couch in one minute, then repeated the whole process (Fig. 7). So the total time for this data collection was 30 minutes with 1800 data points. (The PDA transmits detected data every second.) The person moved in the test environment alone with no interference from other people. The CF-card reader was held in an upright position at all time to ensure that signals from all tags could be more accurately detected.

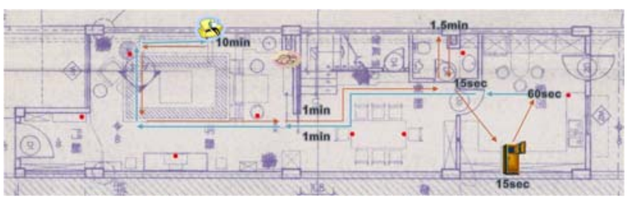


Fig. 7. Test Moving Route

The collected RSSI data sequences from five of the 10 tags are shown in Fig. 8. The five tags are *Side Table A*, *Side Table B*, *Dining Table B*, *Toilet*, and *Bracelet*. In Fig. 8 (a), the tag was not detected (with an RSSI value of zero) in the two squared periods because the person was going away from the living room. Fig. 8 (b) is similar to Fig. 8 (a) except that the RSSI values are lower and noisier when the person stayed in the living room. The noisy signal was caused by the furniture existing between Side Table B and the couch. In Fig. 8 (c), the signal was even noisier than Fig. 8 (b) when the person was in the couch. This is because that there is a division between the living room and the dining room. On the other hand, the RSSI values of the Dining Table B tag were higher when the person moved to the toilet and the kitchen. From Fig. 8 (d), it is very obvious to know at what time the person was in the toilet. In Fig. 8 (e), it shows that the bracelet tag was with the person during the test. From these signals, we are sure that the collected RSSI values can represent the movement of the person. The RSSI values from the tags change with the distances between the tags and the reader (the person).

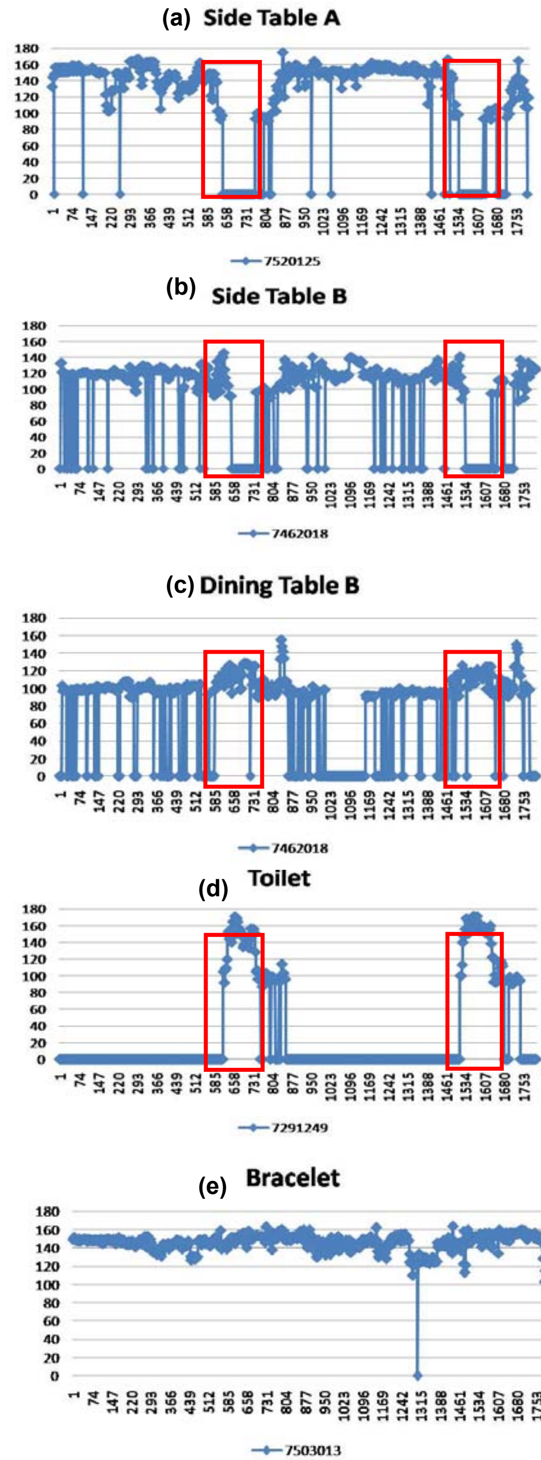
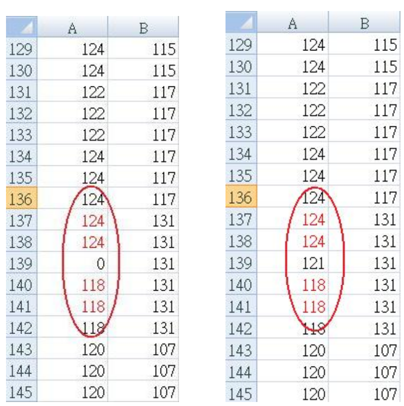


Fig. 8. Collected Data from the 30-minute Test Movement

The RSSI values can be unstable and noisy sometimes. This is why we need to preprocess the collected data before they are sent to the clustering process. From Fig. 8, we see noises in the collected signals. Most importantly, the signals dropped to zero sometimes even when the person was not moving. Also, furniture and divisions between rooms cause interference to the signals. Moreover, the reader is *directional*. Detected signals can be different when we point the reader in different directions. This is why the person had to keep the reader in an upright position to have more accurate data collection.

To provide the clustering algorithm a better set of data, we need to preprocess the collected data. First, the sudden zeros are smoothed out. Secondly, the data are sampled in two different ways for our modeling of short-term behavior and long-term behavior. In Fig. 9 (a), a sudden zero is shown and it can be smoothed out to 121 as in Fig. 9 (b). We use a voting mechanism to decide if a zero is a "sudden" zero. If the majority of the previous two samples and the subsequent two samples are not zero, it means that the current zero should not be zero. Thus it is smoothed out by averaging the RSSI values of the four neighbors. In this example, (124+124+118+118)/4=121. The smoothed signal of Fig. 8 (a) is shown in Fig. 10.



	A	B
129	124	115
130	124	115
131	122	117
132	122	117
133	122	117
134	124	117
135	124	117
136	124	117
137	124	131
138	124	131
139	0	131
140	118	131
141	118	131
142	118	131
143	120	107
144	120	107
145	120	107

	A	B
129	124	115
130	124	115
131	122	117
132	122	117
133	122	117
134	124	117
135	124	117
136	124	117
137	124	131
138	124	131
139	121	131
140	118	131
141	118	131
142	118	131
143	120	107
144	120	107
145	120	107

(a) (b)

Fig. 9. RSSI values with a sudden zero (a) before smoothing (b) after smoothing

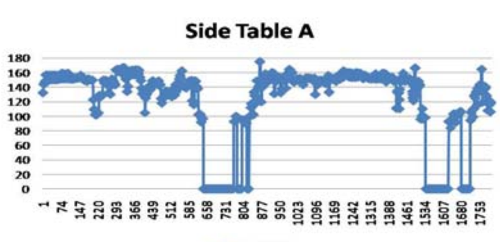
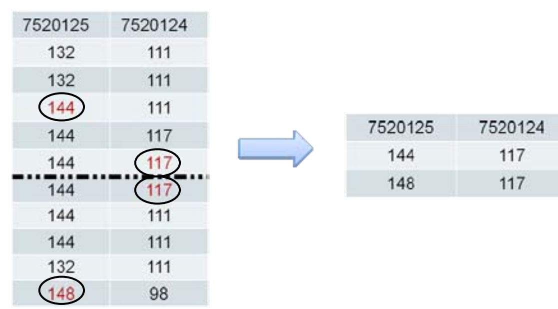


Fig. 10. Result with sudden zeros smoothed out for Side Table A

To model short-term and long-term behavior, we need to sample the data with different sampling intervals. This is to reduce the number of data points in a *windowed* data from the time sequences. Abnormal short-term behavior could be "falling down." On the other hand, abnormal long-term behavior could be "staying in the toilet for too long" or "not opening the refrigerator in five hours." So the short-term behavior modeling is based on windowed data of a 5-minute period and long–term behavior modeling is based on windowed data of a 5-hour period.

As mentioned above, data sampling is used to reduce the complexity of behavior modeling by clustering. However, uniform sampling seems not appropriate here. The person can move from one point to another then back in a few seconds. We might lose the information of this movement if the data are simply sampled by a fixed interval. Thus we believe that it is better to retain the maximal RSSI value in the sampled interval so that movement information can be best preserved. Fig. 11 shows an example for such a sampling process. The data are sampled every five seconds. 7520125 and 7520124 are the tag numbers.



7520125	7520124
132	111
132	111
144	111
144	117
144	117
144	117
144	111
144	111
132	111
148	98

7520125	7520124
144	117
148	117

Fig. 11. Data Sampling with the Maximum Retained

5. Discussion and Conclusion

We have proposed to use the clustering technique to build a human behavior model solely based on RFID data. Preliminary results show that the RFID data from active RFID tags can really grasp human movement in an environment. Next, we will develop a real-time system to collect data for a longer period of time, say, one week, build a model by clustering, and test the model with designed abnormal behavior data. If the model successfully detects the abnormal behavior, we will then seek elderly volunteers to further check the system in their daily life.

Furthermore, in this research, a person's behavior is only represented by his/her movement in the environment. We also plan to attach passive tags to objects that are normally used in daily life. With these data from passive tags, we will be able to understand a person's behavior more precisely.

References

[1] Crepaldi, R,, Harris, A.F., Kooper, R., Kravets, R., Maselli, G., Petrioli, C., Zorzi, "Managing Heterogeneous Sensors and Actuators in Ubiquitous Computing Environments," *Proc. International Conference on Mobile Computing and Networking*, Montreal, Quebec, Canada, 2007, pp. 35--42.

[2] Xiao, Y., Shen, X., Sun, B., Cai, L., "Security and Privacy in RFID and Applications in Telemedicine," *IEEE Communications Magazine*, 2006, vol. 44, Issue 4, pp. 64--72.

[3] Ho, L., Moh, M., Walker, Z., Hamada, T., Su, C.-F., "A Prototype on RFID and Sensor Networks for Elder Healthcare," *Proc. ACM SIGCOMM workshop on Experimental Approaches to Wireless Network Design and Analysis*, Philadelphia, Pennsylvania, USA, 2005, pp. 70--75.

[4] Lee, B., Kim, H., "A Design of Context aware Smart Home Safety Management Using Networked RFID and Sensor," *Proc. First IFIP WG6.2 Home Networking Conference*, Paris, France, 2007, pp. 215--224.

[5] Jih, W.-R., Hsu, J.Y., Tsai, T.-M., "Context-Aware Service Integration for Elderly Care in a Smart Environment," *Proc. AAAI Workshop on Modeling and Retrieval of Context Retrieval of Context*, The AAAI Press, Menlo Park, California, pp. 44--48.

[6] RFID Journal, http://www.rfidjournal.com .

[7] Jiang, W., Yu, D., Ma, Y., "A Tracking Algorithm in RFID Reader Network," *Proc. Japan-China Joint Workshop on Frontier of Computer Science and Technology*, 2006, pp. 164--171.

[8] Aydin, E., Oktem, R., Dincer, Z. Akbulut, I.K., "Study of an RFID Based Moving Object Tracking System," *RFID Eurasia*, Istanbul, 2007, pp. 1--5.

[9] Sangwan, R.S., Qiu, R.G., "Using RFID Tags for Tracking Patients, Charts and Medical Equipment within an Integrated Health Delivery Network", *Proc. 2005 IEEE Networking, Sensing and Control,* 2005, pp. 1070--1074.

[10] Jia, S., Sheng, J., Chugo, D., Takase, K., "Obstacle Recognition for a Mobile Robot Indoor Environments using RFID and Stereo Vision," *Proc. International Conference on Mechatronics and Automation*, 2007, pp. 2789--2794.

[11] Prasithsangaree, P., Krishnamurthy, P., Chrysanthis, P.K., "On Indoor Position Location with Wireless LANs," *Proc. 13th IEEE Int'l Symp. On Personal, Indoor, & Mobile Radio Communications*, Lisbon, Portugal, 2002, vol. 2, pp. 720--724.

[12] Liau, W.H., *Inhabitant Tracking and Service Provision in an Intelligent e-Home via Floor Load Sensors*, Master's Thesis, Department of Computer Science and Information Engineering, Nation Taiwan University, Taiwan, 2005.

[13] Hara, K., Omori, T, Ueno, R., "Detection of Unusual Human Behavior in Intelligent House," *Proc. 12th IEEE Workshop on Neural Networks for Signal Processing*, 2002, pp. 697--706.

[14] Chung, P.-C., Liu, C.-D., "A Daily Behavior Enabled Hidden Markov Model for Human Behavior Understanding," *Pattern Recognition*, Elsevier, vol. 40, pp. 1572--1580.

[15] Lin, C.P., Tai, J.C., Song, K.T., "Traffic Monitoring based on Real-Time Image Tracking," *Proc. IEEE International Conference on Robotics and Automation*, 2003, vol. 2, pp. 2091-- 2096.

[16] Lin, C.-H., Liu, J.-C., Ho, C.-H., "Anomaly Detection Using LibSVM Training Tools," *Proc. International Conference on Information Security and Assurance*, 2008, pp. 166--171.

[17] Wang, Y., Wong, J., Miner, A., "Anomaly Intrusion Detection Using One Class SVM," *Proc. Fifth Annual IEEE SMC Information Assurance Workshop*, 2004, pp. 358--364.

[18] Ghosh, A.K., Schwartzbard, A., "A Study in Using Neural Network for Anomaly and Misuse Detection," *Proc. 8th Conference on USENIX Security Symposium*, Washington, D.C., 1999, vol. 8, pp. 12.

[19] Zizka, J., Hroza, J., Pouliquen, B., Ignat, C., Steinberger, R., "The Selection of Electronic Text Documents Supported by Only Positive Examples," *JADT*, 2006.

[20] Mitra, P., Murthy, C.A., Pal, S.K., "Unsupervised Feature Selection Using Feature Similarity," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002, vol. 24, no. 3, pp. 301--312.